2023 • 27 • 6

Print ISSN 2500-0462
Online ISSN 2500-3259

Сетевое издание

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.
Периодичность 8 выпусков в год
DOI 10.18699/VJGB-23-65

Учредители

Сибирское отделение Российской академии наук

Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук»

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

2023 • 27 • 6

Print ISSN 2500-0462
Online ISSN 2500-3259

Online edition

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

Founded in 1997
Published 8 times annually
DOI 10.18699/VJGB-23-65

Founders

Siberian Branch of the Russian Academy of Sciences

Federal Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences

The Vavilov Society of Geneticists and Breeders

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

СОДЕРЖАНИЕ • 2023 • 27 • 6

Экологическая генетика

Тест-системы и вакцинопрофилактика

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2023 • 27 • 6

Ecological genetics

Test kits and preventive vaccination

Вавиловский журнал генетики и селекции. 2023;27(6):545-552
DOI 10.18699/VJGB-23-66

Original Russian text https://vavilovj-icg.ru/

Development and characterization of a line with substitution of chromosome 4B of wheat *Triticum aestivum* L. on chromosome 4H^{mar} of wild barley *Hordeum marinum* ssp. *gussoneanum* (4*x*)

L.A. Pershina[1, 2] ✉, N.V. Trubacheeva[1, 2], V.K. Shumny[1], E.D. Badaeva[3]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Kurchatov Genomics Center of ICG SB RAS, Novosibirsk, Russia
[3] Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow, Russia
✉ pershina@bionet.nsc.ru

Abstract. Introgressive hybridization is the main method of broadening the genetic diversity of bread wheat. Wild barley *Hordeum marinum* ssp. *gussoneanum* Hudson (2*n* = 4*x* = 28) has useful agronomical traits, such as high resistance to stress factors, that could be a potential source of new genes for bread wheat improvement. This study aimed to evaluate the possibility of introgression of *H. marinum* chromosomes into the genome of bread wheat using an incomplete amphiploid *H. marinum* ssp. *gussoneanum* (4*x*)–*T. aestivum* (Pyrotrix 28) (2*n* = 54) carrying the cytoplasm of wild barley. For this purpose, we crossed the line of bread wheat variety Pyrotrix 28 with an incomplete amphiploid, and then selected cytogenetically stable 42-chromosome plants with a high level of fertility in hybrid progeny. Genomic *in situ* hybridization (GISH) revealed a pair of *H. marinum* chromosomes in the genome of these plants. C-banding analysis confirmed that bread wheat chromosome 4B was replaced by wild barley chromosome 4H^{mar}. SSR markers *Xgwm368* and *Xgwm6* confirmed the absence of chromosome 4B, and EST markers *BAWU808* and *BAW112* identified chromosome 4H^{mar} in the genome of the isolated disomic wheat-barley substitution line. The study of this line showed that the substitution of chromosome 4B with chromosome 4H^{mar} resulted in a change of some morphological traits. It included intense anthocyanin coleoptile coloration, specific for *H. marinum*, as well as a lack of purple coloration of the ears in the leaf sheath, specific for Pyrotrix 28. Line 4H^{mar}(4B) showed increased performance for several traits, including plant height, number of spikes and tillers per plant, spikelet and grain number in the main spike, grain number per plant, but it had decreased values of 1000-grain weight compared to wheat. Cytogenetic stability and fertility of line 4H^{mar}(4B) indicated a high compensation ability of barley 4H^{mar} for wheat chromosome 4B and confirmed their homeology.
Key words: *Hordeum marinum* ssp. *gussoneanum*; bread wheat; wheat-barley substitution line 4H^{mar}(4B).

For citation: Pershina L.A., Trubacheeva N.V., Shumny V.K., Badaeva E.D. Development and characterization of a line with substitution of chromosome 4B of wheat *Triticum aestivum* L. on chromosome 4H^{mar} of wild barley *Hordeum marinum* ssp. *gussoneanum* (4*x*). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6): 545-552. DOI 10.18699/VJGB-23-66

Получение и изучение линии с замещением хромосомы 4B пшеницы *Triticum aestivum* L. на хромосому 4H^{mar} дикого ячменя *Hordeum marinum* ssp. *gussoneanum* (4*x*)

Л.А. Першина[1, 2] ✉, Н.В. Трубачеева[1, 2], В.К. Шумный[1], Е.Д. Бадаева[3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия
[3] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
✉ pershina@bionet.nsc.ru

Аннотация. Интрогрессивная гибридизация является основным методом расширения генетического разнообразия мягкой пшеницы. В качестве источника новых генов для мягкой пшеницы может служить дикий ячмень *Hordeum marinum* ssp. *gussoneanum* Hudson (2*n* = 4*x* = 28), который характеризуется высокой устойчивостью к стрессовым факторам. Настоящая работа посвящена изучению возможности использования неполного амфиплоида *H. marinum* ssp. *gussoneanum* (4*x*)–*T. aestivum* (Пиротрикс 28) (2*n* = 54), носителя цитоплазмы дикого ячменя, в качестве источника хромосом *H. marinum* для их интрогрессии в геном мягкой пшеницы. С этой целью получены гибриды между линией сорта мягкой пшеницы Пиротрикс 28 (далее П28) и неполным амфиплоидом,

а затем среди потомков гибридов проведен отбор цитогенетически стабильных 42-хромосомных растений с высоким уровнем фертильности. С использованием GISH-анализа обнаружено наличие пары хромосом *H. marinum* в геноме этих растений. По результатам С-окрашивания хромосом установлено, что у этой гибридной линии произошло замещение хромосомы 4В мягкой пшеницы на хромосому дикого ячменя 4H^{mar}. С помощью хромосом-специфичных SSR-маркеров *Xgwm368* и *Xgwm6* подтверждено отсутствие хромосомы 4В мягкой пшеницы, а с применением EST-маркеров *BAWU808* и *BAW112* – наличие хромосомы 4H^{mar} в геноме выделенной дисомной пшенично-ячменной линии. Изучение этой линии показало, что замещение хромосомы 4В мягкой пшеницы на хромосому дикого ячменя 4H^{mar} привело к изменению ряда признаков: сильно выраженной антоциановой окраске колеоптиле, характерной для дикого ячменя *H. marinum*, а также отсутствию пурпурной окраски ушек у основания листьев, которая проявляется у линии сорта пшеницы П28. По высоте растений, числу стеблей и колосьев в растении, числу колосков и зерен в главном колосе, а также числу зерен в растении линия 4H^{mar}(4B) превосходит родительскую линию П28, а по массе 1000 зерен ей уступает. Цитогенетическая стабильность и фертильность линии 4H^{mar}(4B) указывают на высокую компенсационную способность хромосомы 4H^{mar} ячменя по отношению к хромосоме 4В мягкой пшеницы и гомеологию между этими хромосомами.
Ключевые слова: *Hordeum marinum* ssp. *gussoneanum*; мягкая пшеница; пшенично-ячменная замещенная линия 4H^{mar}(4B).

Introduction

Introgressive hybridization is the main method of broadening the genetic diversity of bread wheat. Identification of new genetic resources for agronomically important traits, such as biotic and abiotic stress resistance, bread-making quality, increased content of microelements in grain and others is an important task (Molnár-Láng, Linc, 2015; Hao et al., 2020). The species of the genus *Hordeum* L., including annual and perennial barley grasses, as well as grain-type species, are considered a potential source to introgress genes for these traits (Garthwaite et al., 2005; Rubiales, Moral, 2011).

The progamous and embryonic incompatibility is strongly displayed in wheat × barley hybridizations and intergeneric hybrids were produced after the technique to overcome these difficulties had been developed (Kruse, 1973). To date, substitution, addition and translocation lines were developed by hybridization between bread wheat and *H. vulgare* (Molnár-Láng et al., 2014), *H. spontaneum* (Taketa, Takeda, 2001), *H. chilense* (Rey et al., 2015), *H. californicum* (Fang et al., 2014). Alloplasmic lines (*H. vulgare*)-*T. aestivum* were developed and used in breeding to create promising forms and cultivars of wheat (Pershina et al., 2018). In addition, a new grain crop, tritordeum, was produced from hybridization between durum wheat (*Triticum durum*) and a wild barley, *H. chilense* (Martín et al., 1999; Martín, 2017).

The ability to cross with wheat is also shown by two subspecies of the sea barley complex *H. marinum*: herbaceous annuals *H. marinum* Hudson ssp. *marinum* (2*x*) and *H. marinum* Hudson ssp. *gussoneanum* (Parl.) Thell. (2*x*, 4*x*) (Pershina et al., 2004; Islam et al., 2007). Due to their high resistance to salinity and waterlogging, these species are able to grow in saline meadows and marshes along sea coasts (Garthwaite et al., 2005; Islam et al., 2007). At the same time, salinity tolerance in *H. marinum* was estimated to be higher than in other species of Triticeae (Garthwaite et al., 2005). In addition, *H. marinum* ssp. *gussoneanum* (=*H. geniculatum* All.) (2*n* = 28) is resistant to drought and sudden temperature changes (Kobylyanskiy, 1967), and samples with a high protein content in seeds were also isolated (Pershina et al., 2009).

Amphiploids carrying wild barley cytoplasm were produced from hybrids between *H. marinum* ssp. *gussoneanum* (4*x*) and *T. aestivum* (Pershina et al., 2004) and between *H. marinum* (2*x*) and *T. aestivum* (Islam et al., 2007). The salt- and waterlogging tolerances of amphiploids were intermediate to those of their parents (Islam et al., 2007; Malik et al, 2009) and reduced grain yield was observed due to the negative influence of the *H. marinum* cytoplasm. To eliminate this cytoplasm effect, it is necessary to produce euplasmic hybrid genotypes, in which wheat is the maternal parent. However, the use of *H. marinum* ssp. *gussoneanum* as a pollinator when crossing with wheat is extremely difficult due to the limited amount of pollen in the small flowers of *H. marinum.*

This paper presents the results of using the incomplete amphiploid *H. marinum* ssp. *gussoneanum–T. aestivum* (2*n* = 54) with wild barley cytoplasm as a source of *H. marinum* chromosomes for their introgression into the bread wheat genome. The euplasmic wheat-barley disomic substitution line 4H^{mar}(4B) isolated among the progenies of these hybrids was also studied.

Materials and methods

Plant material. The line formed from one plant of the wheat variety Pyrotrix 28 (maternal parent, P28) was crossed with individual plants of an incomplete amphiploid L-503 *H. marinum* ssp. *gussoneanum–T. aestivum* (P28) (2*n* = 54). The incomplete amphiploid isolated from the progeny of a cytogenetically unstable amphiploid *H. marinum* ssp. *gussoneanum–T. aestivum* (2*n* = 68–70) carried 42 wheat chromosomes and 12 *H. marinum* chromosomes, with the exception of 5H^{mar} (Trubacheeva et al., 2019).

Plants were grown from seeds set in two hybrid combinations (P28 × 503 plant (p)5) and (P28 × 503 plant (p)10), and their progenies were studied in F_2–F_6 generations. Each generation was formed from the seeds of the most productive plant of the previous generation. Plants were characterized by seed set in the main spike, dividing them according to the level of fertility into groups: CS – completely sterile (no seeds); PS – partially sterile (1–9 seeds); PF – partially fertile (10–19 seeds); F – fertile (20–30 seeds); FF – fully fertile (more than 30 seeds). Starting from F_2, the number of chromosomes was counted in plants. The chromosome composition in metaphase I (MI) was determined in euploid

($2n$ = 42) hybrid plants F_5 (P28 × 503p10). Cytogenetically stable euploid plants were studied to identify *H. marinum* chromosomes. Hybridization and the study of the hybrid progeny were carried out in a hydroponic greenhouse.

The isolated wheat-barley substitution line 4H^{mar}(4B) was studied in the field. The plants were grown on plots 1 m wide, 20 plants per row, with a distance of 25 cm between rows. The control was the parental line P28. Each genotype was grown in three rows arranged in a randomized order. The lines were evaluated according to morphological traits. During harvesting, plant height, tiller number, number of productive spikes, main spike length, spikelet number per main spike, grain number per main spike and per plant were estimated. Statistical analyses were conducted in Microsoft Excel 2007. Single-factor analysis and the calculation of the least significant difference were used (Dospekhov, 1985).

Cytological analysis, genomic *in situ* hybridization and C-banding. Analysis of chromosome number was performed according to the standard Felgen preparation method using root tips of plants grown in a hydroponic greenhouse. The MI chromosome configuration was examined in pollen mother cells (PMCs) using the 2 % acetocarmine smear method. GISH was performed according to previously published protocols (Trubacheeva et al., 2019). Total genomic *H. marinum* ssp. *gussoneanum* DNA was labelled by nick translation with biotin-16-dUTP and used as a probe. The signals were observed and captured using AxioImagerM1 (Zeiss, Germany) fluorescence microscope. The work was performed at the Collective Center for Microscopic Analysis of Biological Objects (ICG SB RAS, Novosibirsk, Russia). C-banding was performed according to a previously published protocol (Badaeva et al., 1994).

DNA marker analysis. Two expressed sequence tag (EST-PCR) markers, *BAWU808* and *BAW112* (Trubacheeva et al., 2019), were used to identify chromosome 4H^{mar} of wild barley in the genome of the isolated line. The absence of wheat chromosome 4B in the substitution line 4H^{mar}(4B) was confirmed using chromosome-specific SSR markers *Xgwm368* and *Xgwm6* (Röder et al., 1998). The primer sequences were reported in (Trubacheeva et al., 2019. Amplifications were carried out using a Bio-Rad T-100 Thermal Cycler PCR. Amplification products were resolved in 2 % w/v agarose gels and visualized with ethidium bromide. Gel images were captured using a gel documentation system Gel Doc XR+ (Bio-Rad, USA).

Results

Identification of euploid plants among the progeny of hybrids (P28 × L-503)

As a result of crossing the line P28 with incomplete amphiploid L-503 in four hybrid combinations, shriveled seeds set with a frequency of 3.1 to 6.6 %, from which F_1 plants were grown (Table 1).

The shriveled seeds set in F_1 plants germinated in the hybrid (P28 × 503p5) with a frequency of 13.9 %, and in the hybrid (P28 × 503p10) with a frequency of 55.5 %. Starting from the second generation, the fertility level was studied in the progeny of these hybrids, estimated by the number of seed set per main spike, and the number of chromosomes was analyzed (Table 2).

In the second generation of the hybrid combination (P28 × 503p5), partially fertile (33 %) plants ($2n$ = 46), as well as sterile (45 %) and partially sterile (22 %) plants were observed, among which 46- and 48-chromosome plants were identified. The third generation, derived from a partially fertile plant ($2n$ = 46), also consisted of sterile (31 %), partially sterile (44 %) and partially fertile (25 %) plants. The most productive in F_3 was a 44-chromosome plant, among the progeny of which along with sterile (18 %), partially sterile (41 %) and partially fertile (27 %), fertile (14 %) plants were also identified in F_4. Among the progeny of the fourth generation, 42- and 44-chromosome plants were found. In F_5 and F_6 generations derived from fertile 42-chromosome plants, in addition to euploids ($2n$ = 42), aneuploids ($2n$ = 41; $2n$ = 43) were also identified. At the same time, segregation into plants with different levels of fertility, including sterile ones, was still observed (see Table 2). Fully fertile plants were obtained only in F_6 (6 %).

When selecting for productivity, the hybrid combination (P28 × 503p10) had a more accelerated formation of 42-chromosomal cytotypes with a high level of fertility. Fertile 46-chromosome plants (20 %) were already identified in F_2 along with sterile (13 %), partially sterile (27 %) and partially fertile (49 %) plants. The remaining groups of plants were represented by cytotypes with chromosome numbers $2n$ = 46, 46 + t, 46 + 2t.The progenies of the 46-chromosome plant in F_3 were sterile (6 %), partially sterile (38 %), partially fertile (25 %), and 42-chromosome fertile (31 %) plants. In F_3, in addition to euploids, aneuploids were found, including those with telocentric chromosomes (($2n$ = 43, 43 + t, 44 + t, 44 + 2t).

Among 26 F_4 plants derived from a 42-chromosome plant of the third generation, fertile plants prevailed (62 %), including plants with full fertility (11 %). In total, in F_4 of (P28 × 503p10), completely sterile (8 %) plants and plants with a low level of fertility (partially sterile – 11 % and partially fertile – 8 %) only accounted for less than one third. The remaining plants were fertile (62 %) and fully fertile (11 %). In F_4 cytotypes with $2n$ = 41, 42, 42 + t were identified among the studied plants. As a result of the selection of the most productive 42-chromosome plants, the fifth and sixth generations were represented only by fertile plants (in F_5 – 70 %, and in F_6 – 16 %) and fully fertile plants (in F_5 – 30 %, and in F_6 – 84 %) (see Table 2).

In order to select meiotically stable 42-chromosome plants for further work, three lines derived from individual plants of the P28 × 503p10 hybrid from F_6 were characterized by chromosome configuration at the first meiotic metaphase (MI). Study of meiosis in pollen mother cells (PMC) has shown a high cytological stability of euploid lines. All the studied plants of the line (P28 × 503p10) F_6 p1 and the major part of the plants of the line (P28 × 503p10) F_6 p2 (79 %) and the line (P28 × 503p10) F_6 p3 (89 %) had 21″ bivalents (Table 3). The identified violations are associated with the presence of univalents (20″ + 2′), including those involving telocentric chromosomes (21″ + 2t′).

Table 1. Seed set per spike, number of grains in F_1 plants and their germination in hybrid combination (P28 × L-503)

Hybrid combination	Number and frequency* (%) of hybrid grains in a pollinated spike	Number of		Germination of grains, %
		F_1 plants studied	grains in F_1 plant	
P28 × 503p5	1 (3.3)	1	129	13.9
P28 × 503p9	1 (3.1)	1	2	0
P28 × 503p10	2 (6.6)	2	18	55.5
P28 × 503p22	1 (4.0)	1	0	–

* From the number of pollinated flowers.

Table 2. Fertility level and number of chromosomes in F_1–F_6 of the hybrid combination *T. aestivum* (P28) (2*n* = 42) × (*H. marinum* ssp. *gussoneanum* × P28) (2*n* = 54)

Generation	Number of plants analysed	Number and frequency (%) of plants					Number of chromosomes
		FS 0	PS (1–9)#	PF (10–19)#	F (20–30)#	FF (more 30)#	
P28 × 503p5							
F_2	18	8 (45 %)	4 (22 %)	6 (33 %)	0	0	46*, 48
F_3	16	5 (31 %)	7 (44 %)	4 (25 %)	0	0	44*, 46
F_4	22	4 (18 %)	9 (41 %)	6 (27 %)	3 (14 %)	0	42*, 44
F_5	28	5 (18 %)	6 (21 %)	8 (29 %)	9 (32 %)	0	41, 42*, 43
F_6	32	2 (6 %)	4 (13 %)	10 (31 %)	14 (44 %)	2 (6 %)	41, 42, 43
P28 × 503p10							
F_2	15	2 (13 %)	4 (27 %)	6 (40 %)	3 (20 %)	0	46*, 46 + t, 46 + 2t
F_3	16	1 (6 %)	6 (38 %)	4 (25 %)	5 (31 %)	0	42*, 43, 43 + t, 44 + t, 44 + 2t
F_4	26	2 (8 %)	3 (11 %)	2 (8 %)	16 (62 %)	3 (11 %)	41, 42*, 42 + t
F_5	30	0	0	0	21 (70 %)	9 (30 %)	42*, 42 + t, 42 + 2t
F_6	38	0	0	0	6 (16 %)	32 (84 %)	42, 42 + t

Number of grains in the main spike; * the most productive cytotypes used to form the next generation.

Table 3. Characteristics of meiosis in lines derived from the hybrid combination (P28 × 503p10) F_6

Line	Number of plants examined	Number of plants and chromosome configuration at MI		
		21″	20″ + 2′	21″ + t″
(P28 × 503p10) F_6 p1	18	18 (100 %)	–	–
(P28 × 503p10) F_6 p2	19	15 (79 %)	4 (21 %)	–
(P28 × 503p10) F_6 p3	18	16 (89 %)	–	2 (12 %)

Identification of the wheat–barley disomic substitution line 4H^{mar}(4B)

At the next stage, plants with a 21″ chromosome configuration were included in the work. Using GISH analysis, it was shown that they carry a pair of *H. marinum* chromosomes (Fig. 1).

This result indicates that during self-pollination of the progeny of the hybrid (P28 × 503p10) and the production of euploids, a pair of *H. marinum* chromosomes introgressed into the genome of bread wheat. C-banding identified a pair of chromosomes 4H^{mar} and the absence of a pair of chromosomes 4B in the isolated line indicating that it was disomic for the substitution 4H^{mar} (4B) (Fig. 2, *a*). A telocentric chromosome for the arm of chromosome 4H was identified in a line with 2*n* = 40 + t chromosomes (see Fig. 2, *b*).

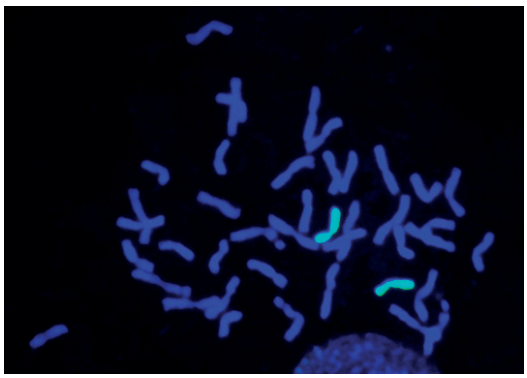

Fig. 1. GISH image of a wheat–barley substitution line. Barley chromosomes are in green, wheat chromosomes are in blue.

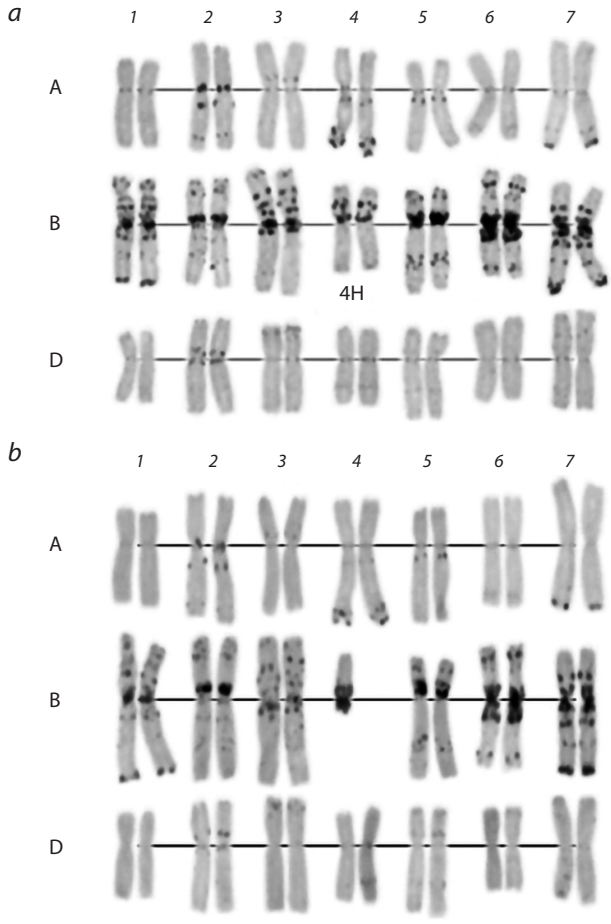

Fig. 2. C-banding of the 4H^{mar}(4B) (2*n* = 42) (*a*) and 4H^{mar} L (2*n* = 40 + t) (*b*) substitution lines.

Molecular analysis confirmed the presence of chromosome 4H^{mar} and the absence of chromosome 4B in the genome of disomic wheat-barley lines isolated from self-pollinated progeny of the hybrid combination P28 × 503p10. Molecular markers were used to confirm both the absence of chromosome 4B (Fig. 3, *a*) and the presence of chromosome 4H^{mar} (see Fig. 3, *b*) in isolated substitution lines. The DNA of the parent line P28 and *H. marinum* was used as a control.

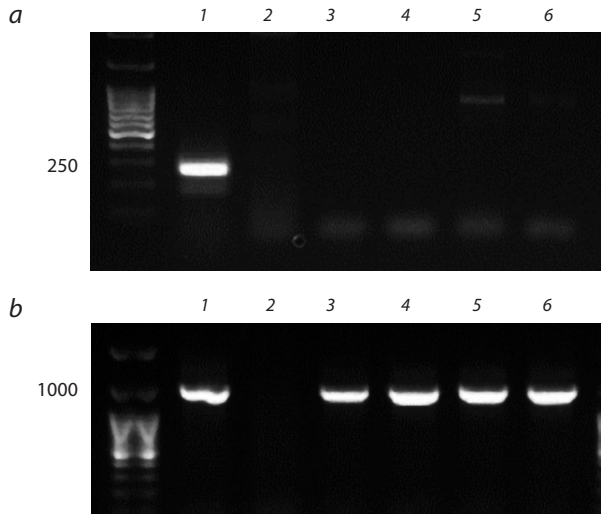

Fig. 3. PCR amplification of SSR marker *Xgwm368* (*a*) and EST marker *BAWU808* (*b*).

a: *1* – *T. aestivum*; *2* – *H. marinum*; *b*: *1* – *H. marinum*; *2* – *T. aestivum*; *3–6* – plants of line 4H^{mar}(4B). The numbers indicate the fragment length in bp.

Fig. 4. Coleoptile (*a*); ears at the base of the leaf (*b*).

1 – line 4H^{mar}(4B); *2* – line P28.

Characterization of the disomic wheat-barley substitution line 4H^{mar}(4B)

Plants of the line 4H^{mar}(4B) differed phenotypically from the parental line P28. Such traits included pronounced anthocyanin coloration of the coleoptile, which is typical for the wild barley *H. marinum* (Fig. 4, *a*), as well as an absence of purple coloration of the ears at the base of the leaves (see Fig. 4, *b*).

Substitution of wheat chromosome 4B for wild barley chromosome 4H^{mar} resulted in the development of viable plants, in which the values of some quantitative traits were higher than those of the parent line P28. The plant height and yield-related traits in the line 4H^{mar}(4B) were shown to be significantly different from those of the parental line (Table 4). Thus, the values of plant height, tiller and spike number per plant, spike length, spikelet number per spike, grain number per spike and per plant of the line 4H^{mar}(4B) were significantly higher than those of P28, while its thousand-kernel weight was less than that in the P28. Figures 5 and 6 show spikes and grains, respectively.

Table 4. Agronomic characteristics of the wheat–barley substitution line 4H^{mar}(4B)

Genotype	Plant height, cm	Tiller number	Spike number	Spike length	Spikelet number per spike	Grain number per spike	Grain number per plant	Thousand-kernel weight
4H^{mar}(4B)	133.63*	9.11*	5.6*	11.9*	22.3*	63.0*	236*	32.51*
П28	123.73	4.5	3.8	8.3	17.0	43.9	138.3	39.77
HCP_{05}	3.63	2.2	0.8	0.94	0.8	4.56	41.32	3.44

* $p < 0.05$.

Fig. 5. Spikes of lines: *1* – P28; *2–4* – 4H^{mar}(4B).

Fig. 6. Grains of line P28 (*1*); line 4H^{mar}(4B) (*2*).

Discussion

Wild relatives of bread wheat and their wheat synthetic derivatives are a valuable resource for introgressive hybridization (Davoyan et al., 2012; Molnár-Láng, Linc, 2015; Li et al., 2018). Tritordium is used in hybridization with wheat to transfer *H. chilense* chromosomes into the wheat genome (Martin, 2017), and triticale is used to transfer rye chromosomes (Shchapova, Kravtsova, 1990).

Earlier in our work, amphiploid *H. marinum* ssp. *gussoneanum–T. aestivum* (2*n* = 70) wild barley cytoplasm was used as a maternal parent in crosses with bread wheat to obtain alloplasmic disomic wheat–barley substitution lines 7H^{mar}(7B), 7H^{mar}(7D), 7H^{mar}L(7D), as well as ditelosomic addition lines 2*n* = 42 + 2t (7H^{mar}L) (Pershina et al., 2004; Trubacheeva et al., 2019). In this work, individual plants of the incomplete amphiploid *H. marinum* ssp. *gussoneanum–T. aestivum* (2*n* = 54) with wild barley cytoplasm were used as pollinators when crossing with the line P28 to introduce the genetic material of *H. marinum* into the euplasmic genetic background of bread wheat. The frequency of hybrid seed set was low, but some of the F_1 hybrid plants were viable, and two F_1 hybrids (P28 × 503p5) and (P28 × 503p10) set seeds that germinated.

The analysis of the obtained data revealed differences in the process of formation and the rate of cytological stabilization between the progeny of hybrids (P28 × 503p5) and (P28 × 503p10). This process was slower in the hybrid combination (P28 × 503p5) compared to the combination (P28 × 503p10). Thus, the progeny of 42-chromosome plants of the hybrid (P28 × 503p5) in the F_5 and F_6 segregated in plants with different levels of fertility, including completely sterile ones. In F_6, only half of 32 plants were fertile and fully fertile. In the hybrid combination (P28 × 503p10), on the contrary, only fertile or fully fertile plants were obtained in the fifth and sixth generations already. In addition, in the progeny of the hybrid (P28 × 503p10), in contrast to (P28 × 503p5), barley telocentric chromosomes appeared. Such results are consistent with the data of the authors who emphasized the uniqueness of the progeny of each hybrid grain as a source of diversity in the development of wide hybrid derivatives (Shchapova, Kravtsova, 1990).

As follows from the results obtained, the process of stabilization of karyotypes of 42-chromosome plants in the progeny of the (P28 × 503p10) F_6 hybrid was accompanied by the substitution of a pair of wheat chromosomes for a pair of *H. marinum* chromosomes. It was shown using GISH, C-banding and molecular analysis that chromosome 4B was substituted by chromosome 4H^{mar} in cytogenetically stable euploid plants. In addition, it was found that telocentrics also belong to the chromosome 4H^{mar}.

The type of wheat-barley substitution was also confirmed by the morphological traits that were exhibited in the line 4H^{mar}(4B). The absence of purple coloration of the ears at the base of the leaves in the line 4H^{mar}(4B) indicates the absence of wheat chromosome 4B, since this trait is controlled by the *Ra2* gene located on this chromosome (Melz, Thiele, 1990) and exhibited in P28. The smaller grain in the substitution line 4H^{mar}(4B) can be associated both with the absence of chromosome 4B, which affects the grain size and shape in wheat (Rahman et al., 2020), and with the presence of chromoso-

me 4H^{mar}, because *H. marinum* belongs to small-seeded barley grasses (Bothmer et al., 1991).

The line 4H^{mar}(4B) had a clear phenotypic marker specific for *H. marinum*. This is anthocyanin coleoptile coloration, which is absent in the parent line P28 and was previously found in the alloplasmic wheat–barley substitution line 7H^{mar}(7D) (Khlestkina et al., 2011). The accumulation of anthocyanin in vegetative organs is positively related to resistance to stress factors, and in wheat, ability to accumulate anthocyanins in the coleoptile is controlled by the *Rc* (red coleoptile) genes (Khlestkina et al., 2011). In this regard, the line 4H^{mar}(4B) may be useful for future studies of resistance to abiotic stresses, because *H. marinum* contains resistance genes (Garthwaite et al., 2005; Islam et al., 2007; Malik et al., 2009).

It has been established that homoeologous group 4 chromosomes of other barley species also possess genes that could be used for wheat breeding. For instance, chromosome 4H^{ch} of *H. chilense* contains genes for resistance to *Septoria tritici* and salt stress (Said, Cabrera, 2009), and chromosome 4H of *H. vulgare* was able to increase water use efficiency associated with drought tolerance of a wheat substitution line (Molnár et al., 2007). The line 4H^{mar}(4B) obtained in our work may have a potential for breeding, since it is characterized by high yield. Thus, the values of the number of spikes, the length of the spike, the number of spikelets per spike, the number of grains per and per plant in the 4H^{mar}(4B) line are higher than those in the line P28. These results, as well as the cytogenetic stability of the line 4H^{mar}(4B), indicated homoeology and high ability of barley 4H^{mar} chromosome to compensate for wheat 4B.

Conclusion

Thus, the efficiency of using the incomplete amphiploid (*H. marinum* ssp. *gussoneanum–T. aestivum*) (2*n* = 54) with wild barley cytoplasm to transfer *H. marinum* chromosomes into euplasmic genetic background of bread wheat has been shown.

References

Badaeva E.D., Badaev N.S., Gill B.S., Filatenko A.A. Intraspecific karyotype divergence in *Triticum araraticum* (*Poaceae*). *Plant Syst. Evol.* 1994;192:117-145. DOI 10.1007/BF00985912.

Bothmer R.V., Jacobsen N., Baden C., Jorgensen R.B., Linde-Laursen I. An Ecogeographical Study of the Genus *Hordeum*. Rome: IBPGR, 1991.

Davoyan R.O., Bebyakina I.V., Davoyan O.R., Zinchenko A.N., Davoyan E.R., Kravchenko A.M., Zubanova Y.S. The use of synthetic forms in the preservation and exploitation of the gene pool of wild common wheat relatives. *Russ. J. Genet. Appl. Res.* 2012;2:480-485. DOI 10.1134/S2079059712060044.

Dospekhov B.A. Methodology of Field Experiments (With the Basics of Statistical Processing of Research Results. Moscow: Agropromizdat Publ., 1985. (in Russian)

Fang Y., Yuan J., Wang Z., Wang H., Xiao J., Yang Z. Development of *T. aestivum* L.–*H. californicum* alien chromosome lines and assignment of homoeologous groups of *Hordeum californicum* chromosomes. *J. Genet. Genomics*. 2014;41(8):439-477. DOI 10.1016/j.jgg.2014.06.004.

Garthwaite A.J., von Bothmer R., Colmer T.D. Salt tolerance in wild *Hordeum* species is associated with restricted entry of Na^+ and Cl^- into the shoots. *J. Exp. Bot*. 2005;56(419):2365-2378. DOI 10.1093/jxb/eri229.

Hao M., Zhang L., Ning S., Huang L., Yuan Z., Wu B., Yan Z., Dai S., Jiang B., Zheng Y., Liu D. The resurgence of introgression breeding, as exemplified in wheat improvement. *Front. Plant Sci*. 2020;11:252. DOI 10.3389/fpls.2020.00252.

Islam S., Malik A.I., Islam A.K.M.R., Colmer T.D. Salt tolerance in a *Hordeum marinum–Triticum aestivum* amphiploid, and its parents. *J. Exp. Bot.* 2007;58(5):1219-1229. DOI 10.1093/jxb/erl293.

Khlestkina E.K., Antonova E.V., Pershina L.A., Soloviev A.A., Badaeva E.D., Borner A.A., Salina E.A. Variability of *Rc* (red coleoptile) alleles in wheat and wheat-alien genetic stock collections. *Cereal Res. Commun*. 2011;39(4):465-474. DOI 10.1556/CRC.39.2011.4.1.

Kobylyanskiy V.D. Biological characters of wild barley species in relation to aims of breeding. *Biologicheskiy Zhurnal Armenii = Armenian Journal of Biology.* 1967;20(10):41-51. (in Russian)

Kruse A. *Hordeum* × *Triticum* hybrids. *Hereditas*. 1973;73(1):157-161. DOI 10.1111/j.1601-5223.1973.tb01078.x.

Li A., Lui D., Yang W., Kishii M. Synthetic hexaploid wheat: yesterday, today, and tomorrow. *Engineering*. 2018;4(4):552-558. DOI 10.1016/j.eng.2018.07.001.

Malik A.I., English J.P., Colmer T.D. Tolerance of *Hordeum marinum* accessions to O_2 deficiency, salinity and these stresses combined. *Ann. Bot.* 2009;103(2):237-248. DOI 10.1093/aob/mcn142.

Martín A. Tritordeum: a man-made cereal. In: 4th Conference of Cereal Biotechnology and Breeding Jointly Organized by EUCARPIA Cereal Section November 6–9, 2017, Budapest, Hungary. Budapest, 2017;5-6.

Martín A., Álvarez J.B., Martín L.M., Barro F., Ballesteros J. The development of tritordeum: a novel cereal for food processing. *J. Cereal Sci*. 1999;30(2):85-95. DOI 10.1006/jcrs.1998.0235.

Melz G., Thiele V. Chromosome locations of genes controlling 'purple leaf base' in rye and wheat. *Euphytica.* 1990;49:155-159. DOI 10.1007/BF00027265.

Molnár-Láng M., Linc G. Wheat-barley hybrids and introgression lines. In: Molnár-Láng M., Ceoloni C., Dolezel J. (Eds.). Alien Introgression in Wheat. Cytogenetics, Molecular Biology and Genomics. Cham: Springer, 2015;315-346. DOI 10.1007/978-3-319-23494-6_12.

Molnár I., Linc G., Dulai S., Nagy E.D., Molnár-Láng M. Ability of chromosome 4H to compensate for 4D in response to drought stress in a newly developed and identified wheat-barley 4H(4D) disomic substitution line. *Plant Breed.* 2007;126(4):369-374. DOI 10.1111/j.1439-0523.2007.01300.

Molnár-Láng M., Linc G., Szakács É. Wheat-barley hybridization: the last 40 years. *Euphytica.* 2014;195:315-329. DOI 10.1007/s10681-013-1009-9.

Pershina L.A., Numerova O.M., Belova L.I., Devyatkina E.P., Rakovtseva T.S., Shumny V.K. Expression of fertility during morphogenesis in self-pollinated backcrossed progenies of barley-wheat amphiploids [*Hordeum geniculatum* All. (2*n* = 28) × *Triticum aestivum* L. (2*n* = 42)] (2*n* = 70). *Russ. J. Genet.* 2004;40:510-514. DOI 10.1023/B:RUGE.0000029153.61243.c2.

Pershina L.A., Devyatkina E.P., Belova L.I., Trubacheeva N.V., Arbuzova V.S., Kravtsova L.A. Features of alloplasmic wheat–barley substitution and addition lines (*Hordeum marinum* subsp. *gussoneanum*)-*Triticum aestivum*. *Russ. J. Genet.* 2009;45(10):1223-1229. DOI 10.1134/S102279540910010X.

Pershina L.A., Belova L.I., Trubacheeva N.V., Osadchaya T.S., Shumny V.K., Belan I.A., Rosseeva L.P., Nemchenko V.V., Abakumov S.N. Alloplasmic recombinant lines (*H. vulgare*)-*T. aestivum* with 1RS.1BL translocation: initial genotypes for production of common wheat varieties. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2018;22(5):544-552. DOI 10.18699/VJ18.393. (in Russian)

Rahman S., Islam S., Yu Z., She M., Nevo E., Ma W. Current progress in understanding and recovering the wheat genes lost in evolution and domestication. *Int. J. Mol. Sci*. 2020;21(16):5836. DOI 10.3390/ijms21165836.

Rey M.-D., Calderón M.-C., Rodrigo M.J., Zacarías L., Alós E., Prieto P. Novel bread wheat lines enriched in carotenoids carrying *Hordeum chilense* chromosome arms in the *ph1b* background. *PLoS One.* 2015;10(8):e0134598. DOI 10.1371/journal.pone.0134598.

Röder M.S., Korzun V., Wendehake K., Plaschke J., Tixier M.H., Leroy P., Ganal M.W. A microsatellite map of wheat. *Genetics*. 1998;149(4):2007-2023. DOI 10.1093/genetics/149.4.2007.

Rubiales D., Moral A. Resistance of *Hordeum chilense* against loose smuts of wheat and barley (*Ustilago tritici* and *U. nuda*) and its expression in amphiploids with wheat. *Plant Breed*. 2011;130(1):101-103. DOI 10.1111/j.1439-0523.2010.01818.x.

Said M., Cabrera A. A physical map of chromosome 4H^{ch} from *H. chilense* containing SSR, STS and EST-SSR molecular markers. *Euphytica*. 2009;167:253-259. DOI 10.1007/s10681-009-9895-6.

Shchapova A.I., Kravtsova L.A. Cytogenetics of Wheat-rye Hybrids. Novosibirsk: Nauka Publ., 1990. (in Russian)

Taketa S., Takeda K. Production and characterization of a complete set of wheat-wild barley (*Hordeum vulgare* ssp. *spontaneum*) chromosome addition lines. *Breed Sci.* 2001;51(3):199-206. DOI 10.1270/jsbbs.51.199.

Trubacheeva N.V., Badaeva E.D., Osadchaya T.S., Pershina L.A. Use of *H. vulgare* EST markers, GISH and C-banding to study bread wheat–*H. marinum* subsp. *gussoneanum* (2*n* = 28) introgression lines. *Cereal Res. Commun*. 2019;47(4):593-603. DOI 10.1556/0806.47.2019.37.

ORCID ID

L.A. Pershina orcid.org/0000-0002-9941-2026
N.V. Trubacheeva orcid.org/0000-0002-6701-6811
V.K. Shumny orcid.org/0000-0003-1939-6140
E.D. Badaeva orcid.org/0000-0001-7101-9639

Acknowledgements. This research was co-funded by Russian Foundation for Basic Research Grant No. 20-016-00196 and Budget Project FWNR-2022-0017.
We thank the Microscopic Center of the Siberian Branch of the Russian Academy of Sciences (http://www.bionet.nsc.ru/microscopy/) for granting access to microscopic equipment.

Conflict of interest. The authors declare no conflict of interest.

Received March 20, 2023. Revised May 17, 2023. Accepted May 18, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):553-564
DOI 10.18699/VJGB-23-67

Original Russian text https://vavilovj-icg.ru/

Molecular cytological analysis of alien introgressions in common wheat lines created by crossing of *Triticum aestivum* with *T. dicoccoides* and *T. dicoccum*

O.A. Orlovskaya[1], I.N. Leonova[2, 3] ✉, L.A. Solovey[1], N.I. Dubovets[1]

[1] Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, Belarus
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Department of Genetics and Selection, Novosibirsk State Agricultural University, Novosibirsk, Russia
✉ leonova@bionet.nsc.ru

Abstract. Wild and domesticated emmer (BBAA, 2*n* = 28) are of significant interest for expanding the genetic diversity of common wheat as sources of a high protein and microelement grain content, resistance to many biotic and abiotic factors. Particular interest in these species is also determined by their close relationship with *Triticum aestivum* L., which facilitates interspecific hybridization. The objective of this work was to analyze the nature of alien introgressions in hybrid lines from crossing common wheat varieties with *T. dicoccoides* and *T. dicoccum*, and to assess the effect of their genome fragments on the cytological stability of introgression lines. A C-banding technique and genotyping with SNP and SSR markers were used to determine localization and length of introgression fragments. Assessment of cytological stability was carried out on the basis of chromosome behavior in microsporogenesis. A molecular cytogenetic analysis of introgression wheat lines indicated that the inclusion of the genetic material of wild and domesticated emmer was carried out mainly in the form of whole arms or large fragments in the chromosomes of the B genome and less extended inserts in the A genome. At the same time, the highest frequency of introgressions of the emmer genome was observed in chromosomes 1A, 1B, 2B, and 3B. The analysis of the final stage of meiosis showed a high level of cytological stability in the vast majority of introgression wheat lines (meiotic index was 83.0–99.0 %), which ensures the formation of functional gametes in an amount sufficient for successful reproduction. These lines are of interest for the selection of promising material with agronomically valuable traits and their subsequent inclusion in the breeding process.
Key words: common wheat; *Triticum aestivum*; *T. dicoccoides*; *T. dicoccum*; introgression lines; C-banding; SSR analysis; SNP analysis; microsporogenesis; cytological stability.

For citation: Orlovskaya O.A., Leonova I.N., Solovey L.A., Dubovets N.I. Molecular cytological analysis of alien introgressions in common wheat lines created by crossing of *Triticum aestivum* with *T. dicoccoides* and *T. dicoccum*. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):553-564. DOI 10.18699/VJGB-23-67

Молекулярно-цитологический анализ интрогрессивных линий, полученных от скрещивания мягкой пшеницы *Triticum aestivum* с *T. dicoccum* и *T. dicoccoides*

О.А. Орловская[1], И.Н. Леонова[2, 3] ✉, Л.А. Соловей[1], Н.И. Дубовец[1]

[1] Институт генетики и цитологии Национальной академии наук Беларуси, Минск, Республика Беларусь
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Кафедра генетики и селекции, Новосибирский государственный аграрный университет, Новосибирск, Россия
✉ leonova@bionet.nsc.ru

Аннотация. Дикая и культурная полбы (BBAA, 2*n* = 28) представляют значительный интерес для расширения генетического разнообразия мягкой пшеницы как источники генетических факторов, определяющих высокое содержание белка и микроэлементов в зерне и устойчивость ко многим биотическим и абиотическим факторам. Особое внимание к этим видам обусловлено их близким родством с *Triticum aestivum*, что облегчает межвидовую гибридизацию. Целью настоящей работы были анализ характера чужеродных интрогрессий у гибридных линий, полученных от скрещивания сортов мягкой пшеницы с *T. dicoccoides* и *T. dicoccum*, и оценка их влияния на цитологическую стабильность. Для установления локализации и протяженности фрагментов интрогрессии использовали С-бэндинг и генотипирование маркерами SNP и SSR. Оценку цитологической стабильности проводили на основе изучения поведения хромосом в микроспорогенезе. Молекулярно-цитогенетический анализ интрогрессивных линий пшеницы свидетельствует о том, что включение генетического материала дикой и культурной полб осуществляется преимущественно в виде целых плеч или крупных фрагментов в хромосомах генома B и менее протяженных вставок в геноме A. При этом наибольшая частота интрогрессий фрагментов

генома полб наблюдалась в хромосомах 1A, 1B, 2B, 3B. Изучение заключительной стадии мейоза показало высокий уровень цитологической стабильности у большинства интрогрессивных линий пшеницы (мейотический индекс составил 83.0–99.0 %), что обеспечивает формирование у них функциональных гамет в количестве, достаточном для успешной репродукции. Данные линии представляют интерес для отбора перспективного материала с хозяйственно ценными признаками с последующим включением их в селекционный процесс.
Ключевые слова: мягкая пшеница; *Triticum aestivum*; *T. dicoccoides*; *T. dicoccum*; интрогрессивные линии; С-бэндинг; SSR-анализ; SNP-анализ; микроспорогенез; цитологическая стабильность.

Introduction

Common wheat *Triticum aestivum* L. (BBAADD, 2*n* = 42) is one of the most important cereal crops and a major source of calories for most of the world's population. In addition to food purposes, wheat is used in the pulp and paper and chemical industries, for the production of ethanol, and flour milling waste and feed grains are used as livestock feed. Breeding bread wheat for high-yielding varieties has significantly reduced the level of genetic diversity compared to wild relatives (Xie, Nevo, 2008; Nevo, Chen, 2010; Budak et al., 2013). At present, other species of the genus *Triticum* are increasingly used to expand the bread wheat gene pool (Jaradat, 2013; Liu et al., 2019; Orlovskaya et al., 2020).

Wild tetraploid wheat, or wild emmer wheat *T. dicoccoides* Schwein f. (BBAA, 2*n* = 28) appeared as a result of spontaneous hybridization between diploid species: *T. urartu* Thum. (AA, 2*n* = 14) and unknown close relative of *Aegilops speltoides* Tausch. (SS, 2*n* = 14) (Dvorak et al., 1993; Peng et al., 2011). It is assumed that wild emmer participated in the formation of domesticated emmer *T. dicoccum* (Schrank.) Schuebl (BBAA, 2*n* = 28). The formation of common wheat occurred as a result of natural hybridization of a tetraploid species from the genus *Triticum* (BBAA) and a diploid species *Ae. tauschii* Coss., a D genome donor (Petersen et al., 2006).

Wild and domesticated emmer are of significant interest in expanding the genetic diversity of common wheat. Many accessions of *T. dicoccoides* are known to be adapted to unfavorable environments (Peleg et al., 2005; Nevo, Chen, 2010), characterized by a high protein and microelements content (Cakmak et al., 2004; Uauy et al., 2006; Wang Z. et al., 2018). More than 20 genes and QTLs have been identified in the genomes of *T. dicoccum* and *T. dicoccoides* for resistance to powdery mildew, leaf rust, yellow rust, and Fusarium (Peng et al., 2000; Xie, Nevo, 2008). Interest in wild and domesticated emmer is also due to phylogenetic relationship with common wheat. However, despite the closeness of emmer genomes to common wheat genomes A and B, the transfer of alien chromatin into cultivars may be accompanied by introgression of genetic material that negatively affects agronomically important traits.

Cytological and molecular methods are effective tools for chromosomal identification of foreign chromatin in the common wheat genome. One of them is differential staining of mitotic chromosomes (C-banding), which makes it possible, based on a comparison of C-banding patterns in hybrids and initial parental forms, to reveal structural transformations of the karyotype, indicating the introgression of foreign chromatin. However, many cereal genomes contain an insignificant amount of heterochromatin, and closely related species often have a similar C-banding pattern, which limits the application of this method (Surzhikov et al., 2007; Dedkova et al., 2009).

Molecular markers are effective for detecting of structural changes in low-heterochromatin genomes and identifying of short introgressed fragments, among which SSR and SNP markers are most widely used (Zhou et al., 2013; Jorgensen et al., 2017). At present molecular genetic maps of the chromosomes of hexaploid wheat and wild emmer have been constructed based on SSR and SNP markers specific for the A, B, and D genomes of *T. aestivum* (Röder et al., 1998; Pestsova et al., 2000; Wang S. et al., 2014; Maccaferri et al., 2015). The use of molecular markers increases the efficiency of foreign introgression monitoring.

Previously, we studied the nature of foreign substitutions and translocations and the process of stabilization of hybrid lines obtained by crossing common wheat varieties with *T. kiharae* (A^tA^tGGDD, 2*n* = 42) (Orlovskaya et al., 2020). It has been shown that the introgression of the genetic material of *T. kiharae* occurs as whole chromosomes or large fragments (centric and terminal translocations). The objective of this work was to analyze the nature of alien introgressions in hybrid lines obtained from hybridization of common wheat varieties with *T. dicoccoides* and *T. dicoccum*, and to assess their effect on the cytological stability of introgression lines.

Materials and methods

We used nine F_{10} lines obtained at the Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, from crossing common wheat varieties *T. aestivum* (Rassvet, Festivalnaya, and Pitic S62) with emmer accessions from the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR) collection *T. dicoccoides* k-5199 and *T. dicoccoides* (the origin of this accession is unknown), and *T. dicoccum* k-45926. Line 29 (Rassvet × *T. dicoccoides* k-5199); lines 11-1, 13-3, 15-7-2, and 16-5 (*T. dicoccoides* × Festivalnaya); lines 206-2-2 and 213-1-2 (Pitic S62 × *T. dicoccum* k-45926) and lines 1-3 and 2-7 (Festivalnaya × *T. dicoccum* k-45926) were created by self-pollination of F_1 hybrids and subsequent generations and selected for molecular cytogenetic studies based on the assessment of the inheritance of morphological and productivity traits in F_1–F_9 generations.

The preparation of cytological plates and the procedure of C-banding were performed according to the method of E.D. Badaeva et al. (1994). Identification of individual chromosomes of A-, B- and D-genomes was carried out in accordance with the generalized ideogramme of differentially stained chromosomes developed by E.D. Badaeva et al. (1990). Stained slides were analyzed using Amplival microscope (Carl Zeiss, Jena, Germany). Selected metaphase plates were photographed using the Leica DC 300 digital video camera.

Processing of the obtained images was carried out using graphics editor Adobe Photoshop 2017.

Genomic DNA was isolated from the seedlings of 5–7-days old as described in E.S. Skolotneva et al. (2017). Genotyping with SNP markers was performed using Illumina Infinium 20K chip technology (TraitGenetics, Germany, http://www.traitgenetics.com). SSR markers (WMC, GWM, and GDM) were used to clarify the chromosomal location and the length of introgression fragments, using two or more markers per chromosome arm. Polymerase chain reaction (PCR) protocols for SSR markers are described in M.S. Röder et al. (1998). Separation of PCR fragments was performed on an ABI PRISM 3100 automatic sequencer (Applied Biosystems, USA). The fragment size was calculated using the ABI GeneScan software (version 2.1). Putative chromosomal localization was determined based on wheat chromosome consensus maps constructed using SSR and SNP markers (Somers et al., 2004; Wang S. et al., 2014).

Microsporogenesis was studied on temporary squashed preparations. Spikes were cut before leaving the leaf sheath and fixed in the ethanol-acetic mixture (3:1). A day after fixation, the material was transferred to 70 % ethyl alcohol, where it was stored before analysis at $t = +2–4$ °C. Acetoorcein (2 %) was used as a dye. For each cross combination and initial forms, 30 plates of metaphase I and 50–100 microsporocytes of the following stages of meiosis (anaphase I and II, metaphase II, tetrads) were analyzed. The slides were analyzed on the microscope Amplival (Carl Zeiss) with Apochromate lens 100x aperture 1.32 MI.

Statistical data analysis was carried out using STATISTICA v. 10 (http://statsoft.ru/) and MS Excel 2010.

Results

C-banding

In this study, karyotyping of five hybrid lines obtained with the involvement of common wheat cultivars and two accession of *T. dicoccoides* was carried out. Of these, two lines (11-1 and 13-3) were tetraploid ($2n = 4x = 28$), while others were stabilized at the hexaploid ploidy level ($2n = 6x = 42$).

Both *T. dicoccoides* accessions had an almost identical C-banding pattern, while differences between varieties were observed in the degree of expression and the presence/absence of a number of telomeric and intercalary blocks of heterochromatin, which ensured the individuality of their karyotypes. The vast majority of these differences were found in the heterochromatin-rich genome B. As for the A-genome, polymorphism between parental forms was noted only in two chromosomes – 4A and 6A, and only in the *T. dicoccoides* × Festivalnaya cross combination: chromosome 4A of *T. dicoccoides* has a large telomeric block of heterochromatin in the long arm, while Festivalnaya has a more pronounced subtelomeric block. The 6A chromosome of *T. dicoccoides* is also distinguished by the presence of a bright intercalary block in the proximal region of the long arm. Comparison of the obtained patterns of differential chromosomes staining in hybrid lines of wheat and the corresponding parental forms made it possible to identify the presence of introgressions of the genetic material of *T. dicoccoides* in all five lines (Table 1).

In line 29, the distribution of heterochromatin blocks characteristic of wild emmer was noted in the proximal region of the long arm of chromosome 1B and in the distal regions of the short arm of chromosome 2B and the long arm of 5B (Fig. 1).

It is not possible to determine the size of the *T. dicoccoides* genome fragments included in the cv. Rassvet genome due to the identity of the C-banding patterns of the parental forms in the adjacent regions of the chromosomes. In this regard, in Fig. 1 and in all subsequent Figures depicting the karyotypes of the studied lines, arrows mark only the localization sites of alien fragments. Changes in the C-banding pattern were also detected in chromosome 3B, both in the long and short arms, which gave us reason to assume that in this case, the whole chromosome of *T. dicoccoides* was introgressed. As for chromosome 4B, in all parental forms this chromosome had an identical pattern of staining, which did not allow us to conclude the possible fact of the genetic material exchange between these homoeologs of wheat and *T. dicoccoides* in any of the hybrid lines.

In the karyotype of line 11-1, the C-banding pattern typical for *T. dicoccoides* was observed in the distal region of the long

Table 1. Chromosomal localization of emmer genetic material in introgression wheat lines according to C-banding and genotyping data using SSR and SNP markers

Cross combination	Line	C-banding	SSR and SNP markers
Rassvet × *T. dicoccoides* k-5199	29	1BL, 2BS, 3B, 5BL	1BL, 2BS, 3B, 5B, 6AL
T. dicoccoides × Festivalnaya	11-1	1B, 2BL, 3BS, 4AL, 5BL	1AL, 1BL, 2BL, 4AL, 4B, 5BL, 6AL, 6BL, 7A, 7B
	13-3	2BL, 3B, 4AL, 6AL, 6BS, 7BS	1A, 1BL, 2A, 2B, 3A, 3B, 4A, 4B, 5A, 5BS, 6AL, 6B, 7A, 7B
	15-7-2	2BL, 3BL, 5BL	1AL, 1BL, 2B, 3A, 3BL, 4AL, 4B, 5AL, 5BL, 6BL, 7AL, 7BL
	16-5	2BL, 3BL, 5BL, 6BS	1A, 1B, 2B, 3A, 3BL, 4AL, 4B, 5BL, 6AL, 6B, 7A, 7B
Pitic S62 × *T. dicoccum* k-45926	206-2-2	1BL, 7A, 7BL	1AL, 1BL, 2AL, 2B, 3AL, 3BL, 4AL, 4BL, 5AL, 5BL, 6AL, 6BS, 7A, 7BL
	213-1-2	2BL	1BL, 2AL, 2BL, 3AL, 3BL, 4AL, 5AL, 5BL, 6A, 7AL, 7BL
T. dicoccum k-45926 × Festivalnaya	1-3	1AL, 1BS, 2AL, 6BL	1AL, 1BS, 2AL, 2B, 3AS, 3BL, 4AS, 5AL, 5BL, 6AS, 6BL, 7AS, 7B
	2-7	1BS, 2AS, 3B, 4AL, 6BL, 7BL	1AS, 1BS, 2AS, 2BL, 3A, 3B, 4AL, 5A, 5BL, 6AS, 6BL, 7AS, 7BL

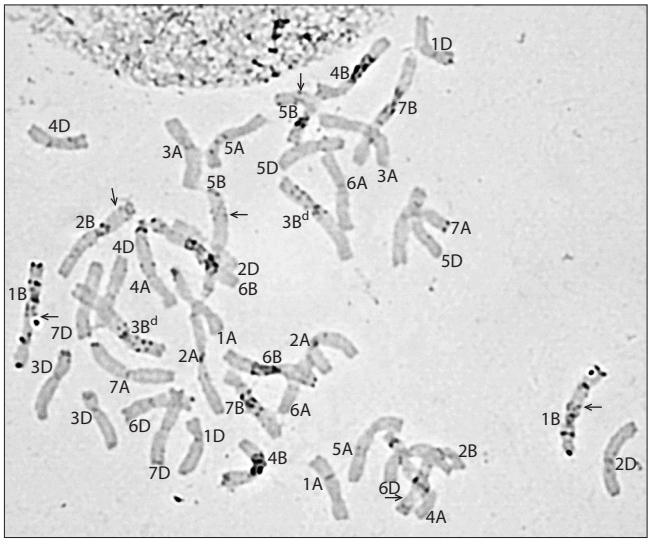

Fig. 1. Karyotype of introgression wheat line 29 Rassvet × *T. dicoccoides* k-5199.

Here and in Fig. 2: locations of introgressed fragments of wild emmer are indicated by arrows.

arm of chromosome 4A, as well as in the central regions of the short arm of 3B and long arms of chromosomes 2B and 5B (Fig. 2, *a*). Taking into account that 4A homoeolog of wild emmer and Festivalnaya differ only in the C-banding pattern of the distal region of the long arm, it is not possible to infer unequivocally whether the whole emmer chromosome, or the long arm, or only its distal part was introgressed. At the same time, chromosome 1B, judging by the distribution of heterochromatin blocks in both arms, most likely belongs entirely to *T. dicoccoides*. In other cases, the insertion of emmer chromatin fragments is more likely.

The karyotype of line 13-3 (see Fig. 2, *b*) demonstrated changes in the C-banding pattern in the distal region of the long arm of chromosome 4A similar to changes in line 11-1. In addition, chromosome 6A has a bright intercalary block typical to *T. dicoccoides* in the proximal region of the long arm. As for the chromosomes of the genome B, introgressions of *T. dicoccoides* were identified in 2BL (presumably the whole arm belongs to wild emmer), and in the distal 6BS and proximal 7BS regions. Chromosome 3B, according to the C-banding pattern, belongs entirely to wild emmer. Of particular note is the fact that this line contains a heteromorphic pair of chromosomes 6B, in which only one of the homoeologs is characterized by a change in the C-banding pattern in the distal region of the short arm. At the same time, the size of the introduced emmer fragment is unclear, since, as in the case of chromosome 4A, polymorphism in the distribution of heterochromatin blocks in 6B homologues was noted only in the above region.

For line 15-7-2, introgression of the *T. dicoccoides* genetic material was found in the long arm of chromosome 2B, as well as in the proximal regions of the long arms of chromosomes 3B and 5B (see Fig. 2, *c*).

A similar set of recombinant chromosomes (2B, 3B, and 5B) was found in line 16-5 with the only difference that chromosome 5B contained a fragment of emmer chromatin not from the proximal, but from distal region of the long arm. In addition, a fragment of *T. dicoccoides* chromatin was identified in the distal region of the short arm of chromosome 6B (see Fig. 2, *d*).

It should be noted that introgression of the genetic material of *T. dicoccoides* into the genome of common wheat occurred in accordance with the homoeology of chromosomes. In the course of stabilization of karyotypes in each homoeologous group, identical variants of resulting recombinant chromosomes were selected, and as a result, all introduced fragments of the emmer genome are present in disomic state (the only exception is 6B chromosome in line 13-3). No chromatin exchanges between non-homoeologous chromosomes were found in the studied material.

The study of the nature of introgressions of domesticated emmer *T. dicoccum* k-45926 into the genome of common wheat was carried out using the material of four hybrid lines developed with the involvement of wheat varieties Festivalnaya and Pitic S62. Analysis of the karyotype of the emmer wheat demonstrated the presence of the T7AL-5BS.5BL translocation resulting from the transfer of a fragment of the long arm of chromosome 7A to the distal region of the short arm of chromosome 5B. The other structural transformations were found in chromosome 7A with a deletion of the distal part of the long arm. According to the literature data, both types of aberrant chromosomes are widespread among *T. dicoccum* accessions growing in the Mediterranean and Western Europe (Dedkova et al., 2007). Similar structural rearrangements of chromosomes were also noted in some *T. dicoccoides* genotypes (Badaeva et al., 2007).

When comparing the obtained patterns of differential staining of chromosomes in hybrid wheat lines and corresponding parental forms, the presence of introgressions of the *T. dicoccum* genetic material was established in all the lines under study (see Table 1). In the karyotype of line 1-3 from the *T. dicoccum* × Festivalnaya cross combination, C-banding pattern typical for *T. dicoccum* was observed in the distal regions of long arms of chromosomes 1A, 2A, and 6B. The size of introgressed fragments was approximately equal to half of the arm. For chromosome 1B, the insertion of a short arm of domesticated emmer was identified (Fig. 3, *a*).

The second line from the same cross combination is characterized by a larger amount of introgressed emmer chromatin: in the karyotypes of all analyzed plants, emmer chromatin was found in chromosome 3B, in chromosome 1B, near telomeres of 6B and 7B chromosomes, in the distal region of chromosome 2A, and also the distal region of the long arm of chromosome 4A (see Fig. 3, *b*).

In the karyotype of line 206-2-2 from the cross combination Pitic S62 × *T. dicoccum* k-45926, chromosome 7A, characteristic of the original emmer accession, was present with a deletion of the distal fragment of the long arm. The similarity of C-banding patterns in the short arms of this chromosome in the parental wheat cultivar and emmer does not allow us to make an unambiguous conclusion whether this chromosome is recombinant or completely belongs to emmer. We are more inclined to the second option. In addition, the C-banding pattern typical of *T. dicoccum* k-45926 was noted in the distal

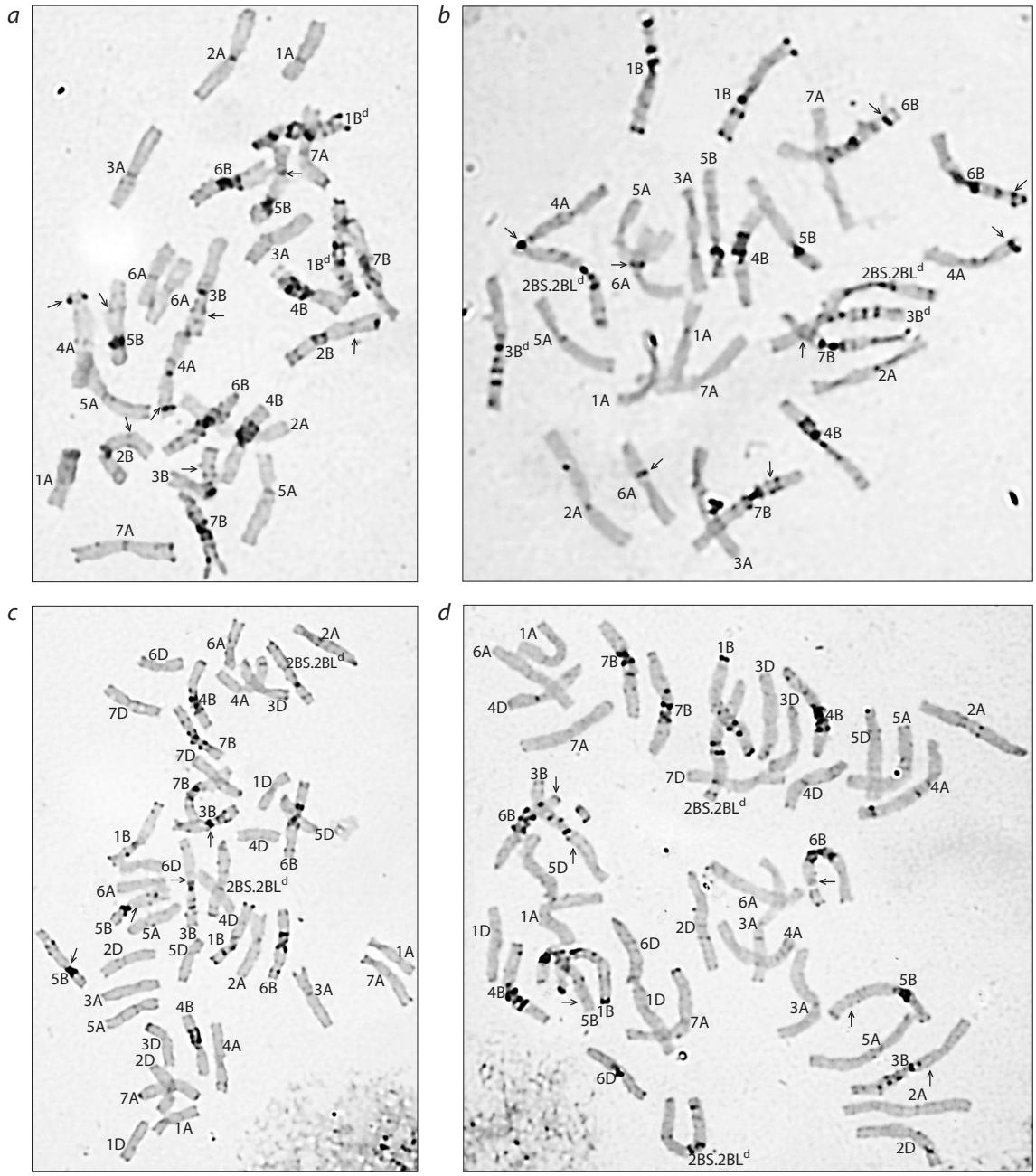

Fig. 2. Karyotypes of introgression wheat lines of *T. dicoccoides* × Festivalnaya cross combination. Lines: *a* – 11-1; *b* – 13-3; *c* – 15-7-2; *d* – 16-5.

Intact arms of *T. dicoccoides* are indicated with a superscript "d".

regions of chromosomes 1BL and 7BL (Fig. 4, *a*). The smallest amount of *T. dicoccum* k-45926 genetic material (long arm of chromosome 2B) introduced into the wheat genome was found in line 213-1-2 (see Fig. 4, *b*). As in case of the introgression of the wild emmer genetic material, all introduced fragments of the domesticated emmer genome are present in the karyotypes of hybrid lines in the disomic state.

Molecular analysis

The number of SNP markers mapped on different chromosomes of A-, B- or D-genomes varied significantly with the smallest number noted for chromosomes of the 4th homoeological group (Supplementary Material 1)[1]. More than 50 % of SNP markers of A and B genomes revealed polymorphism between *T. dicoccum*, *T. dicoccoides* and parental wheat varieties. A high amplification level of SNP markers of the D genome was also noted in *T. dicoccum* and *T. dicoccoides*. However, it is currently not possible to establish their chromosomal localization in the genome of wild and domesticated emmer.

Despite the high coverage of chromosomes with SNP markers, in the distal regions of some chromosomes (1AS, 1AL, 2BS, 3BL, 4AS, 4BL, 6AL, and 7AS) the number of polymor-

[1] Supplementary Materials 1 and 2 are available in the online version of the paper: https://vavilov.elpub.ru/jour/manager/files/Suppl_Orlovskaya_Engl_27_6.pdf

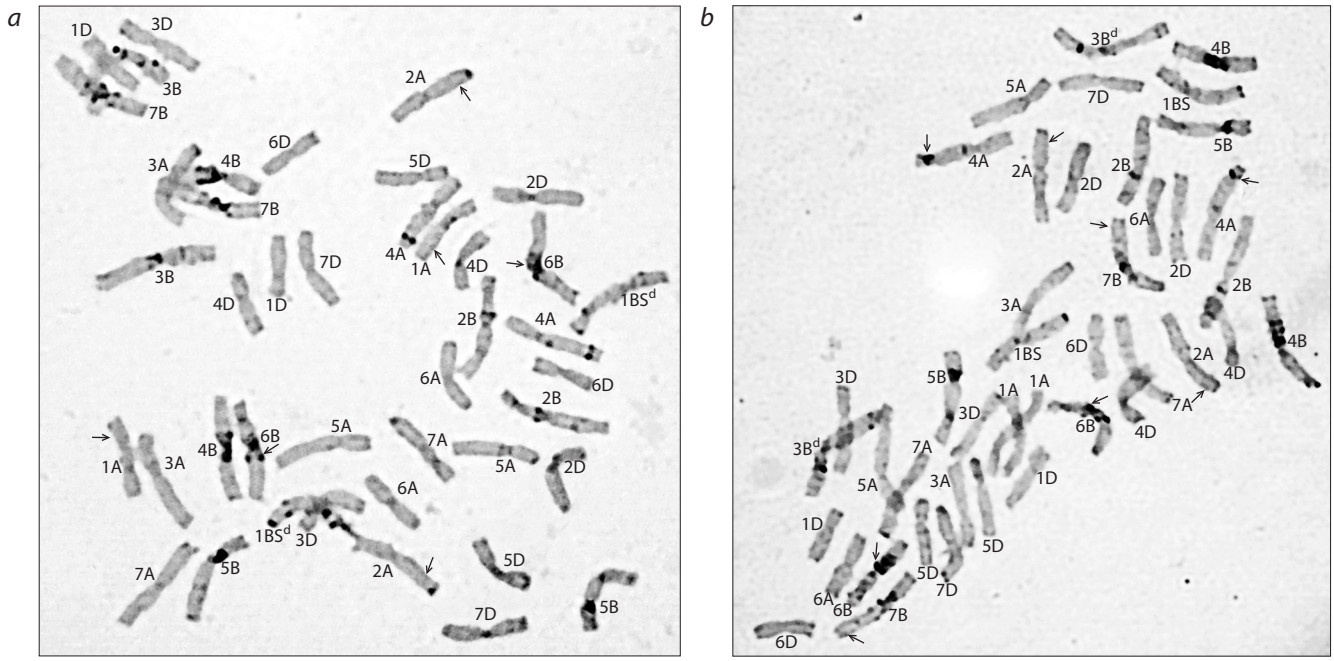

Fig. 3. Karyotypes of introgression wheat lines of *T. dicoccum* k-45926 × Festivalnaya combination: *a* – line 1-3; *b* – line 2-7.

Here and in Fig. 4: introduced chromosomes and whole arms of *T. dicoccum* k-45926 are marked with a superscript "d". Arrows indicate location of introgression fragments of domesticated emmer.

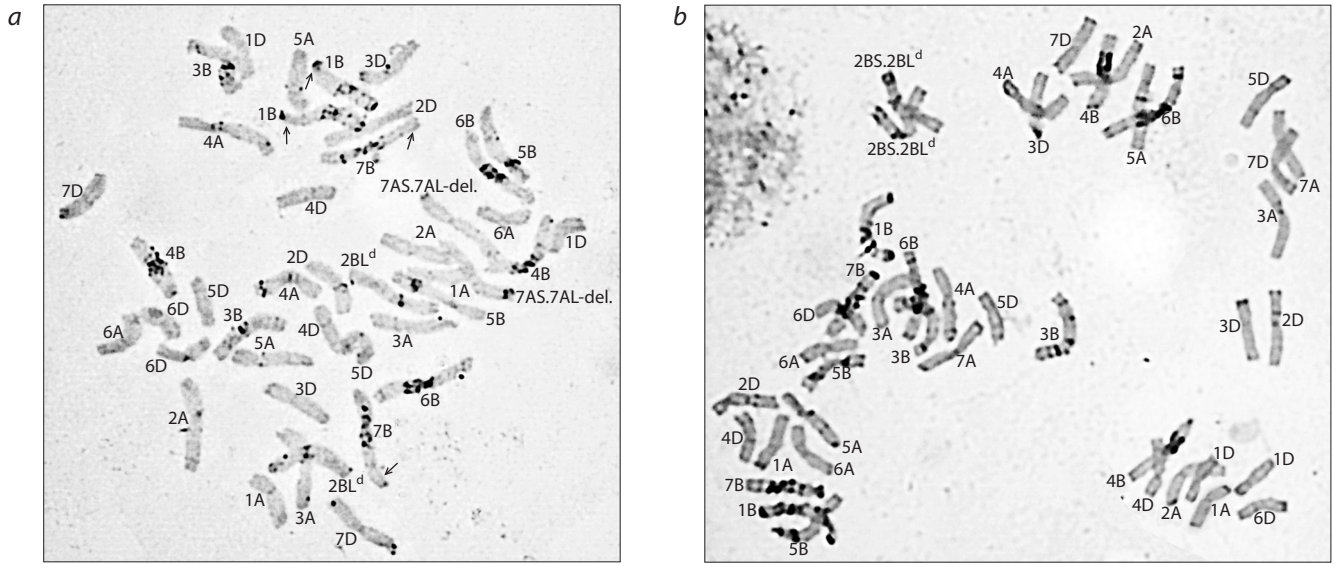

Fig. 4. Karyotypes of Pitic S62 × *T. dicoccum* k-45926 introgression wheat lines: *a* – line 206-2-2; *b* – line 213-1-2.

phic markers was insufficient to determine the completeness of substitution and fragment lengths. SSR markers were additionally used to correct the length of introgressed fragments. Analysis of the polymorphism of SSR markers indicates that almost all of the markers used are polymorphic, while the absence of amplification fragments for D-genome markers in tetraploid species was established (see Supplementary Material 1).

Genotyping of introgression lines of wheat and initial parental forms indicates that all the studied lines had recombination events involving chromosomes of wheat relatives or their fragments (see Table 1). At the same time, the frequency of substitutions and translocations and the length of introgressed fragments depend on the hybrid combination and the direction of crossing (see Table 1, Supplementary Material 2). Comparison of the amplification spectra of SNP and SSR markers demonstrated the presence of the wild emmer genetic material in most common wheat chromosomes, except for line 29 (Rassvet × *T. dicoccoides* k-5199). Only five fragments were found in the genome of this line, and their localization coincides with the data of cytological analysis. Perhaps this is due to the direction of crossing, since in this case the wheat

variety was used as a maternal component in contrast to the lines of the *T. dicoccoides* × Festivalnaya combination, where bread wheat was used as a pollinator.

Two lines from the *T. dicoccoides* × Festivalnaya combination (11-1 and 13-3) stabilized at the tetraploid level of ploidy; however, the nature of the recombination events of these lines was different. Fragments of introgression in line 13-3 are, as a rule, longer than in line 11-1. In line 11-1, according to molecular analysis, a complete replacement of the long arms of chromosomes 1B, 4A, 5B, 6A, and 7A is assumed; in other cases, inserts of foreign chromatin were insignificant (see Supplementary Material 2). In lines 15-7-2 and 16-5, alien chromatin was not found in chromosome 6A and 5A, respectively, and both lines have no introgressions in chromosome 2A.

However, these lines differed in the localization and extent of introgression fragments. Thus, in line 16-5, the inclusion of the genetic material of the *T. dicoccoides* B-genome was carried out in the form of larger fragments than in line 15-7-2 (see Supplementary Material 2). As for recombination events in the genome A, in line 16-5 a significant part of the 7A chromosome is replaced by the *T. dicoccoides* chromosome; 3AS and 4AL also originated from wild emmer. In the remaining A-genome chromosomes, *T. dicoccoides* segments were of small length and were found mainly in long arms. In line 15-7-2, small introgression fragments in the A-genome were found only in the long arms of chromosomes. An exception was chromosome 3A in which foreign chromatin was found in both arms (see Supplementary Material 2).

Line 206-2-2 (Pitic S62 × *T. dicoccum* k-45926) contained the genetic material of domesticated emmer in all chromosomes of A- and B-genomes; in line 213-1-2, introgression was not detected in chromosomes 1A, 4B, and 6B. Recombination events in the lines of this cross combination were of similar character (see Supplementary Material 2). For lines 1-3 and 2-7 of *T. dicoccum* k-45926 × Festivalnaya, molecular analysis did not show the presence of the alien genetic material only in chromosome 4B, while in chromosomes 1A, 2A, 2B, 3B, 4A and 7B fragments of domesticated emmer were localized in different arms (see Supplementary Material 2).

Microsporogenesis

The analysis of the chromosome behavior at the metaphase I stage showed that the number of chromosomes forming bivalents in all introgression lines exceeded 90 % and in the majority approached 100 % (Table 2). The highest level of chromosome pairing was noted in line 16-5 – 100 %. Only one cell with two univalents out of all studied pollen mother cells (PMC) was found in lines 29 and 1-3; single cells with two univalents were found in lines 13-3, 15-7-2, 213-1-2, and 2-7. In line 11-1, 73.3 % of cells with disorders were found out of the number of analyzed PMCs, and the number of univalents in them varied from two (33.3 %) to six (10.0 %). It should be noted that line 11-1 is also characterized by the highest number of open bivalents among the studied genotypes, which indicates weakening of chromosome pairing (see Table 2). The subsequent stages of meiosis proceeded with minor disturbances, which led to the formation of normal tetrads at the final stage (Table 3).

The exception was line 11-1 with the meiotic index of only 53.65 %, which is consistent with the data obtained from the analysis of the behavior of chromosomes at the metaphase I. It is the line that is characterized by the lowest level of synapsis among the studied introgression lines (see Table 2). Lines with a high rate of chromosomal associations in the early stages of meiosis, as a rule, had a higher value of the meiotic index (83.0–99.0 %).

At the final stage of meiosis, along with normal tetrads, abnormal tetrads with micronuclei of various sizes are formed, the number of which varied in the studied material from 1 to 6, but most often tetrads with 1–2 micronuclei were formed. It should be noted that the spectrum of abnormalities in more stable lines was much smaller. For example, in line 2–7 with a high meiotic index, only one cell with one micronucleus was found. PMCs containing six micronuclei were noted only for the least stable lines 11-1 and 206-2-2, and their frequency was very low (0.91–2.0 % of the total number of analyzed cells). In line 13-3, single triads and pentads were observed. The number of cells with such disorders was only 1.11 %. The appearance of triads is often explained by the presence of an autonomous spindle at metaphase I or metaphase II, the absence of kinetochore fibrils, or abnormal premature cytokinesis in prophase II (Sosnikhina et al., 2007).

It should be noted that a sufficiently high level of chromosome synapsis was already established by us in F_2 hybrids from crossing emmer wheat with common wheat varieties (Orlovskaya et al., 2010). The number of chromosomes included in the bivalents in these hybrids was at the level of 90.9–99.3 %. As a rule, only two chromosomes did not enter the mating process. Despite the rather high level of chromosome pairing of F_2 hybrids in metaphase I, the subsequent stages of meiosis in this generation proceeded with significant abnormalities. As a result, the percentage of normal tetrads (meiotic index as an indicator of the normal meiosis) was very low, ranging from 8 to 20 %. The obtained results indicate that by the tenth generation there was a significant stabilization of the meiosis, which ensures the formation of a sufficient number of functional gametes for the successful reproduction of the developed hybrid material.

Discussion

The data obtained indicate a high frequency of introgression of the genetic material of emmer wheat into the genome of the common wheat cultivar. It should be noted that the data of SNP and SSR analyses confirm the results of differential staining of chromosomes. Thus, the highest frequency of incorporation of the genetic material of *T. dicoccum* and *T. dicoccoides* into the common wheat genome was found for chromosomes 1B, 2B, and 3B, both according to the results of molecular analysis and C-banding.

Identification of introgressions in chromosomes of the A-genome using C banding, as already noted, is difficult due to the small number of diagnostic blocks and low polymorphism of A-genome chromosomes. In addition, wild and domesticated emmer wheat are tetraploid ($2n = 4x = 28$) with the BBAA genomic structure, where both genomes are homologous to the corresponding *T. aestivum* genomes ($2n = 6x = 42$; BBAADD genome). As a result, these species

Table 2. Average frequencies of different chromosome associations in metaphase I of F_{10} common wheat introgression lines and their parents

Cross combination/ parental varieties	Line	Bivalent, pcs			Chromosome number in bivalent, %	Univalent, pcs
		closed	opened	Total		
Rassvet × *T. dicoccoides* k-5199	29	19.23 ± 0.22	1.73 ± 0.22	20.96 ± 0.03	99.76	0.07 ± 0.06
T. dicoccoides × Festivalnaya	11-1	7.37 ± 0.33	5.30 ± 0.26	12.67 ± 0.23	90.48	2.47 ± 0.35
	13-3	12.23 ± 0.26	1.67 ± 0.24	13.90 ± 0.06	99.29	0.20 ± 0.11
	15-7-2	19.50 ± 0.22	1.23 ± 0.21	20.73 ± 0.08	98.7	0.53 ± 0.16
	16-5	19.67 ± 0.19	1.33 ± 0.20	21.00 ± 0	100	0
Pitic S62 × *T. dicoccum* k-45926	206-2-2	18.30 ± 0.28	2.40 ± 0.28	20.70 ± 0.1	98.57	0.60 ± 0.2
	213-1-2	19.27 ± 0.26	1.60 ± 0.24	20.87 ± 0.08	99.37	0.27 ± 0.16
T. dicoccum k-45926 × Festivalnaya	1-3	19.50 ± 0.18	1.47 ± 0.18	20.97 ± 0.03	99.84	0.07 ± 0.06
	2-7	19.50 ± 0.23	1.43 ± 0.22	20.93 ± 0.05	99.68	0.13 ± 0.09
Pitic S62		19.80 ± 0.19	1.17 ± 0.19	20.97 ± 0.03	99.84	0.07 ± 0.06
Rassvet		20.03 ± 0.16	0.97 ± 0.16	21.00 ± 0	100.0	0
Festivalnaya		19.50 ± 0.21	1.43 ± 0.21	20.97 ± 0.03	99.76	0.06 ± 0.06
T. dicoccum k-45926		12.97 ± 0.24	0.97 ± 0.24	13.97 ± 0.03	99.76	0.07 ± 0.06
T. dicoccoides		12.63 ± 0.18	1.30 ± 0.18	13.93 ± 0.07	99.52	0.07 ± 0.06
T. dicoccoides k-5199		12.20 ± 0.49	1.50 ± 0.43	13.70 ± 0.21	97.86	0.60 ± 0.43

Table 3. Characterization of meiotic stages in common wheat introgression lines and their parents

Cross combination/ parental varieties	Line	Number of normal PMCs, %			Meiotic index, %
		Anaphase I	Metaphase II	Anaphase II	
Rassvet × *T. dicoccoides* k-5199	29	68.00	81.67	82.86	93.00
T. dicoccoides × Festivalnaya	11-1	49.18	41.82	19.51	53.65
	13-3	80.0	74.55	80.0	92.33
	15-7-2	78.30	81.40	82.00	86.20
	16-5	85.00	88.57	85.00	89.09
Pitic S62 × *T. dicoccum* k-45926	206-2-2	62.9	77.14	88.57	83.0
	213-1-2	86.0	85.0	85.37	92.0
T. dicoccum k-45926 × Festivalnaya	1-3	82.86	90.00	90.01	96.00
	2-7	83.33	86.25	82.86	99.00
Pitic S62		87.14	89.19	80.0	87.00
Rassvet		80.00	96.67	91.12	99.00
Festivalnaya		82.50	90.91	88.32	84.55
T. dicoccum k-45926		91.25	92.50	95.00	97.50
T. dicoccoides		91.67	80.00	82.86	96.25
T. dicoccoides k-5199		82.35	81.67	85.00	95.71

have the same pattern of genome-specific C-bands, and the differences necessary for monitoring of alien introgressions may only concern permanent blocks of heterochromatin that are polymorphic in size, as well as non-permanent C-bands that vary in the presence/absence and size. This greatly complicates the detection of the introgression of emmer chromatin in the common wheat genome by C-banding. Therefore, for a more accurate description of chromosomal rearrangements, it is optimal to use both molecular and cytological methods. This has been demonstrated in different studies where hybrid forms from crossing different types of wheat, barley and triticale were investigated (Silkova et al., 2006; Mattera et al., 2015; Adonina et al., 2022).

The results of marker analysis indicate that the level of polymorphism and information content of SNP markers is lower compared to SSR markers (see Supplementary Material 1). This is supported by literature data obtained on various plant species (Singh et al., 2013; Garcia et al., 2018; Tereba, Konecka, 2021). However, a decrease in information content is compensated by a higher level of coverage of molecular genetic maps of chromosomes with SNP markers and a low level of null alleles in distant species.

A high level of SSR marker polymorphism is noted in most studies on the genetic diversity of varieties and hybrid forms of cereal crops (Jlassi et al., 2021; Pour-Aboughadareh et al., 2022). According to many authors, the amplification of the D-genome specific SSR markers in the genomes of alien species is 10–30 % (Salina et al., 2006; El-Rawy, Hassan, 2021).

It can be noted that recombination events in lines developed with the involvement of wild emmer occurred much more often in the long arms of chromosomes, which is consistent with the data obtained in the study of SNP polymorphism of 445 recombinant lines from crossing durum wheat with a sample of wild emmer (Jorgensen et al., 2017).

The high frequency of introgressions of the genetic material of *T. dicoccum* and *T. dicoccoides* into the genome of *T. aestivum* revealed in this work is a consequence of the similarity of the subgenomes of common wheat and emmer. Studies by a large group of scientists on the comparative analysis of the A-, B-, D-genomes of common wheat and its diploid and tetraploid relatives showed a high degree of homology between the corresponding genomes of related species (Petersen et al., 2006; IWGSC, 2014). At the same time, obvious differences between the species of the genus *Triticum* have been identified in the form of the loss of genetic material, the formation of new genes, and duplications resulting from evolutionary processes (IWGSC, 2014; Bariah et al., 2020). In addition, some studies have found silencing or alteration of gene function (Ozkan et al., 2001; Kashkush et al., 2002; Feldman, Levy, 2012). During the cultivation of wild emmer, changes were found both in morphological features and in the structure of the genome. For example, the genome size of domesticated emmer was slightly reduced compared to wild emmer (12.87 and 12.91 pg, respectively) (Eilam et al., 2008). All this leads to differences in the nucleotide sequence in the homologous chromosomes of *T. aestivum* and *T. dicoccoides*, which in turn affects the frequency of recombination events.

In our study, the lowest level of introgression was noted for chromosome 2A; for two out of five lines, no rearrangements were found in chromosomes 3A (see Table 1), which is consistent with the literature data. Thus, the analysis of the nucleotide polymorphism of the A- and B-genomes of *T. aestivum* and *T. dicoccoides* revealed significant differences between the 2A chromosomes of these related species. The average nucleotide polymorphism in this chromosome of *T. aestivum* and *T. dicoccoides* was 0.56 and 0.83, respectively, and the average number of haplotypes per locus was 1.85 and 2.25, respectively (Akhunov et al., 2010). Significant differences were also found between chromosomes 3A and 4A of *T. aestivum* and *T. dicoccoides* compared to other chromosomes (Akhunov et al., 2010). It should be noted that a comparative analysis of the subgenomes of common wheat and diploid and tetraploid relatives of *T. aestivum* showed some differences in the sequence of genes on chromosomes 2A and 7B, which suggests the presence of small translocations or introgressions that occurred during evolution (IWGSC, 2014).

Taking into account the homoeology of the A- and B-genomes of *T. aestivum* with similar genomes of *T. dicoccoides* and *T. dicoccum*, one should expect a rather high level of chromosome pairing in metaphase I of meiosis of F_1 hybrids followed by the formation of reciprocal exchanges between the regions of homologous wheat and emmer chromosomes. Analysis of the metaphase I stage revealed a high level of bivalent pairing of chromosomes in all studied F_{10} introgression lines (see Table 2).

It is known that the long arm of chromosome 5B contains the *Ph1* locus, which is the main regulator of chromosome synapsis and prevents mating of homoeologs (Riley, Chapman, 1958; Naranjo, 2012). The lack of activity of this locus in diploid relatives of wheat indicates its occurrence as a result of structural changes in chromosome 5B after polyploidization (Chapman, Riley, 1970). *Ph1* candidate gene (*C-Ph1*) has been identified, whose silencing resulted in formation of multivalents (Bhullar et al., 2014). Three homoeologous copies of *C-Ph1* were found on chromosomes 5A, 5B and 5D.

The nucleotide sequence of homoeologous genes has a similarity of about 90 % and differs, as a rule, by insertions and deletions, which lead to changes in the amino acid sequence of the protein. In addition, significant differences were found in the level of expression of homoeologous genes during different stages of the meiotic cycle. For *C-Ph1* on chromosome 5B, the highest level of activity was noted during metaphase I, the lowest level was observed during anaphase I and the absence of activity at subsequent stages. The 5A copy was expressed during anaphase I, dyad and tetrad stages. The 5D copy showed the highest activity during early stages of meiosis (interphase and prophase stages) (Bhullar et al., 2014). In our study, all lines, with the exception of line 13-3, revealed structural changes in the long arm of chromosome 5B; however, as a rule, these transformations do not cause significant weakening of homologue synapsis. A negative effect of chromatin introgression of the wild emmer 5B chromosome is most pronounced in line 11-1, which may be due to the length of the translocated fragment (Supplementary Material 2).

For lines developed with the involvement of domesticated emmer, a higher meiotic index was noted than for lines based on wild emmer. The maximum meiotic index was found for lines 2-7 and 1-3 of *T. dicoccum* k-45926 × Festivalnaya (see

Table 3). This is probably due to the closer similarity between the A and B genomes of *T. aestivum* and *T. dicoccum*. The highest stability among the lines with the insertion of the wild emmer genetic material was found for line 29 (93.0 %), the level of chromosome pairing of which was one of the highest (see Table 2). Line 16-5 with 100 % of the number of chromosomes included in bivalents was characterized by a lower value for this indicator (89.09 %). This fact can be explained by differences in the number of introgressed fragments in lines 29 and 16-5. Thus, line 29 contains fragments of the foreign genome in 5 chromosomes, and line 16-5 in 12 chromosomes.

Correlations between the number of alien fragments and the meiosis abnormality degree are also shown in the publications of other researchers (Gordeeva et al., 2009; Zeng et al., 2013). In addition, one should not exclude an influence of the genotype of the original wheat variety on the cytological stability of the studied lines. Thus, the meiotic index of the Rassvet variety (99.0 %) was higher than that of the Festivalnaya variety (84.55 %), and line 29 developed with the involvement of a more stable variety had the highest percentage of normal tetrads among introgression lines (see Table 3).

All lines with a high meiotic index contain alien fragments in chromosomes 1B, 2B, and 3B; most lines – in chromosomes 5B and 6B and are of interest for expanding of the wheat gene pool. For example, a functional allele of the *Gpc-B1* gene associated with a high content of protein and some microelements in wheat grain were found on chromosome 6B of wild emmer (Uauy et al., 2006). In wild emmer accessions, genes associated with an increased content of total grain protein were also found on chromosomes 2A, 5B, 6B, and 7B (Ohm et al., 2010). Powdery mildew resistance genes were mapped on chromosomes 2B (Zhang et al., 2010) and 5B (Xue et al., 2012) of wild emmer, and drought resistance genes were mapped to 5B loci (Akpinar et al., 2015). Currently, using the GWAS and RNA-seq approach, genetic factors associated with the accumulation of protein and minerals in wild emmer have been identified (Liu et al., 2021; Gong et al., 2022).

Conclusion

Thus, the results of molecular cytogenetic analysis of introgression wheat lines indicate that the insertion of the genetic material of wild and domesticated emmer, whose genomes are homologous to common wheat genomes, occurs mainly in the form of whole arms or large fragments in the chromosomes of the B-genome and less extended inserts in genome A. At the same time, the highest frequency of introgression emmer genome was observed in chromosomes 1A, 1B, 2B, and 3B.

The use of karyotyping methods in combination with screening of hybrid lines with molecular markers makes it possible to obtain extensive information on chromosomal rearrangements and the sizes of alien introgressed fragments. Analysis of the final stage of meiosis showed a high level of cytological stability in the vast majority of studied wheat lines. It should be noted that the lines characterized by an insignificant number of anomalies at the early stages of microsporogenesis, as a rule, had a higher value of the meiotic index. Introgression lines with a normal course of meiosis will be used in further studies (identification of genes that control resistance to biotic and abiotic stressors, high grain quality, etc.) in order to identify promising material for inclusion in the breeding process.

References

Adonina I.G., Shcherban A.B., Zorina M.V., Mehdiyeva S.P., Timonova E.M., Salina E.A. Genetic features of triticale-wheat hybrids with vaviloid-type spike branching. *Plants*. 2022;11(1):58. DOI 10.3390/plants11010058.

Akhunov E.D., Akhunova A.R., Anderson O.D., Anderson J.A., Blake N., Clegg M.T., Coleman-Derr D., Conley E.J., Crossman C.C., Deal K.R., Dubcovsky J., Gill B.S., Gu Y.Q., Hadam J., Heo H., Huo N., Lazo G.R., Luo M.C., Ma Y.Q., Matthews D.E., McGuire P.E., Morrell P.L., Qualset C.O., Renfro J., Tabanao D., Talbert L.E., Tian C., Toleno D.M., Warburton M.L., You F.M., Zhang W., Dvorak J. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics.* 2010;11:702. DOI 10.1186/1471-2164-11-702.

Akpinar B.A., Yuce M., Lucas S., Vrána J., Burešová V., Doležel J., Budak H. Molecular organization and comparative analysis of chromosome 5B of the wild wheat ancestor *Triticum dicoccoides. Sci. Rep.* 2015;5:10763. DOI 10.1038/srep10763.

Badaeva E.D., Sozinova L.F., Badaev N.S., Muravenko O.V., Zelenin A.V. "Chromosomal passport" of *Triticum aestivum* L. em Thell. cv. Chinese Spring and standartization of chromosomal analysis of cereals. *Cereal Res. Commun.* 1990;18(4):273-281.

Badaeva E.D., Badaev N.S., Gill B.S., Filatenko A.A. Intraspecific karyotype divergence in *Triticum araraticum* (*Poaceae*). *Plant Syst. Evol.* 1994;192:117-145. DOI 10.1007/BF00985912.

Badaeva E.D., Dedkova O.S., Gay G., Pukhalskyi V.A., Zelenin A.V., Bernard S., Bernard M. Chromosomal rearrangements in wheat: their types and distribution. *Genome.* 2007;50(10):907-926. DOI 10.1139/g07-072.

Bariah I., Keidar-Friedman D., Kashkush K. Identification and characterization of large-scale genomic rearrangements during wheat evolution. *PLoS One.* 2020;15(4):e0231323. DOI 10.1371/journal.pone.0231323.

Bhullar R., Nagarajan R., Bennypaul H., Sidhu G.K., Sidhu G., Rustgi S., von Wettstein D., Gill K.S. Silencing of a metaphase I-specific gene results in a phenotype similar to that of the Pairing homeologous 1 (*Ph1*) gene mutations. *Proc. Natl. Acad. Sci. USA.* 2014;111(39):14187-14192. DOI 10.1073/pnas.1416241111.

Budak H., Kantar M., Kurtoglu K.Y. Drought tolerance in modern and wild wheat. *Sci. World J.* 2013;2013:548246. DOI 10.1155/2013/548246.

Cakmak I., Torun A., Millet E., Feldman M., Fahima T., Korol A., Nevo E., Braun H.I., Özkan H. *Triticum dicoccoides*: an important genetic resource for increasing zinc and iron concentration in modern cultivated wheat. *Soil Sci. Plant Nutr.* 2004;50(7):1047-1054. DOI 10.1080/00380768.2004.10408573.

Chapman V., Riley R. Homoeologous meiotic chromosome pairing in *Triticum aestivum* in which chromosome 5*B* is replaced by an alien homoeologue. *Nature.* 1970;226(5243):376-377. DOI 10.1038/226376a0.

Dedkova O.S., Badaeva E.D., Mitrofanova O.P., Bilinskaya E.N., Pukhalskiy V.A. Analysis of intraspecific diversity of cultivated emmer *Triticum dicoccum* (Schrank) Schuebl. using C-banding technique. *Russ. J. Genet.* 2007;43(11):1271-1285. DOI 10.1134/S1022795407110105.

Dedkova O.S., Badaeva E.D., Amosova A.V., Martynov S.P., Ruanet V.V., Mitrofanova O.P., Pukhal'skiy V.A. Diversity and the origin of the European population of *Triticum dicoccum* (Schrank) Schuebl. as revealed by chromosome analysis. *Russ. J. Genet.* 2009;45(9):1082-1091. DOI 10.1134/S1022795409090099.

Dvorak J., Terlizzi P., Zhang H.B., Resta P. The evolution of polyploidy wheats: identification of the A genome donor species. *Genome.* 1993;36(1):21-31. DOI 10.1139/g93-004.

Eilam T., Anikster Y., Millet E., Manisterski J., Feldman M. Nuclear DNA amount and genome downsizing in natural and synthetic allopolyploids of the genera *Aegilops* and *Triticum*. *Genome.* 2008;51(8):616-627. DOI 10.1139/G08-043.

El-Rawy M.A., Hassan M.I. Assessment of genetic diversity in durum and bread wheat genotypes based on drought tolerance and SSR markers. *Plant Breed. Biotech.* 2021;9(2):89-103. DOI 10.9787/PBB.2021.9.2.89.

Feldman M., Levy A. Genome evolution due to allopolyploidization in wheat. *Genetics.* 2012;192(3):763-774. DOI 10.1534/genetics.112.146316.

García C., Guichoux E., Hampe A. A comparative analysis between SNPs and SSRs to investigate genetic variation in a juniper species (*Juniperus phoenicea* ssp. *turbinata*). *Tree Genet. Genomes.* 2018; 14:87. DOI 10.1007/s11295-018-1301-x.

Gong F., Qi T., Zhang T., Lu Y., Liu J., Zhong X., He J., Li Y., Zheng Y., Liu D., Huang L., Wu B. Comparison of the agronomic, cytological, grain protein characteristics, as well as transcriptomic profile of two wheat lines derived from wild emmer. *Front. Genet.* 2022;12: 804481. DOI 10.3389/fgene.2021.804481.

Gordeeva E.I., Leonova I.N., Kalinina N.P., Salina E.A., Budashkina E.B. Comparative cytological and molecular analysis of common wheat introgression lines containing genetic material of *Triticum timopheevii* Zhuk. *Russ. J. Genet.* 2009;45(12):1428-1437. DOI 10.1134/S1022795409120047.

International Wheat Genome Sequencing Consortium (IWGSC). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science.* 2014;345(6194):1251788. DOI 10.1126/science.1251788.

Jaradat A.A. Wheat landraces: a mini review. *Emir. J. Food Agric*. 2013;25(1):20-29. DOI 10.9755/ejfa.v25i1.15376.

Jlassi I., Bnejdi F., Saadoun M., Hajji A., Mansouri D., Ben-Attia M., El-Gazzah M., El-Bok S. SSR markers and seed quality traits revealed genetic diversity in durum wheat (*Triticum durum* Desf.). *Mol. Biol. Rep.* 2021;48(4):3185-3193. DOI 10.1007/s11033-021-06385-y.

Jorgensen C., Luo M.-C., Ramasamy R., Dawson M., Gill B.S., Korol A.B., Distelfeld A., Dvorak J. A high-density genetic map of wild emmer wheat from the Karaca Dağ region provides new evidence on the structure and evolution of wheat chromosomes. *Front. Plant Sci.* 2017;8:1798. DOI 10.3389/fpls.2017.01798.

Kashkush K., Feldman M., Levy A.A. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics.* 2002; 160(4):1651-1659. DOI 10.1093/genetics/160.4.1651.

Liu J., Huang L., Wang C., Liu Y., Yan Z., Wang Z., Xiang L., Zhong X., Gong F., Zheng Y., Liu D., Wu B. Genome-wide association study reveals novel genomic regions associated with high grain protein content in wheat lines derived from wild emmer wheat. *Front. Plant Sci.* 2019;10:464. DOI 10.3389/fpls.2019.00464.

Liu J., Huang L., Li T., Liu Y., Yan Z., Tang G., Zheng Y., Liu D., Wu B. Genome-wide association study for grain micronutrient concentrations in wheat advanced lines derived from wild emmer. *Front. Plant Sci.* 2021;12:651283. DOI 10.3389/fpls.2021.651283.

Maccaferri M., Ricci A., Salvi S., Milner S.G., Noli E., Martelli P.L., Casadio R., Akhunov E., Scalabrin S., Vendramin V., Ammar K., Blanco A., Desiderio F., Distelfeld A., Dubcovsky J., Fahima T., Faris J., Korol A., Massi A., Mastrangelo M.A., Morgante M., Pozniak C., N'Diaye A., Xu S., Tuberosa R. A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 2015; 13(5):648-663. DOI 10.1111/pbi.12288.

Mattera M.G., Ávila C.M., Atienza S.G., Cabrera A. Cytological and molecular characterization of wheat-*Hordeum chilense* chromosome 7H^{ch} introgression lines. *Euphytica.* 2015;203:165-176. DOI 10.1007/s10681-014-1292-0.

Naranjo T. Finding the correct pathner: the meiotic courtship. *Scientfica (Cairo).* 2012;2012:509073. DOI 10.6064/2012/509073.

Nevo E., Chen G. Drought and salt tolerances in wild relatives for wheat and barley improvement. *Plant Cell Environ*. 2010;33(4): 670-685. DOI 10.1111/j.1365-3040.2009.02107.x.

Ohm J.B., Klindworth D.L., Hareland G.A., Faris J.D., Elias E.M., Xu S.S. Variation in kernel characteristics and protein molecular weight distribution of Langdon durum wild emmer wheat chromosome substitution lines. *J. Cereal Sci.* 2010;52(2):207-214. DOI 10.1016/j.jcs.2010.05.007.

Orlovskaya O.A., Koren L.V., Khotyleva L.V. Cytological characteristic of wheat hybrids produced by remote hybridization of the *Triticeae* tribe. *Izvestiya Natsionalnoy Akademii Nauk Belarusi. Seriya Biologicheskikh Nauk = Proceedings of the National Academy of Sciences of Belarus. Biological Series*. 2010;4:50-54. (in Russian)

Orlovskaya O., Dubovets N., Solovey L., Leonova I. Molecular cytological analysis of alien introgressions in common wheat lines derived from the cross of *Triticum aestivum* with *T. kiharae. BMC Plant Biol.* 2020;20(Suppl. 1):201. DOI 10.1186/s12870-020-02407-2.

Ozkan H., Levy A., Feldman M. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops–Triticum*) group. *Plant Cell.* 2001;13(8):1735-1747. DOI 10.1105/TPC.010082.

Peleg Z., Fahima T., Abbo S., Krugman T., Nevo E., Yakir D., Saranga Y. Genetic diversity for drought resistance in wild wheat and its ecogeographical association. *Plant Cell Environ.* 2005;28(2):176-191. DOI 10.1111/j.1365-3040.2005.01259.x.

Peng J.H., Fahima T., Röder M.S., Huang Q.Y., Dahan A., Li Y.C., Grama A., Nevo E. High-density molecular map of chromosome region harboring stripe-rust resistance genes *YrH52* and *Yr15* derived from wild emmer wheat, *Triticum dicoccoides. Genetica.* 2000; 109(3):199-210. DOI 10.1023/a:1017573726512.

Peng J.H., Sun D.H., Nevo D. Domestication evolution, genetics and genomics in wheat. *Mol. Breed.* 2011;28:281-301. DOI 10.1007/s11032-011-9608-4.

Pestsova E., Ganal M.W., Röder M.S. Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome.* 2000;43(4):689-697. DOI 10.1139/gen-43-4-689.

Petersen G., Seberg O., Yde M., Berthelsen K. Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Mol. Phylogenet. Evol*. 2006;39(1):70-82. DOI 10.1016/j.ympev.2006.01.023.

Pour-Aboughadareh A., Poczai P., Etminan A., Jadidi O., Kianersi F., Shooshtari L. An analysis of genetic variability and population structure in wheat germplasm using microsatellite and gene-based markers. *Plants.* 2022;11(9):1205. DOI 10.3390/plants11091205.

Riley R., Chapman V. Genetic control of cytologically diploid behavior of hexaploid wheat. *Nature.* 1958;182:713-715. DOI 10.1038/182713a0.

Röder M.S., Korzun V., Wendehake K., Plaschke J., Tixier M.H., Leroy P., Ganal M.W. A microsatellite map of wheat. *Genetics.* 1998; 149(4):2007-2023. DOI 10.1093/genetics/149.4.2007.

Salina E.A., Leonova I.N., Efremova T.T., Röder M.S. Wheat genome structure: translocations during the course of polyploidization. *Funct. Integr. Genomics.* 2006;6(1):71-80. DOI 10.1007/s10142-005-0001-4.

Silkova O.G., Dobrovolskaya O.B., Dubovets N.I., Adonina I.G., Kravtsova L.A., Roeder M.S., Salina E.A., Shchapova A.I., Shumny V.K. Production of wheat-rye substitution lines and identification of chromosome composition of karyotypes using C-banding, GISH, and SSR markers. *Russ. J. Genet.* 2006;42(6):645-653. DOI 10.1134/S1022795406060093.

Singh N., Choudhury D.R., Singh A.K., Kumar S., Srinivasan K., Tyagi R.K., Singh N.K., Singh R. Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PLoS One.* 2013;8(12):e84136. DOI 10.1371/journal.pone.0084136.

Skolotneva E.S., Leonova I.N., Bukatich E.Yu., Salina E.A. Methodical approaches to identification of effective wheat genes providing broad-spectrum resistance against fungal diseases. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2017;21(7):862-869. DOI 10.18699/VJ17.307. (in Russian)

Somers D.J., Isaac P., Edwards K. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 2004;109(6):1105-1114. DOI 10.1007/s00122-004-1740-7.

Sosnikhina S.P., Mikhailova E.I., Tikholiz O.A., Tsvetkova N.V., Lovtsyus A.V., Sapronova O.S., Fedotova Yu.S., Kolomiets O.L., Bogdanov Yu.F. Expression and inheritance of a desynaptic phenotype with impaired homologous synapsis in rye. *Russ. J. Genet.* 2007; 43(10):1193-1200. DOI 10.1134/S1022795407100146.

Surzhikov S.A., Knüpffer H., Kilian B. Chromosomal rearrangements in wheat: their types and distribution. *Genome.* 2007;50(10):907-926. DOI 10.1139/G07-072.

Tereba A., Konecka A. Comparison of microsatellites and SNP markers in genetic diversity level of two scots pine stands. *Environ. Sci. Proc.* 2021;3(1):4. DOI 10.3390/IECF2020-07776.

Uauy C., Distelfeld A., Fahima T., Blechl A., Dubcovsky J. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science.* 2006;314(5803):1298-1301. DOI 10.1126/science.1133649.

Wang S., Wong D., Forrest K., Allen A., Chao S., Huang B.E., Maccaferri M., … Tuberosa R., Lawley C., Mikoulitch I., Cavanagh C., Edwards K.J., Hayden M., Akhunov E. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 2014;12(6):787-796. DOI 10.1111/pbi.12183.

Wang Z., Huang L., Wu B., Hu J., Jiang Z., Hu J., Qi P., Zheng Y., Liu D. Characterization of an integrated active *Glu-1Ay* allele in common wheat from wild emmer and its potential role in flour improvement. *Int. J. Mol. Sci.* 2018;19(4):923. DOI 10.3390/ijms19040923.

Xie W., Nevo E. Wild emmer: genetic resources, gene mapping and potential for wheat improvement. *Euphytica.* 2008;164:603-614. DOI 10.1007/s10681-008-9703-8.

Xue F., Ji W., Wang C., Zhang H., Yang B. High-density mapping and marker development for the powdery mildew resistance gene *PmAS846* derived from wild emmer wheat (*Triticum turgidum* var. *dicoccoides*). *Theor. Appl. Genet.* 2012;124(8):1549-1560. DOI 10.1007/s00122-012-1809-7.

Zeng J., Cao W., Hucl P., Yang Y., Xue A., Chi D., Fedak G. Molecular cytogenetic analysis of wheat – *Elymus repens* introgression lines with resistance to Fusarium head blight. *Genome.* 2013;56(1):75-82. DOI 10.1139/gen-2012-0130.

Zhang H., Guan H., Li J., Zhu J., Xie C., Zhou Y., Duan X., Yang T., Sun Q., Liu Z. Genetic and comparative genomics mapping reveals that a powdery mildew resistance gene *Ml3D232* originating from wild emmer co-segregates with an NBS-LRR analog in common wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 2010;121(8):1613-1621. DOI 10.1007/s00122-010-1414-6.

Zhou W., Jiang Y., Zhang W., Xu G., Rong J. Characterization of large chromosome segment introgressions from *Triticum turgidum* subsp. *dicoccoides* into bread wheat with simple sequence repeat markers. *Crop Sci.* 2013;53(4):1555-1565. DOI 10.2135/cropsci2012.05.0287.

ORCID ID

O.A. Orlovskaya orcid.org/0000-0002-1187-1317
I.N. Leonova orcid.org/0000-0002-6516-0545
N.I. Dubovets orcid.org/0000-0002-5980-181X

Acknowledgements. Cytological analysis was done with financial support of the Belarusian Republican Foundation for Fundamental Research (project B18R-028). Analysis of introgression lines with SSR and SNP markers was supported by budgetary project FWNR-2022-0017 of the Federal Research Center "Institute of Cytology and Genetics SB RAS".

Conflict of interest. The authors declare no conflict of interest.

Received April 3, 2023. Revised May 24, 2023. Accepted May 26, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):565-571
DOI 10.18699/VJGB-23-68

Genome-wide association study for charcoal rot resistance in soybean harvested in Kazakhstan

A. Zatybekov[1], S. Abugalieva[1], S. Didorenko[2], A. Rsaliyev[3], A. Maulenbay[3], C. Fang[4], Y. Turuspekov[1]

[1] Institute of Plant Biology and Biotechnology, Almaty, Kazakhstan
[2] Kazakh Research Institute of Agriculture and Plant Growing, Almalybak, Almaty Region, Kazakhstan
[3] Research Institute for Biological Safety Problems, Gvardeiskiy (Otar), Zhambyl Region, Kazakhstan
[4] School of Life Sciences, Guangzhou University, Guangzhou, China
yerlant@yahoo.com

Abstract. Charcoal rot (CR) caused by the fungal pathogen *Macrophomina phaseolina* is a devastating disease affecting soybean (*Glycine max* (L.) Merrill.) worldwide. Identifying the genetic factors associated with resistance to charcoal rot is crucial for developing disease-resistant soybean cultivars. In this research, we conducted a genome-wide association study (GWAS) using different models and genotypic data to unravel the genetic determinants underlying soybean resistance to charcoal rot. The study relied on a panel of 252 soybean accessions, comprising commercial cultivars and breeding lines, to capture genetic variations associated with resistance. The phenotypic evaluation was performed under natural conditions during the 2021–2022 period. Disease severity and survival rates were recorded to quantify the resistance levels in the accessions. Genotypic data consisted of two sets: the results of genotyping using the Illumina iSelect 6K SNP (single-nucleotide polymorphism) array and the results of whole-genome resequencing. The GWAS was conducted using four different models (MLM, MLMM, FarmCPU, and BLINK) based on the GAPIT platform. As a result, SNP markers of 11 quantitative trait loci associated with CR resistance were identified. Candidate genes within the identified genomic regions were explored for their functional annotations and potential roles in plant defense responses. The findings from this study may further contribute to the development of molecular breeding strategies for enhancing CR resistance in soybean cultivars. Marker-assisted selection can be efficiently employed to accelerate the breeding process, enabling the development of cultivars with improved resistance to charcoal rot. Ultimately, deploying resistant cultivars may significantly reduce yield losses and enhance the sustainability of soybean production, benefiting farmers and ensuring a stable supply of this valuable crop.
Key words: soybean; charcoal rot; whole genome resequencing; GWAS; SNP; QTL.

For citation: Zatybekov A., Abugalieva S., Didorenko S., Rsaliyev A., Maulenbay A., Fang C., Turuspekov Y. Genome-wide association study for charcoal rot resistance in soybean harvested in Kazakhstan. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):565-571. DOI 10.18699/VJGB-23-68

Полногеномный анализ ассоциации устойчивости к пепельной гнили сои, выращенной в Казахстане

А. Затыбеков[1], С. Абугалиева[1], С. Дидоренко[2], А. Рсалиев[3], А. Мауленбай[3], Ч. Фанг[4], Е. Туруспеков[1]

[1] Институт биологии и биотехнологии растений, Алматы, Казахстан
[2] Казахский научно-исследовательский институт земледелия и растениеводства, пос. Алмалыбак, Алматинская область, Казахстан
[3] Научно-исследовательский институт проблем биологической безопасности, пос. Гвардейский (Отар), Жамбылская область, Казахстан
[4] Школа наук о жизни, Университет Гуанчжоу, Гуанчжоу, Китайская Народная Республика
yerlant@yahoo.com

Аннотация. Пепельная гниль, вызываемая грибным патогеном *Macrophomina phaseolina*, представляет собой опасное заболевание, поражающее сою (*Glycine max* (L.) Merrill.) во всем мире. Выявление генетических факторов, связанных с устойчивостью к пепельной гнили, имеет важное значение для создания устойчивых к болезням сортов сои. Мы провели полногеномный анализ ассоциации (ПГАА) с использованием различных моделей и генотипических данных, чтобы найти генетические детерминанты, лежащие в основе устойчивости сои к пепельной гнили. В исследовании использовали коллекцию, состоящую из 252 образцов сои, включая коммерческие сорта и селекционные линии, для выявления генетических вариаций, связанных с устойчивостью. Фенотипическую оценку проводили в естественных условиях в период 2021–2022 гг. В работе регистрировали уровень заболевания и показатели выживаемости для количественной оценки уровней устойчивости образцов. Генотипические данные состояли из двух наборов: результаты генотипирования с применением технологии Illumina iSelect 6K SNP, и данные полногеномного ресеквенирования. Полногеномный анализ ассоциации был выполнен с помощью четырех различных моделей (MLM, MLMM, FarmCPU и BLINK) на платформе GAPIT. В результате были идентифицированы SNP-маркеры 11 локусов количественных признаков, ассоциированных с устойчивостью к пепельной гнили. Гены-кандидаты в пределах идентифицированных геномных областей были изучены на предмет их функциональной аннотации и потенциальной роли в защитных реакциях растений. Результаты этого

исследования могут внести дополнительный вклад в разработку стратегий молекулярной селекции для повышения устойчивости сортов сои к пепельной гнили. Маркер-опосредованный отбор может быть эффективно применен для ускорения процесса селекции, что позволит создавать сорта с повышенной устойчивостью к пепельной гнили. Использование устойчивых сортов может значительно сократить потери урожая и повысить устойчивость производства сои, что принесет пользу фермерам и обеспечит стабильное производство этой ценной культуры.
Ключевые слова: соя; пепельная гниль; полногеномное ресеквенирование; ПГАА; SNP; ЛКП.

Introduction

Soybean (*Glycine max* (L.) Merrill.) is one of the most important legumes in the world due to the high nutritional value and protein content of seeds (Pratap et al., 2016). According to the Agencies for Strategic Planning and Reforms of the Republic of Kazakhstan Bureau of National Statistics, soybean was grown on 127.7 thousand hectares in Kazakhstan in 2022 (https://new.stat.gov.kz/ru/industries/business-statistics/stat-forrest-village-hunt-fish/publications/5099/). To further develop the soybean industry in Kazakhstan, the Government of Kazakhstan has announced a new program known as "Northern Soybean" to expand the soybean area to 1.5 million hectares (https://www.gov.kz/article/64601?lang=kk).

The important factor severely limiting soybean productivity is its susceptibility to harmful fungal diseases (Wrather et al., 2010; Bandara et al., 2020). Strategies for the management of soybean fungal diseases include cultural methods, seed-applied fungicides, and biological controls, but these have not been effective or widely adopted, and have provided limited control (Akem, 1996; Hartman et al., 2015). Therefore, genetic resistance may be the most feasible and sustainable method by which to manage fungal diseases (Lin et al., 2022). Breeding for resistance is difficult because most diseases are quantitatively inherited and controlled by multiple genes. However, modern genomic methodologies may help elucidate resistance mechanisms and identify resistant genotypes to improve breeding programs (St. Clair, 2010). For instance, breeding projects can be coupled with genome-wide association studies (GWASs), as this approach can efficiently identify candidate genes for disease screening.

Genome-wide association studies are becoming a routine approach in the search for marker-trait associations (Korte, Farlow, 2013) and can be efficiently applied to assess genetic variations of important agronomic traits, including disease resistance (Iquira et al., 2015). These studies use high-density single-nucleotide polymorphism (SNP) arrays and variable populations to enhance the mapping resolution, which drastically improves the identification of putative causal genes (Song et al., 2013; Zhang et al., 2015). Although there are possibilities in respect of GWASs for the prediction of false positive associations, applying variable statistical algorithms can be instrumental in controlling these. For instance, a mixed linear model (MLM) that uses population structure and kinship matrices will significantly regulate inflation (Kaler et al., 2020). However, this method may also remove true genes as background noise in certain studies of complex traits associated with the population structure. To overcome this obstacle, the Bayesian information and linkage-disequilibrium iteratively nested keyway (BLINK) employs a multiple-loci test method along with a fixed-effect model (FEM), Bayesian information criteria, and linkage disequilibrium information (Huang et al., 2019). Our previous GWAS of 182 soybean accessions for resistance to fungal diseases (Zatybekov et al., 2018) allowed the identification of 15 marker-trait associations (MTAs) for resistance to fusarium root rot (FUS, caused by *Fusarium* spp.), frogeye leaf spot (FLS, caused by *Cercospora sojina*), and brown spot (BS, caused by *Septoria glycines*).

In Kazakhstan, more than ten soybean fungal diseases have been identified (Mombekova et al., 2013; Didorenko et al., 2014; Zatybekov et al., 2018), and one category of these comprises root rot diseases (charcoal rot, phytophthora root rot, fusarium root rot, etc.). The expansion of these studies is an obvious necessity in order to examine the genetic background associated with the resistance to harmful pathogens in soybean. One of the local harmful root rot diseases is charcoal rot (CR) caused by *Macrophomina phaseolina*, a soil- and seed-borne polyphagous fungus (Paris et al., 2006).

Currently, only partial resistance has been recorded in soybean, while complete resistance to *M. phaseolina* has not been reported in any plant species (Paris et al., 2006; Pawlowski et al., 2015). In addition, it has been reported that no similar markers or genes have been found when comparing field and greenhouse studies, suggesting that CR resistance in soybean has a complex molecular mechanism. Therefore, the search for more valuable resistance sources from which to identify resistance genes should be continued. The main purpose of this study was to identify MTAs for charcoal rot resistance in a collection of 252 accessions from major soybean growing regions from around the world using the GWAS approach.

Materials and methods

The soybean collection analyzed in this study consisted of 252 accessions, including 31 released cultivars and prospective breeding lines from Kazakhstan (Zatybekov et al., 2020). Accessions from 20 countries were represented in the collection and separated into five origin groups: Western and Eastern Europe, North America, East Asia, and Kazakhstan (Zatybekov et al., 2020).

The collection was assessed in the experimental plots of the Kazakh Research Institute of Agriculture and Plant Growing (south-eastern Kazakhstan) in the 2021–2022 period. Plants were grown in one-meter-long rows with a 30 cm distance between adjacent rows and 5 cm between plants within rows. The assessment of field resistance to CR was based on a generally accepted five-point scale: 1 – microsclerotia are not visible in the tissue (I-immune); 2 – very few microsclerotia are visible in the core, vascular tissue, or under the epidermis, and the vascular tissue has not changed color (R-resistant); 3 – vascular tissue is partially discolored, and microsclerotia partially cover the tissue (MR-moderately resistant); 4 – vascular tissue is discolored with numerous microsclerotia embedded in the tissue, and microsclerotia are also visible under the outer epi-

dermis in stem and root sections (MS-moderately susceptible); and 5 – vascular tissue darkened due to the large amount of microsclerotia both inside and outside the tissues of the stem and root (S-susceptible) (Mengistu et al., 2007). The field experiments for CR resistance were conducted in triplicate and in randomized order. The results were analyzed using Statistical Package for the Social Sciences (SPSS 22.0.0.0) (https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software) computer programs.

The genotyping data consisted of two sets of SNP data. The first set (Set 1) was developed using the soybean 6K SNP Illumina iSelect array (Song et al., 2013) at Traitgenetics GmbH (Gatersleben, Germany). DNA samples were extracted and purified from single seeds of individual cultivars using commercial kits (Qiagen, CA, USA). The DNA concentration for each sample was adjusted to 50 ng/µl. SNP genotype analysis was carried out using Illumina Genome Studio software (GS V2011.1). The second set (Set 2) was developed at the Department of the School of Life Sciences, Guangzhou University, China, using whole-genome resequencing (WGRS) technology based on the Illumina HiSeq X Ten system (Lu et al., 2020). For each of the accessions in the panel of 252, at least five µg of DNA was used to construct a sequencing library with an Illumina TruSeq DNA Sample Prep Kit, according to the manufacturer's instructions.

The SNP datasets were filtered using a 10 % cutoff for missing data, and markers with minor allele frequency ≥ 0.05 were considered for the genome-wide association studies. Numbers of hypothetical groups ranging from $k = 1$ to 10 were assessed using 50,000 burn-in iterations followed by 100,000 recorded Markov-chain iterations. The sampling variance of the population structure inference was estimated for each k using STRUCTURE software (Pritchard et al., 2000) with five independent runs. The delta K value (ΔK) was estimated using Structure Harvester (Evanno et al., 2005). The Q-matrix was developed based on the final k-values. Population genetic analysis was conducted using two sets of SNP data to construct a neighbor-joining tree with TASSEL software and further visualization using the iTOL online platform (https://itol.embl.de/).

The GWAS for soybean resistance to CR in Southeast Kazakhstan was conducted using MLM (Yu et al., 2006), a multiple-locus mixed linear model (MLMM) (Segura et al., 2012), fixed and random model circulating probability unification (FarmCPU) (Liu et al., 2016), and BLINK models (Huang et al., 2019) using GAPIT V3 software (Wang, Zhang, 2021). The rMVP package (Yin et al., 2021) was used for the visualization with $P \leq 0.0002$ thresholds. A QQ plot was applied to evaluate the distribution of observed p-values compared to the expected distribution under the null hypothesis of no association between genetic markers and the trait of interest (Ehret, 2010).

Results

Disease resistance

Field trial results obtained at the experimental stations of the Southeast region suggested a clear difference in the development of charcoal rot. During the two-year period, the spread and effect of CR pathogens on plants were stronger in 2021

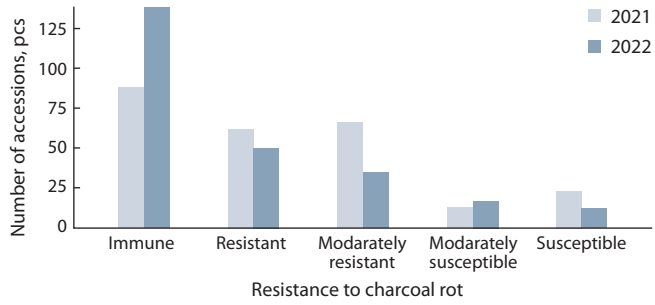

Fig. 1. Resistance of the soybean collection to charcoal rot in the Southeast region of Kazakhstan.

(Fig. 1, Supplementary Material 1)[1]. The results indicate that the group of susceptible accessions comprised 23 samples (9.1 %) in 2021 and 12 samples (4.8 %) in 2022 (see Fig. 1).

Five genotypes were susceptible to resistance to CR during the two years of the study. Among them were three accessions from East Asia, i. e., cultivars Jin nong 62, Dong doe 027, and Mei feng 18 from China, and line 1034 from Korea. In addition, two accessions from East Europe (Osobliva from Ukraine and CH 147020-1 from Belarus) were susceptible to charcoal rot.

Genetic variation in the soybean collection based on two sets of SNP markers

The final data after filtering consisted of 4495 (Set 1) and 44,385 (Set 2) polymorphic SNPs. Set 1 consisted of 77.98 % transitions and 22.02 % transverse variants, and Set 2 comprised 71.88 % transitions and 28.12 % transverse variants. The average length of chromosomes was 47.4 Mb for both sets, and the average number of SNPs per chromosome was 222.1 for Set 1 and 2219.3 for Set 2. The chromosome length ranged from 37.3 Mb in Gm16 to 62.1 Mb in Gm18 for Set 1, and from 34.7 Mb in Gm11 to 58 Mb in Gm18 for Set 2. The number of markers per chromosome varied from 185 in Gm11 to 295 in Gm13 for Set 1, and from 1438 in Gm11 to 3438 in Gm18 for Set 2. The average density of the SNP map was one marker every 213 Kb for Set 1 and every 22 Kb for Set 2.

The Structure Harvester results divided the studied collection into three clusters based on data from Set 1 and Set 2 (Fig. 2).

No clear separation of accessions depending on the origin was recorded, regardless of the set chosen. The largest group of accessions was distributed in the third cluster, which consisted of 165 genotypes in Set 1 and 168 samples in Set 2. Based on two sets of genotyping data, a phylogenetic tree was constructed using the neighbor-joining method (Fig. 3). Population analysis based on resequencing data showed a clear division into four populations (see Fig. 3, Set 2).

Association mapping

The GWAS of soybean resistance to CR was conducted using two sets of genotypic data (Set 1 and Set 2) and four models (MLM, MLMM, FarmCPU, and BLINK). The results of the GWAS using four models are shown in Supplementary Material 2. The comparative results of the associations suggest that

[1] Supplementary Materials 1–4 are available in the online version of the paper: https://vavilovj-icg.ru/download/pict-2023-27/appx21.pdf

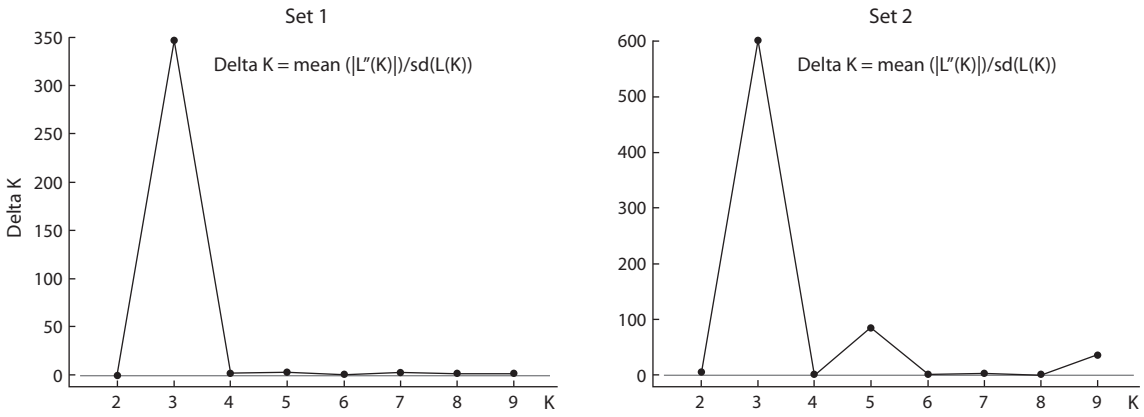

Fig. 2. Determination of delta K (ΔK) using Structure Harvester software based on two genotypic datasets.
Set 1 – data based on genotyping using Illumina iSelect array (4651 SNP); Set 2 – data based on resequencing technology (44,385 SNP).

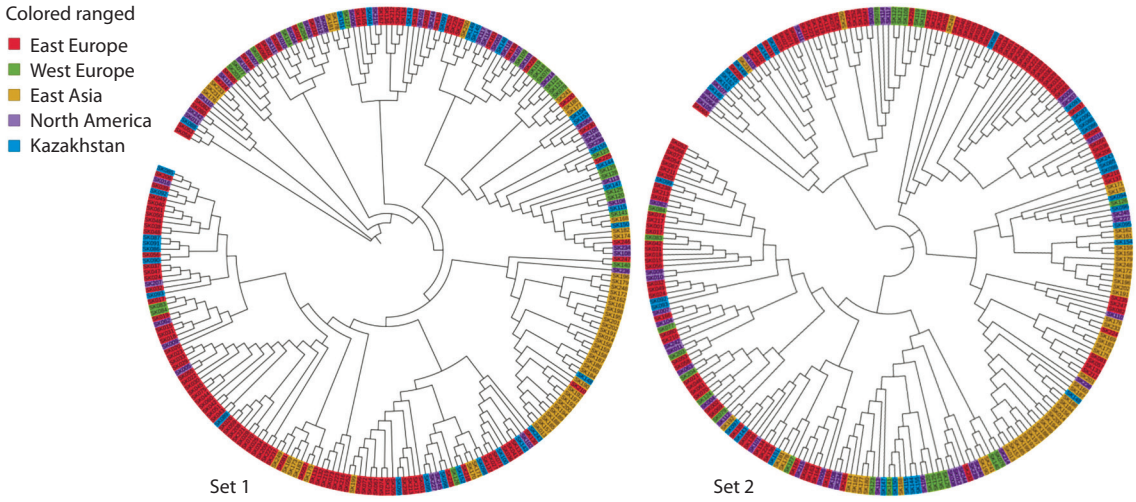

Fig. 3. Neighbor-joining tree based on two sets of genotypic data.
Set 1 – data based on genotyping using Illumina iSelect array (4651 SNP); Set 2 – data based on resequencing technology (44,385 SNP).

the most informative results were obtained using the MLMM, FarmCPU, and BLINK models (Table 1).

The application of BLINK using Set 1 allowed the identification of five quantitative trait loci (QTLs) associated with CR resistance, while the usage of Set 2 expanded the detection to 11 QTLs that were significant at the threshold of $P \leq 0.002$ (see Supplementary Material 3, Table 1). The QQ plot confirms the reliability of the associations (see Supplementary Material, 3, *b*).

The physical position of each critical SNP marker in the MTAs was overlaid with the positions of known QTLs (https://soybase.org/search/qtllist_by_symbol.php) (Table 2). Particularly, the assessment of the 11 QTLs listed in Table 2 indicated that 8 of them were reported in published reports in respect of plant resistance studies.

The largest numbers of QTLs associated with CR resistance using Set 2 were identified in chromosomes 7, 9, and 15 (Supplementary Material 4). Analysis of the genome physical locations of associated SNP markers revealed that all identified SNPs were part of the coding DNA sequence (CDS) (Table 3).

Table 1. Comparison of identified numbers of SNPs and quantitative trait loci (QTLs) according to GWAS models based on two sets of genotypic data

Genotypic data	GWAS model	Number of identified SNP markers	Number of QTLs
Set 1	MLM	5	5
	MLMM	5	5
	FarmCPU	5	5
	BLINK	5	5
Set 2	MLM	61	11
	MLMM	63	11
	FarmCPU	63	11
	BLINK	63	11

Table 2. The list of identified QTLs using the BLINK model

QTL ID	SNP ID	Chr	Position	QTL interval	*p*-value			Allele	Effect	Known QTLs*
					2021	2022	Average			
qMac.ph 2-1	S02_330064	2	330064	230834-662455	1.6161E–04	5.8456E–04	4.6828E–06	A/G	0.337	–
qMac.ph 3-1	S03_35294437	3	35294437	34928316-35364002		9.8024E–05		A/G	0.436	SCN 33-7, 44-15
qMac.ph 7-1	S07_5459756	7	5459756	5459595-5459756			2.7346E–04	A/G	–0.338	Phytoph 14-8
qMac.ph 7-2	S07_15585048	7	15585048	15585048			1.1160E–04	A/G	0.364	Sclero 10-1
qMac.ph 8-1	S08_3858712	8	3858712	3858712	1.2103E–04			G/A	–0.437	SDS 13-13, 15-2, 16-3, Fus lesion length 1-1
qMac.ph 9-1	S09_7368126	9	7368126	7368126			2,2932E–04	A/T	0.483	Sclero 1-3
qMac.ph 9-2	S09_37287856	9	37287856	36928447-37712255	4.7031E–04		2.8502E–04	A/G	0.391	Sclero 8-3, SDS 18-3
qMac.ph 15-1	S15_9318442	15	9318442	8797665-9906427	8.3066E–05			G/A	0.449	Asian rust 2-4, Fus lesion length 1-4, SDS disease index 21-2
qMac.ph 15-2	S15_49773459	15	49773459	49772578-49773474	5.9168E–05		3.6647E–04	C/T	0.259	–
qMac.ph 16-1	S16_3274861	16	3274861	2706211-3726463	2.5021E–04			G/A	–0.545	–
qMac.ph 19-1	S19_46398331	19	46398331	45583506-46401031	1.5684E–04		3.9965E–04	A/C	0.344	Phytoph 14-8

* Based on the QTL list on SoyBase (https://soybase.org/search/qtllist_by_symbol.php).

Table 3. Physical positions of identified QTLs in the soybean genome

QTL ID	Chr	QTL position interval	Genes in the QTL interval	Peak genes	Annotation of peak genes
qMac.ph 2-1	2	230834-662455	Glyma_02G001800-Glyma_G006100	Glyma_02G002900	Methyltransferase
				Glyma_02G002500*	Protein argonaute 10
qMac.ph 3-1	3	34928316-35364002	Glyma_03G134100-Glyma_G137400	Glyma_03G137000	Magnesium chelatase
				Glyma_03G135900*, Glyma_03G136400*	NB-ARC domain-containing protein
qMac.ph 7-1	7	5459595-5459756	Glyma_07G061500	Glyma_07G061500	BAG domain-containing protein
qMac.ph 7-2	7	15585048	Glyma_07G132100	Glyma_07G132100	Kinesin-like protein
qMac.ph 8-1	8	3858712	Glyma_08G049400	Glyma_08G049400	RING-type E3 ubiquitin transferase
qMac.ph 9-1	9	7368126	Glyma_09G071600	Glyma_09G071600	Protein TIFY
qMac.ph 9-2	9	36928447-37712255	Glyma_09G146900-Glyma_G154100	Glyma_09G151300	Serine/threonine-protein kinase
qMac.ph 15-1	15	8797665-9906427	Glyma_15G112100-Glyma_G124700	Glyma_15G118800	CCAAT transcription factor
				Glyma_15G118400*, Glyma_15G124500*	HMA domain-containing protein, MLO-like protein
qMac.ph 15-3	15	49772578-49773474	Glyma_15G263900	Glyma_15G263900	Tubulin beta chain
qMac.ph 16-1	16	2706211-3726463	Glyma_16G027600-Glyma_G039700	Glyma_16G034900	Receptor-like serine/threonine-protein kinase
				Glyma_16G031700*, Glyma_16G033900*	AAA + ATPase domain-containing protein, ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase
qMac.ph 19-1	19	45583506-46401031	Glyma_19G198800-Glyma_G209400	Glyma_19G209300	PPR_long domain-containing protein
				Glyma_19G198800*, Glyma_19G198900*	Late embryogenesis abundant protein LEA-2 subgroup domain-containing protein

* Genes associated with deseases resistance function.

Each SNP in an intergenic position was considered for possible functional annotation based on the proximity of closely located genes.

Discussion

The analysis of soybean CR development in Southeast Kazakhstan revealed a strong environmental influence on the distribution of pathogens and plant tolerance to the disease. Previous CR resistance studies in soybean were unsuccessful in identifying major genes that completely resist this pathogen (Coser et al., 2017), indicating that the resistance is quantitative and may rely on the efficiency of minor QTLs. Therefore, searching for resistant genotypes and applying powerful genomic tools are the obvious priorities for successful CR resistance breeding in soybean. The current work was conducted to identify QTLs associated with CR resistance in soybean using a genome-wide association studies.

The study was based on using two genotyping sets (Set 1 and Set 2) that drastically differed in the number of SNPs in the same soybean collection. Set 1 included 4651 SNPs and was extensively used in our previous GWAS projects (Zatybekov et al., 2020), while Set 2 included 44,385 SNPs, which is roughly ten times higher than the number in Set 1. As expected, the power of the GWAS using Set 2 (11 QTLs) was higher than that of Set 1 (five QTLs) (see Table 1) and relied on WRGS technology, which allows more in-depth searching and discovery of new genes associated with agronomic traits, as well as the study of evolutionary mutations in the genome. Most of the identified statistically significant QTLs were detected using the BLINK and FarmCPU models, which were applied among four different statistical approaches, including MLM and MLMM. In this study, BLINK and FarmCPU were the most successful approaches (see Supplementary Material 2), and additional estimations, including QQ plots, confirmed that these models generally result in fewer false positives and identify more true positives (Huang et al., 2019).

Assessment of the identified QTL locations showed that, in most cases, they are distant from each other, or detected on different chromosomes. Analysis of the SNPs in the identified MTAs revealed nine proteins associated with the immune response to pathogens. The qMac.ph 3-1 interval contains two genes (*Glyma_03G0135900* and *Glyma_03G0136400*), both coding NB-ARC domain-containing protein, which is associated with resistance to fungal pathogens (Van Ooijen et al., 2008). The qMac.ph 15-1 interval contains gene *Glyma_15G118400* coding HMA domain-containing protein, which is associated with regulation of the defense response to fungi.

In addition, the gene *Glyma_15G124500*, which codes MLO-like protein associated with powdery-mildew resistance (Shen et al., 2012), is also located in this interval. The qMac.ph 16-1 interval contains two genes, *Glyma_16G031700* and *Glyma_16G033900*, associated with the defense response. The late embryogenesis abundant protein LEA-2 subgroup coding two genes, *Glyma_19G198800* and *Glyma_19G198900*, was located in an interval of qMac.ph 19-1 and is associated with disease resistance. These examples of SNPs in MTAs, including presumably newly identified genetic factors that have not been matched to previously reported factors in the literature, require additional validation studies. Thus, the identified QTLs may facilitate the discovery of new genes for disease resistance and a better understanding of genotype × environment interaction patterns. The identified SNP markers (see Tables 2 and 3) for each of the detected QTLs of CR resistance can be efficiently used in marker-associated selection projects in soybean.

Conclusion

Our GWAS of soybean resistance to CR has provided important insights into the genetic determinants underlying resistance to this devastating disease. By employing two sets of genotypic data with different SNP density levels and utilizing four GWAS models, we identified eleven QTLs that were statistically associated with CR resistance. The GWAS revealed the complexity of the genetic architecture underlying resistance to CR, indicating the involvement of multiple genes and molecular pathways. The identified genomic regions can serve as valuable targets for further functional validation and exploration of their specific roles in plant defense responses against *Macrophomina phaseolina*. Overall, the study represents a significant step in understanding the genetic basis of soybean resistance to charcoal rot. The knowledge gained from this research may further contribute to developing resilient soybean cultivars, ensuring a stable and sustainable supply of this essential crop while minimizing the economic and environmental impacts of charcoal rot.

References

Akem C.N. Management of Soybean Diseases. IITA Research Guide 40. Training Program. Ibadan, Nigeria: International Institute of Tropical Agriculture (IITA), 1996.

Bandara A.Y., Weerasooriya D.K., Bradley C.A., Allen T.W., Esker P.D. Dissecting the economic impact of soybean diseases in the United States over two decades. *PLoS One*. 2020;15(4):e0231141. DOI 10.1371/journal.pone.0231141.

Coser S.M., Chowda Reddy R.V., Zhang J., Mueller D.S., Mengistu A., Wise K.A., Allen T.W., Singh A., Singh A.K. Genetic architecture of charcoal rot (*Macrophomina phaseolina*) resistance in soybean revealed using a diverse panel. *Front. Plant Sci*. 2017;8:1626. DOI 10.3389/fpls.2017.01626.

Didorenko S.V., Sagitov A.O., Kudaibergenov M.S. Main diseases on crops of soybean and methods of dealing with them. *Agroalem*. 2014;8(61):42-46. (in Russian)

Ehret G.B. Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. *Curr. Hypertens. Rep*. 2010;12:17-25. DOI 10.1007/s11906-009-0086-6.

Evanno G., Regnaut S., Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol*. 2005;14(8):2611-2620. DOI 10.1111/j.1365-294X.2005.02553.x.

Hartman G.L., Rupe J.C., Sikora E.J., Domier L.L., Davis J.A., Steffey K.L. Compendium of Soybean Diseases and Pests. St. Paul, Minnesota: The American Phytopathological Society, 2015. DOI 10.1094/9780890544754.

Huang M., Liu X., Zhou Y., Summers R.M., Zhang Z. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience*. 2019;8(2):giy154. DOI 10.1093/gigascience/giy154.

Iquira E., Humira S., Francois B. Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biol*. 2015;15:5. DOI 10.1186/s12870-014-0408-y.

Kaler A.S., Gillman J.D., Beissinger T., Purcell L.C. Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. *Front. Plant Sci*. 2020;10:1794. DOI 10.3389/fpls.2019.01794.

Korte A., Farlow A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*. 2013;9:29. DOI 10.1186/1746-4811-9-29.

Lin F., Chhapekar S.S., Vieira C.C., Da Silva M.P., Rojas A., Lee D., Liu N., Pardo E.M., Lee Y.-C., Dong Z., Pinheiro J.B., Ploper L.D., Rupe J., Chen P., Wang D., Nguyen H.T. Breeding for disease resistance in soybean: a global perspective. *Theor. Appl. Genet*. 2022; 135:3773-3872. DOI 10.1007/s00122-022-04101-3.

Liu X., Huang M., Fan B., Buckler E.S., Zhang Z. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet*. 2016;12(2):e1005767. DOI 10.1371/journal.pgen.1005767.

Lu S., Dong L., Fang C., Liu S., Kong L., Cheng Q., Chen L., Su T., Nan H., Zhang D., Zhang L., Wang Z., Yang Y., Yu D., Liu X., Yang Q., Lin X., Tang Y., Zhao X., Yang X., Tian C., Xie Q., Yuan X., Tian Z., Liu B., Weller J.L., Kong F. Stepwise selection on homeologous *PRR* genes controlling flowering and maturity during soybean domestification. *Nat. Genet*. 2020;52(4):428-436. DOI 10.1038/s41588-020-0604-7.

Mengistu A., Ray J.D., Smith J.R., Paris R.L. Charcoal rot disease assessment of soybean genotypes using a colony-forming unit index. *Crop Sci*. 2007;47(6):2453-2461. DOI 10.2135/cropsci2007.04.0186.

Mombekova G.A., Shemshurova O.N., Seitbattalova A.I., Aitkhozhina N.A., Bekmakhanova N.E. Phytopathogens of sugar beet and soybean cultivated in soil and climatic conditions of Almaty region. *Vestnik NAN RK = Bulletin of the National Academy of Sciences of the Republic of Kazakhstan.* 2013;4:8-11. (in Russian)

Paris R.L., Mengistu A., Tyler J.M., Smith J.R. Registration of soybean germplasm line DT97–4290 with moderate resistance to charcoal rot. *Crop Sci*. 2006;46(5):2324-2325. DOI 10.2135/cropsci2005.09.0297.

Pawlowski M.L., Hill C.B., Hartman G.L. Resistance to charcoal rot identified in ancestral soybean germplasm. *Crop Sci*. 2015;55(3): 1230-1235. DOI 10.2135/cropsci2014.10.0687.

Pratap A., Gupta S.K., Kumar J., Mehandi S., Pandey V.R. Soybean. In: Breeding Oilseed Crops for Sustainable Production. Ch. 12. Academic Press, 2016;293-315. DOI 10.1016/b978-0-12-801309-0.00012-4.

Pritchard J.K., Stephens M., Rosenberg N.A., Donnelly P. Association mapping in structured populations. *Am. J. Hum. Genet*. 2000;67(1): 170-181. DOI 10.1086/302959.

Segura V., Vilhjalmsson B., Platt A., Korte A., Seren U., Long Q., Nordborg M. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat. Genet*. 2012;44(7):825-830. DOI 10.1038/ng.2314.

Shen Q., Zhao J., Du C., Xiang Y., Cao J., Qin X. Genome-scale identification of MLO domain-containing genes in soybean (*Glycine max* L. Merr.). *Genes Genet. Syst*. 2012;87(2):89-98. DOI 10.1266/ggs.87.89.

Song Q., Hyten D.L., Jia G., Quigley C.V., Fickus E.W., Nelson R.L., Cregan P.B. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS One*. 2013;8(1):e54985. DOI 10.1371/journal.pone.0054985.

St. Clair D.A. Quantitative disease resistance and quantitative resistance loci in breeding. *Annu. Rev. Phytopathol*. 2010;48:247-268. DOI 10.1146/annurev-phyto-080508-081904.

Van Ooijen G., Mayr G., Kasiem M.M.A., Albrecht M., Cornelissen B.J., Takken F.L. Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.* 2008;59(6): 1383-1397. DOI 10.1093/jxb/ern045.

Wang J., Zhang Z. GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *Genom. Proteom. Bioinform*. 2021;19(4):629-640. DOI 10.1016/j.gpb.2021.08.005.

Wrather A., Shannon G., Balardin R., Carregal L., Escobar R., Gupta G.K., Ma Z., Morel W., Ploper D., Tenuta A. Effect of diseases on soybean yield in the top eight producing countries in 2006. *Plant Health Prog*. 2010;11(1). Online. DOI 10.1094/PHP-2010-0125-01-RS.

Yin L., Zhang H., Tang Z., Xu J., Yin D., Zhang Z., Yuan X., Zhu M., Zhao S., Li X. rMVP: A memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genom. Proteom. Bioinform.* 2021;19(4):619-628. DOI 10.1016/j.gpb.2020.10.007.

Yu J., Pressoir G., Briggs W., Bi I., Yamasaki M., Doebley J., McMullen M., Gaut B., Nielson D., Holland J., Kresovich S., Buckler E. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet*. 2006;38(2): 203-208. DOI 10.1038/ng1702.

Zatybekov A., Abugalieva S., Didorenko S., Rsaliyev A., Turuspekov Y. GWAS of a soybean breeding collection from South East and South Kazakhstan for resistance to fungal diseases. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2018;22(5):536-543. DOI 10.18699/VJ18.392.

Zatybekov A., Abugalieva S., Didorenko S., Turuspekov Y. Effect of population size on genome-wide association study of agronomic traits in soybean. *Proc. Latv. Acad. Sci*. 2020;74(4):244-251. DOI 10.2478/prolas-2020-0039.

Zhang J., Song Q., Cregan P.B., Nelson R.L., Wang X., Wu J., Jiang G.L. Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics*. 2015;16(1):217. DOI 10.1186/s12864-015-1441-4.

ORCID ID

A. Zatybekov orcid.org/0000-0003-4310-5753
S. Abugalieva orcid.org/0000-0002-9748-507X
S. Didorenko orcid.org/0000-0002-2223-0718
A. Rsaliyev orcid.org/0000-0002-9921-6076
A. Maulenbay orcid.org/0000-0001-9535-2997
C. Fang orcid.org/0000-0003-0564-7586
Y. Turuspekov orcid.org/0000-0001-8590-1745

Acknowledgements. The authors would like to acknowledge the funding from the Ministry of Science and Higher Education of the Republic of Kazakhstan, under the science and technology program O.001 "Biological safety of the Republic of Kazakhstan: assessment of threats, scientific and technical basis for their prevention and elimination".

Conflict of interest. The authors declare no conflict of interest.

Received June 8, 2023. Revised July 11, 2023. Accepted July 24, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):572-581
DOI 10.18699/VJGB-23-69

Original Russian text https://vavilovj-icg.ru/

Powdery mildew resistance of apple clonal rootstocks from the collection of the Michurinsk State Agrarian University

I.N. Shamshin[1, 2] ✉, M.L. Dubrovsky[1, 2], A.A. Trifonova[2], K.V. Boris[2], A.M. Kudryavtsev[2]

[1] Michurinsk State Agrarian University, Michurinsk, Russia
[2] Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, Russia
✉ ivan_shamshin@mail.ru

Abstract. Apple clonal rootstocks are the basis of modern intensive horticulture, providing a rapid increase in yield and convenience of fruit trees cultivation. Production of clonal rootstocks under high humidity often causes powdery mildew infection caused by the pathogenic fungus *Podosphaera leucotricha* Salm., which significantly reduces the productivity of stoolbed. Growing powdery mildew resistant genotypes is the most appropriate way to combat this disease and allows reducing the use of fungicides. To accelerate the search for resistant forms, molecular markers associated with resistance genes have been developed. However, these markers have not been used to study clonal rootstocks. The aims of the work were the field assessment of powdery mildew resistance of apple clonal rootstocks from the collection of the Michurinsk State Agrarian University and the screening of the collection for *Pl-1*, *Pl-2*, *Pl-w* and *Pl-d* resistance genes. The results of a three-year field evaluation of powdery mildew resistance of 80 rootstocks allowed us to distinguish five main groups ranging from very low to highly resistant. A group of 57 accessions was classified as powdery mildew resistant. The search for resistance genes was performed using the AT20 SCAR (*Pl-1* gene), OPU02 SCAR (*Pl-2* gene), EM DM01 (*Pl-d* gene), and EM M02 (*Pl-w* gene) markers. The *Pl-d* and *Pl-1* genes identified in 33 (41.25 %) and 31 (38.75 %) accessions, respectively, were the most common in the collection. The *Pl-w* gene was detected only in two accessions. Identification of the *Pl-2* gene with the OPU02 SCAR marker did not reveal a fragment of the expected size. Thirty accessions with different powdery mildew resistance scores had two genes, *Pl-1* and *Pl-d*, and highly resistant forms G16 and 14-1 had a combination of the *Pl-d* and *Pl-w* genes. These accessions can be used as donors of powdery mildew resistance for breeding new apple clonal rootstocks.
Key words: clonal rootstocks; *Podosphaera leucotricha* Salm.; molecular markers; resistance genes.

For citation: Shamshin I.N., Dubrovsky M.L., Trifonova A.A., Boris K.V., Kudryavtsev A.M. Powdery mildew resistance of apple clonal rootstocks from the collection of the Michurinsk State Agrarian University. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):572-581. DOI 10.18699/VJGB-23-69

Исследование устойчивости клоновых подвоев яблони из коллекции Мичуринского государственного аграрного университета к мучнистой росе

И.Н. Шамшин[1, 2] ✉, М.Л. Дубровский[1, 2], А.А. Трифонова[2], К.В. Борис[2], А.М. Кудрявцев[2]

[1] Мичуринский государственный аграрный университет, Мичуринск, Россия
[2] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
✉ ivan_shamshin@mail.ru

Аннотация. Клоновые подвои яблони являются основой современного интенсивного садоводства, они обеспечивают быстрое наращивание урожая и высокую технологичность возделывания плодовых деревьев. При производстве клоновых подвоев в условиях обильного увлажнения субстрата активно развивается возбудитель мучнистой росы – патогенный гриб *Podosphaera leucotricha* Salm., значительно снижающий продуктивность маточника. Использование устойчивых к мучнистой росе подвойных форм – наиболее целесообразный способ борьбы с этим заболеванием, позволяющий существенно сократить применение химических средств защиты растений. Для ускоренного поиска устойчивых форм разработаны молекулярные маркеры, ассоциированные с генами устойчивости. Однако для изучения клоновых подвоев эти маркеры ранее не применялись. Целью работы было изучение полевой устойчивости к мучнистой росе и выявление форм с генами устойчивости *Pl-1*, *Pl-2*, *Pl-w* и *Pl-d* в крупнейшей отечественной коллекции клоновых подвоев яблони Мичуринского государственного аграрного университета. Результаты трехлетней полевой оценки устойчивости 80 отобранных для анализа подвойных форм позволили выделить пять основных групп образцов: от неустойчивых до высокоустойчивых. Наиболее многочисленной была группа устойчивых форм, включающая 57 образцов. Поиск генов устойчивости проводили с помощью маркеров AT20 SCAR (ген *Pl-1*), OPU02 SCAR (ген *Pl-2*), EM DM01 (ген *Pl-d*) и EM M02 (ген *Pl-w*).

Наибольшее распространение в изученной коллекции имеют гены *Pl-d* и *Pl-1*, обнаруженные у 33 (41.25 %) и 31 (38.75 %) образца соответственно. Ген *Pl-w* был выявлен только у двух форм, а при идентификации гена *Pl-2* с маркером OPU02 SCAR не найдено фрагмента ожидаемого размера. При этом 30 образцов с различным уровнем полевой устойчивости имеют два гена, *Pl-1* и *Pl-d*, а высокоустойчивые образцы G16 и 14-1 – комбинацию генов *Pl-d* и *Pl-w*. Эти образцы могут быть использованы в качестве доноров для создания новых клоновых подвоев яблони, несущих комплекс генов устойчивости к мучнистой росе.
Ключевые слова: клоновые подвои; *Podosphaera leucotricha* Salm.; молекулярные маркеры; гены устойчивости.

Introduction

Powdery mildew is one of the main fungal diseases of apple that cause significant economic damage to modern intensive horticulture. Powdery mildew is caused by the pathogenic fungus *Podosphaera leucotricha* Salm. The disease actively develops in humid warm weather in young orchards and nurseries. The fungus forms a white coating on the plant, which eventually turns brown. All above-ground parts of the plant (leaves, shoots, flowers and fruits) are affected by powdery mildew, leading to significant fruit yield losses (Kozlovskaya et al., 2018). During epiphytotic years, powdery mildew can affect up to 100 % of trees and lead to the loss of more than half of the yield. Up to twenty fungicide treatments a year are required to avoid severe powdery mildew infection, significantly increasing the chemical pressure on orchards (Holb et al., 2017; Höfer et al., 2021).

In our country, powdery mildew is widespread in the Southern Federal District (Yakuba, 2018), where intensive horticulture using dwarfing clonal rootstocks is developed. Due to the high regenerative and rooting ability of shoots and easy vegetative propagation, apple clonal rootstocks are one of the most important components in the production of planting material. Rootstocks provide tree anchorage, water and nutrient uptake and affect many other physiological processes. In modern industrial orchards, clonal rootstocks play a leading role in regulating tree vigor, precocity and also contribute to fruit quality and a rapid increase in yield. The use of dwarfing clonal rootstocks provides convenience of fruit tree cultivation, increases orchards productivity and is more economically efficient for fruit growers.

For the propagation of clonal rootstocks in stoolbeds, good rooting of shoots requires abundant moisture in the substrate, so systematic sprinkler irrigation is often used. Constant high humidity may lead to severe powdery mildew damage of the shoots of uterine bushes, causing a disruption in the normal functioning of the photosynthetic apparatus of leaves and a significant decrease in the productivity of the stoolbed. In this regard, many rootstock breeding programs are aimed at identifying new genotypes resistant to powdery mildew.

The search for sources of resistance to powdery mildew has been carried out since the second half of the XX century. To date, a number of resistance gene sources have been identified: the *Pl-1* gene was first discovered in wild species *Malus* × *robusta*, the *Pl-2* gene was identified in *M. zumi* (Knight, Alston, 1968), *Pl-w* in the cultivar White Angel (Gallott et al., 1985; Simon, Weeden, 1991), *Pl-d* in the hybrid D12 (Visser, Verhaegh, 1976) and *Pl-m* in the hybrid MIS (mildew immune selection) (Dayton, 1977).

To search for powdery mildew resistance gene sources, a number of markers have been developed and tested. Thus, the SCAR marker of the *Pl-1* gene was created and successfully tested based on the RAPD marker OPAT20, and the SCAR marker of the *Pl-2* gene, based on the RAPD marker OPU02 (Markussen et al., 1995; Gardiner et al., 2003). In addition, single nucleotide polymorphisms (SNPs) associated with *Pl-2* powdery mildew resistance were identified and validated (Jänsch et al., 2015; Chagné et al., 2019). Based on the AFLP data, SCAR markers EM M01, EM M02 and EM DM01 of the *Pl-w* and *Pl-d* genes were created (Evans, James, 2003; James et al., 2004). Microsatellite markers *CH03C02* and *CH01D03* were also used to identify the *Pl-d* gene (James et al., 2004).

These markers have been successfully used to screen apple cultivar collections in Germany (Höfer et al., 2021) and the Czech Republic (Patzak et al., 2011), as well as to study wild *Malus orientalis* populations of Iran (Amirchakhmaghi et al., 2018). In our country, studies of the powdery mildew resistance genes distribution in commercial and local apple cultivars, as well as in wild species of the genus *Malus* have been carried out (Suprun et al., 2015; Lyzhin, Saveleva, 2020, 2021). However, apple rootstocks have barely been studied using powdery mildew resistance genes markers.

The Michurinsk State Agrarian University (Michurinsk SAU) is the founder of clonal rootstock breeding in Russia and one of the leading institutions in the world in this area. A unique collection of rootstocks, which is the largest in the country, is maintained here. Rootstocks bred at the Michurinsk SAU are grown in orchards in Russia, Europe, USA and other countries. However, there were no studies of their resistance to powdery mildew, including those using molecular markers. Such studies will help to identify sources and donors of resistance in the collection to create new rootstocks combining several powdery mildew resistance genes.

The aim of the work was to study field resistance of the Michurinsk SAU apple clonal rootstocks collection to powdery mildew and to identify accessions with the *Pl-1*, *Pl-2*, *Pl-w* and *Pl-d* resistance genes.

Materials and methods

Eighty apple clonal rootstocks were studied, including 74 rootstocks of the Michurinsk SAU, three rootstocks from other Russian breeding centers (B7-35, K-1 and Ural-5) and three foreign rootstocks (M9 T337, G16 and Babarabskaya yablonya). *Malus* species from the collection of the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), Belarusian cultivar Diyament and cultivar Evereste from the collection of the Michurinsk SAU were used as references for powdery mildew resistance genes *Pl-1*, *Pl-2*, *Pl-d* and *Pl-w*.

Total genomic DNA was extracted from fresh young leaves using the Quick-DNA Plant/Seed Miniprep Kit (Zymo Research, USA) following the manufacturers' protocol.

Table 1. Markers for powdery mildew resistance genes identification

Gene	Marker	Primer sequences	Fragment size, bp	Annealing temperature, °C
Pl-1	AT20 SCAR	F ATCAGCCCCACATGAATCTCATACC R ACATCAGCCCTCAAAGATGAGAAGT	450/500	62
Pl-2	OPU02 SCAR	F CCGACCGATCAGGAATTGTCACCAG R TCGATTATCACTATGTACGGGAGCA	1700	62
Pl-d	EM DM01	F AGGATAATAATCTATCTTGTAAAGG R CCATTCAGCCCAACGAGT	90	53
Pl-w	EM M02	F CTGCAGACTGTTTGTAAGTTGG R AACTCCTTGATTTCTCCTATTGTT	250	56

Powdery mildew resistance genes were analyzed using the following DNA markers: codominant AT20 SCAR marker (*Pl-1* gene) (Markussen et al., 1995), dominant OPU02 SCAR markers (*Pl-2* gene) (Gardiner et al., 2003), EM DM01 (*Pl-d* gene) (James et al., 2004) and EM M02 (*Pl-w* gene) (Evans, James, 2003). The primers sequences and annealing temperatures are presented in Table 1.

PCR reactions were performed in a SimpliAmp (Applied Biosystems, USA) thermal cycler in a final volume of 15 µl containing 20 ng of genomic DNA, 1.5 mM dNTPs, 2.5 mM $MgSO_4$, 10 pM of each primer (Syntol, Russia), 1 u Taq polymerase and 1x standard PCR buffer (Thermo Fisher Scientific, USA). To determine fragment sizes, molecular markers GeneRuler 100 bp (Thermo Fisher Scientific), 50 bp DNA marker (Dialat Ltd, Russia) and Start 250 (Diaem, Russia) were used. After amplification, PCR products were separated in 2 % agarose gels, stained with ethidium bromide and analyzed on a UV-light box.

Powdery mildew resistance evaluation was carried out in the field for three years (2020–2022) in a nursery with sprinkler irrigation to maintain high substrate moisture. The evaluation was carried out according to the method described in the "Program and Methodology of Variety Studies for Fruit, Berry and Nut Crops" (Sedov, Ogoltsova, 1999), with minor modifications. The resistance rate was evaluated by the presence of powdery mildew lesions on the leaf blade. Ten shoots of each rootstock were analyzed. The assessment of the resistance rate was carried out on a five-point scale, where: 0 points – highly resistant genotype (HR, no lesions were detected), 0.1–1 points – resistant genotype (R, lesions up to 1 % of the leaf blade), 1.1–2 points – moderately resistant (M, lesions up to 10 % of the leaf blade), 2.1–3 points – low-resistant (L, lesions up to 40 % of the leaf blade), 3.1–5 points – very low-resistant (VL, lesions from 40 up to 100 % of the leaf blade). For each analyzed genotype, the average score for three years was calculated (Table 2).

Results

As a result of this work, 80 apple rootstocks were analyzed using markers of four powdery mildew resistance genes, *Pl-1*, *Pl-2*, *Pl-w* and *Pl-d* (see Table 2).

The *Pl-1* gene was mapped in the resistance gene cluster of apple LG XII (Dunemann et al., 2007). To identify this gene, the SCAR marker AT20 was used, which allows detecting two fragments – 450 and 500 bp. The presence of a 450 bp fragment is associated with *Pl-1* resistance (Markussen et al., 1995). The accession of wild species *M.* × *robusta*, from which this gene was initially introgressed, was used as reference.

Using the AT20 SCAR marker, a 450 bp fragment was detected in eight studied rootstocks: 54-118, 2-9-94, 2-9-102, 2-12-36, 4-2-41, 9-1-2, 5-26-127 and 67-5(32). At the same time, this fragment together with a 500 bp fragment was detected in 23 accessions, and in the remaining 49 accessions, only a 500 bp fragment was amplified (see Table 2). An example of *Pl-1* gene identification is shown in Fig. 1.

The *Pl-2* gene was mapped on apple LG XI. The OPU02 SCAR marker was used for *Pl-2* identification (Gardiner et al., 2003). The presence of a dominant *Pl-2* allele is detected by amplification of a 1700 bp PCR fragment (Gardiner et al., 2003). The accession of wild species *M. zumi*, from which this gene was initially derived, was used as reference.

In the studied collection, an expected 1700 bp fragment was not revealed. However, in accessions 14-1 and G16, a 1500 bp fragment was amplified, while a ~2200 bp fragment was detected in the reference *M. zumi* accession (Fig. 2).

The dominant SCAR marker EM DM01 was used to identify the *Pl-d* gene mapped on apple LG XII (James et al., 2004). The expected size of the PCR product of this marker is 90 bp. Apple cultivar Diyament, for which the presence of this gene was previously detected, was used as reference (Kozlovskaya et al., 2018). In the studied collection, the *Pl-d* gene was identified in 33 out of 80 analyzed accessions (see Table 2). An example of *Pl-d* gene identification is shown in Fig. 3.

The *Pl-w* gene was mapped on apple LG VIII. SCAR marker EM M02 was used for *Pl-w* identification (Evans, James, 2003). This marker amplifies a 250 bp fragment indicating the presence of the *Pl-w* gene. The Evereste apple, in which the *Pl-w* gene was previously identified, was used as reference (Patzak et al., 2011). In the studied collection of apple clonal rootstocks, the *Pl-w* gene was detected only in two accessions – G16 and 14-1 (see Table 2, Fig. 4).

As a result of field evaluation of apple clonal rootstocks powdery mildew resistance, all studied accessions were divided into five main groups according to the resistance rate, from high to very low (Fig. 5).

It was shown that 57 accessions are resistant (0.1–1 points) to powdery mildew, which was 71.25 % of the total sample (see Fig. 5). In this group, the lowest average score of powdery mildew resistance (0.1) was noted in rootstocks 62-396

Table 2. Analyzed apple rootstocks, data on the diversity of the *Pl-1*, *Pl-2*, *Pl-d* and *Pl-w* resistance genes and the results of field powdery mildew resistance assessment

No.	Accession	Parentage	AT20 (*Pl-1*)		OPU02 (*Pl-2*)	EM DM01 (*Pl-d*)	EM M02 (*Pl-w*)	Powdery mildew infection ratings, points	Resistance rate*
			450 bp	500 bp	1700 bp	90 bp	250 bp		
1	54-118	B9×13-14	+	–	–	+	–	0.3±0.05	R
2	57-490		+	+	–	+	–	0.4±0.03	R
3	62-396 (B10)	13-14×B9	–	+	–	–	–	0.1±0.02	R
4	70-20-20	57-469×57-344	–	+	–	–	–	1.4±0.1	M
5	83-1-15	64-143×54-118	–	+	–	–	–	0.4±0.04	R
6	Malysh Budagovskogo (MB, 76-6-6)	57-344×57-490	–	+	–	–	–	0.1±0.02	R
7	Budagovsky 9 (B9, Bud9)	M8×Krasniy shtandart	–	+	–	–	–	0.2±0.03	R
8	M9 (T337)	B9 clone	–	+	–	–	–	0.9±0.1	R
9	2-3-2	82-27-6 open pollination	+	+	–	+	–	0.2±0.03	R
10	2-3-3		+	+	–	+	–	0.7±0.06	R
11	2-3-8		+	+	–	+	–	0.3±0.02	R
12	2-3-14		+	+	–	+	–	0.4±0.03	R
13	2-3-17		–	+	–	–	–	4.2±0.2	VL
14	2-3-19		+	+	–	–	–	4.1±0.3	VL
15	2-3-44		+	+	–	+	–	4.5±0.2	VL
16	2-3-49		–	+	–	–	–	1.5±0.1	M
17	2-9-49	82-26-2 open pollination	+	+	–	+	–	0.7±0.1	R
18	2-9-56		–	+	–	–	–	2.2±0.2	L
19	2-9-77		+	+	–	+	–	0.5±0.03	R
20	2-9-90		–	+	–	–	–	0.3±0.02	R
21	2-9-94		+	–	–	+	–	2.8±0.2	L
22	2-9-96		–	+	–	–	–	0.2±0.01	R
23	2-9-102		+	–	–	+	–	1.7±0.1	M
24	2-12-10	82-11-5 open pollination	–	+	–	–	–	1.5±0.1	M
25	2-12-15		+	+	–	+	–	1.3±0.1	M
26	2-12-27		+	+	–	+	–	0.3±0.02	R
27	2-12-34		–	+	–	–	–	0.2±0.02	R
28	2-12-36		+	–	–	+	–	0.2±0.01	R
29	2-15-2	85-8-12 open pollination	+	+	–	+	–	3.6±0.2	VL
30	2-15-15		+	+	–	+	–	2.8±0.1	L
31	3-4-7	62-396 open pollination	–	+	–	–	–	0.8±0.1	R
32	3-10-3	82-11-2×Wealthy	–	+	–	–	–	0.5±0.04	R
33	4-2-3	82-52-6×82-26-2	–	+	–	–	–	0.3±0.02	R
34	4-2-50		+	+	–	+	–	0.4±0.04	R
35	4-6-5	83-11-10 open pollination	+	+	–	+	–	0.3±0.02	R
36	5-21-27	82-27-6×Zhigulevskoe	–	+	–	–	–	0.2±0.01	R
37	5-21-93		+	+	–	+	–	0.3±0.04	R
38	5-24-1	82-26-2×Orlik	–	+	–	–	–	0.2±0.02	R
39	5-27-1	Zhigulevskoe×82-26-2	–	+	–	–	–	0.3±0.01	R
40	5-28-11	82-57-8×*Malus baccata* (L.) Borkh.	–	+	–	–	–	0.4±0.03	R

Table 2 (continued)

No.	Accession	Parentage	AT20 (*Pl-1*)		OPU02 (*Pl-2*)	EM DM01 (*Pl-d*)	EM M02 (*Pl-w*)	Powdery mildew infection ratings, points	Resistance rate*
			450 bp	500 bp	1700 bp	90 bp	250 bp		
41	9-1-1	57-157 × Stroevskoe	–	+	–	–	–	0.3 ± 0.02	R
42	9-1-2		+	–	–	+	–	2.1 ± 0.2	L
43	9-1-3		–	+	–	–	–	4.6 ± 0.3	VL
44	9-1-4		–	+	–	–	–	0.5 ± 0.03	R
45	9-1-5		–	+	–	+	–	0.4 ± 0.03	R
46	9-1-9		–	+	–	–	–	0.2 ± 0.01	R
47	5-26-127	–	+	–	–	+	–	2.4 ± 0.2	L
48	57-146	B9 open pollination	–	+	–	–	–	0.2 ± 0.02	R
49	70-6-8	54-83 × 57-344	+	+	–	+	–	0.3 ± 0.02	R
50	Babarabskaya yablonya	*Malus turkmenorum* Juz. et M. Pop.	–	+	–	–	–	0.2 ± 0.03	R
51	14-1	*Malus sieboldii* Rehd., open pollination	–	+	1500 bp	+	+	0	HR
52	73-9-3	57-545 × 57-366	+	+	–	+	–	0.3 ± 0.03	R
53	62-223	Anoka × B9	–	+	–	–	–	0.2 ± 0.02	R
54	3-3-4	85-6-5 × Spartan	–	+	–	–	–	0.2 ± 0.02	R
55	71-7-22	57-531 × 57-233	+	+	–	+	–	0.2 ± 0.02	R
56	71-3-88	58-257 × B9	–	+	–	–	–	0.2 ± 0.04	R
57	Б7-35	M4 × M9	–	+	–	–	–	2.3 ± 0.1	L
58	71-3-137	58-257 × B9	–	+	–	–	–	0.4 ± 0.03	R
59	7-8-5 (Ural-5)	57-469 open pollination	–	+	–	–	–	0.3 ± 0.03	R
60	K-1	Borovinka × M9	–	+	–	–	–	3.4 ± 0.2	VL
61	58-238	B9 × Naliv Aliy	–	+	–	–	–	0.2 ± 0.02	R
62	71-3-49	58-257 × B9	–	+	–	–	–	0.3 ± 0.02	R
63	64-143	B9 × 57-290	–	+	–	–	–	1.7 ± 0.1	M
64	71-3-150	58-257 × B9	–	+	–	–	–	0.3 ± 0.02	R
65	71-3-195		–	+	–	–	–	0.2 ± 0.03	R
66	85-2-11	3-4-98 × 54-118	+	+	–	+	–	0.3 ± 0.04	R
67	76-6-13	57-344 × 57-490	–	+	–	–	–	0.3 ± 0.04	R
68	G16	Ottawa 3 × *Malus floribunda* Siebold ex Van Houtte	–	+	1500 bp	+	+	0	HR
69	67-5(32)	54-83 open pollination	+	–	–	+	–	1.2 ± 0.2	M
70	75-11-280	B9 open pollination	–	+	–	–	–	0.2 ± 0.03	R
71	57-491	B9 × 13-14	+	+	–	+	–	0.4 ± 0.03	R
72	87-7-12	54-118 × B9	+	+	–	+	–	2.8 ± 0.2	L
73	75-12-23	A2 open pollination	–	+	–	–	–	0.2 ± 0.01	R
74	75-11-232	B9 open pollination	–	+	–	–	–	0.2 ± 0.01	R
75	76-3-6	M27 × B9	–	+	–	–	–	0.4 ± 0.03	R
76	86-6-12	–	–	+	–	–	–	1.1 ± 0.1	M
77	85-11-9	70-5-10 × 54-118	–	+	–	–	–	0.2 ± 0.01	R
78	69-6-217	B9 × Kitayka Rozovaya	+	+	–	+	–	0.3 ± 0.02	R
79	69-28-11	58-257 × B9	–	+	–	–	–	0.2 ± 0.02	R
80	4-2-41	82-52-6 × 82-26-2	+	–	–	+	–	0.2 ± 0.02	R

Table 2 (end)

No.	Accession	Parentage	AT20 (*Pl-1*)		OPU02 (*Pl-2*)	EM DM01 (*Pl-d*)	EM M02 (*Pl-w*)	Powdery mildew infection ratings, points	Resistance rate*
			450 bp	500 bp	1700 bp	90 bp	250 bp		
81	*Malus* × *robusta* **(Carrière) Rehder (к43200)		+	–	–	–	–	–	–
82	*Malus zumi* (Matsum.) Rehder (к41272)		–	–	~2200 bp	–	–	–	–
83	Diyament		–	–	–	+	–	–	–
84	Evereste		–	–	–	–	+	–	–

* HR – highly resistant; R – resistant; M – moderately resistant; L – low-resistant; VL – very low-resistant.
** References for: *Pl-1* gene – *M. robusta* (k43200); *Pl-2* gene – *M. zumi* (k41272); *Pl-d* gene – Diyament; *Pl-w* gene – Evereste.

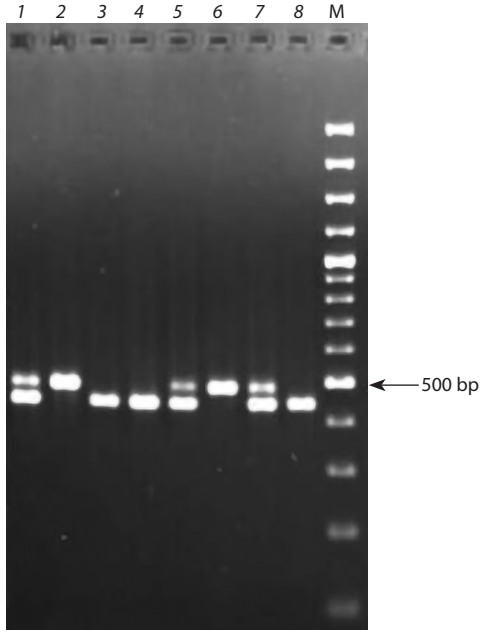

Fig. 1. The results of the *Pl-1* gene identification with the AT20 SCAR marker in apple clonal rootstocks.

1 – 57-490; *2* – B9; *3* – 54-118; *4* – 2-9-94; *5* – 2-12-27; *6* – 5-27-1; *7* – 2-3-19; *8* – *M.* × *robusta* k43200. M – marker GeneRuler 100 bp.

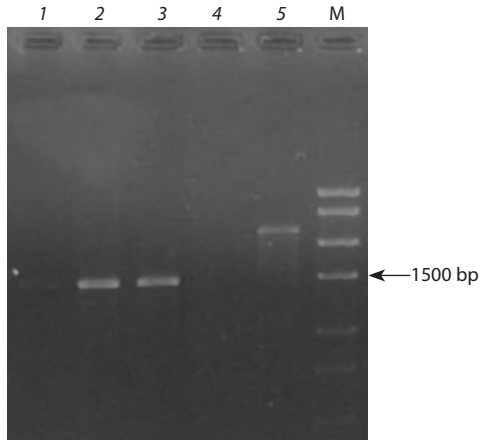

Fig. 2. The results of the *Pl-2* gene identification with the OPU02 SCAR marker in apple clonal rootstocks.

1 – 54-118; *2* – 14-1; *3* – G16; *4* – 9-1-2; *5* – *M. zumi* k41272. M – marker Start 250.

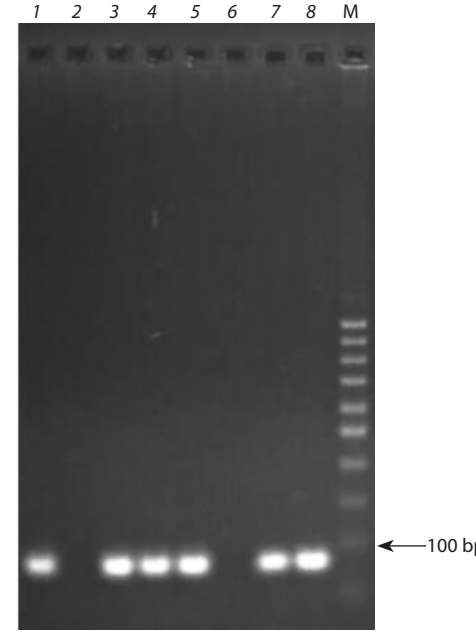

Fig. 3. The results of the *Pl-d* gene identification with the EM DM01 marker in apple clonal rootstocks.

1 – 54-118; *2* – 62-396; *3* – 2-3-8; *4* – 2-3-14; *5* – 2-12-15; *6* – 2-12-10; *7* – 57-490; *8* – Diyament. M – 50 bp DNA marker.

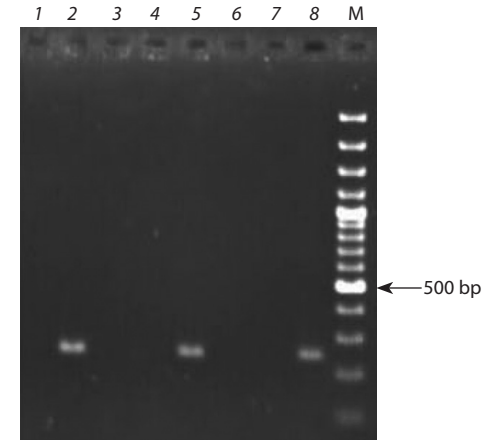

Fig. 4. The results of the *Pl-w* gene identification with the EM M02 marker in apple clonal rootstocks.

1 – 2-3-17; *2* – 14-1; *3* – 7-8-5; *4* – 9-1-2; *5* – G16; *6* – 5-27-1; *7* – 2-3-8; *8* – Evereste. M – marker GeneRuler 100 bp.

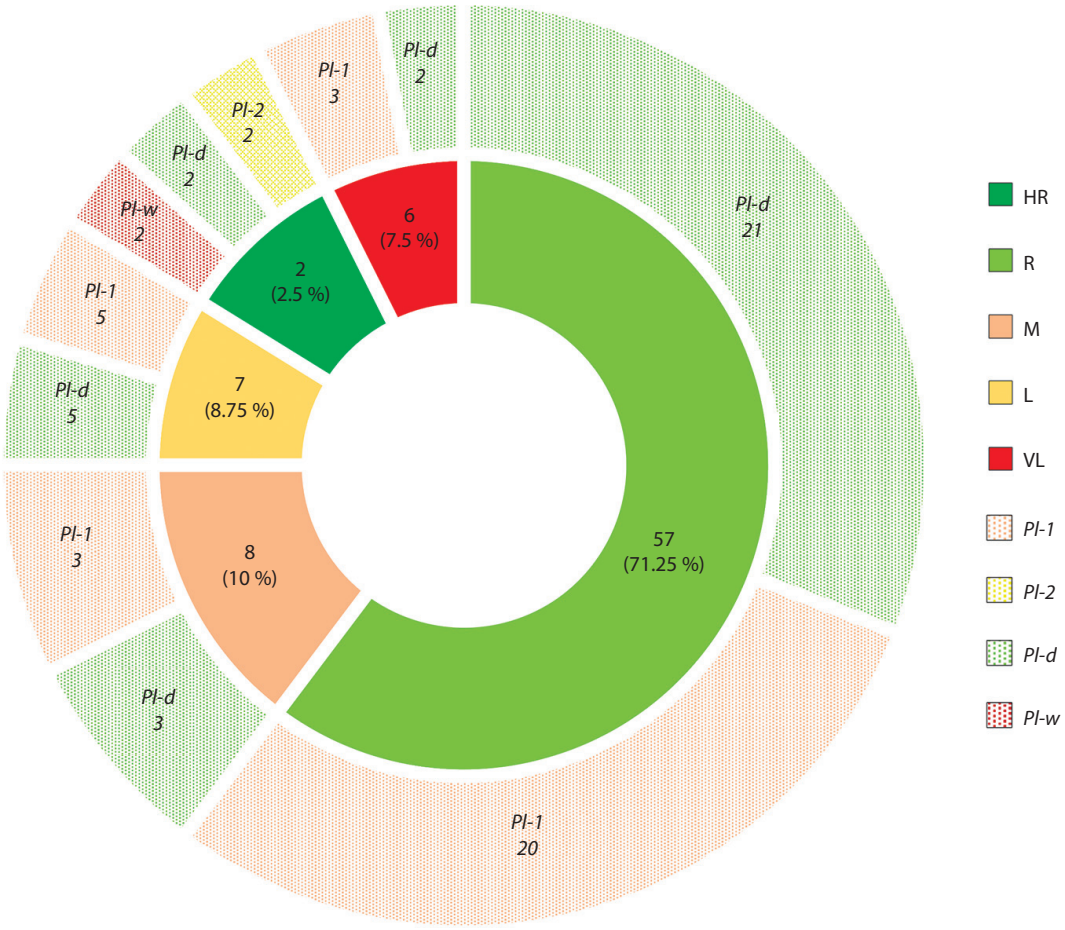

Fig. 5. Proportion of accessions with different powdery mildew resistance in the studied collection according to field resistance evaluation data.

Groups of accessions according to the field resistance rate: HR – highly resistant; R – resistant; M – moderately resistant; L – low-resistant; VL – very low-resistant (inner circle). The number of accessions with the *Pl-1*, *Pl-2*, *Pl-w* and *Pl-d* genes in each group (outer circle).

and 76-6-6, and the maximum (0.9), in rootstock M9 T337 (see Table 2). The group of moderately resistant forms (1.1–2 points) included eight accessions (10 %). The average resistance score for this group varied from 1.1 for 86-6-12 to 1.7 for rootstocks 2-9-102 and 64-143. The proportion of low-resistant (2.1–3 points) and very low-resistant (3.1–5 points) rootstocks was 8.75 %, (from 2.1 for rootstock 9-1-2 to 2.8 for rootstocks 2-9-94, 87-7-12 and 2-15-15) and 7.5 % (from 3.4 for K-1 to 4.6 for 9-1-3), respectively. The group of highly resistant accessions (0 points) was the smallest and included hybrid form 14-1 and rootstock G16 that had not been affected by powdery mildew (see Table 2, Fig. 5).

The distribution of the studied powdery mildew resistance genes in accessions with different field resistance was analyzed (see Fig. 5). The *Pl-d* and *Pl-1* genes were the most widespread in the collection. Thus, the *Pl-d* gene was present in all groups, and the *Pl-1* gene was not detected only in highly resistant accessions. While the *Pl-w* gene, on the contrary, was detected only in two accessions 14-1 and G16, which are highly resistant. In addition, specific 1500 bp PCR products were detected only in these accessions with the *Pl-2* gene marker (see Fig. 2).

Discussion

The analysis of 80 apple clonal rootstocks using markers of powdery mildew resistance genes *Pl-1*, *Pl-2*, *Pl-d* and *Pl-w* allowed assessing the variability of these genes in the Michurinsk SAU collection for the first time.

The most common in the studied collection were the *Pl-d* and *Pl-1* genes identified in 33 (41.25 %) and 31 (38.75 %) accessions, respectively (see Table 2). The *Pl-w* gene was detected only in G16 and 14-1 accessions (see Table 2, Fig. 4). The distribution of the studied powdery mildew resistance genes in the collection may be related to the rootstocks pedigrees, including both wild *Malus* species and commercial apple cultivars and landraces (see Table 2).

Wild apple species are the most promising sources of disease and pest resistance genes for breeding apple cultivars and rootstocks (Pereira-Lorenzo et al., 2018; Solomatin, 2018). However, when breeding clonal rootstocks, during the primary selection of hybrid seedlings, genotypes with good vegetative propagation ability, high rooting ability and increased winter hardiness of the root system are distinguished. The study of resistance to phytopathogens is usually carried out later among the selected rootstocks.

It should be noted that there were practically no studies of apple rootstocks collections using powdery mildew resistance gene markers, and the researchers have been mainly focused on studying apple cultivars and wild *Malus* species. For example, the study using the EM DM01 marker showed a wide distribution of the *Pl-d* gene among old and modern apple cultivars from the collection of the Institute for Fruit Growing, Belarus (Urbanovich et al., 2010). In the study of 145 old and local apple cultivars from the Czech collection, the *Pl-d* gene was detected only in four accessions (Patzak et al., 2011). The study of apple cultivars from the Dresden-Pillnitz gene bank using the microsatellite marker of the *Pl-d* gene also showed its low distribution (Höfer et al., 2021).

The *Pl-d* gene was also found in wild apple species. In the study of 67 *Malu*s species forms, the *Pl-d* gene was identified in seven accessions, and for *M. sieversii* and *M. orientalis*, intraspecific polymorphism was revealed (Lyzhin, Saveleva, 2021). In the study of *M. orientalis* populations growing in Iran, the *Pl-d* gene diversity was also noted (Amirchakhmaghi et al., 2018). Transcaucasia is thought to be one of the centers of origin of low-vigorous clonal rootstocks, where they originated from wild species (Budagovskiy, 1976; Solomatin, 2018). The most common apple species growing in this region is *M. orientalis*. Thus, the distribution of the *Pl-d* gene in the collection may be associated with the origin of rootstocks from wild species, including *M. orientalis*. Another possible factor of the distribution of this gene in the analyzed collection is the use of apple cultivars in breeding rootstocks.

The *Pl-1* gene was the second most common gene in the collection. According to previous studies, this gene is not widely distributed, both among old and commercial apple cultivars (Urbanovich et al., 2010; Patzak et al., 2011; Suprun et al., 2015; Kozlovskaya et al., 2018; Lyzhin, Saveleva, 2020; Höfer et al., 2021). However, the *Pl-1* gene was found in *Malus* species. In a collection of 67 wild *Malus* forms, the *Pl-1* gene was detected in 37.3 % of accessions, including *M.* × *robusta*, the species this gene was initially derived from. At the same time, intraspecific polymorphism for this gene was shown for some species (Lyzhin, Savelyeva, 2021). The presence of the *Pl-1* gene in wild species was also noted in the study by O. Urbanovich et al. (2010). Apparently, the distribution of the *Pl-1* gene in the collection of apple clonal rootstocks is associated with the presence of wild *Malus* species in their pedigrees.

The *Pl-w* gene, identified in two rootstocks using the EM M02 marker, is less distributed in the studied collection. The *Pl-w* gene provides a higher level of resistance than the *Pl-1* and *Pl-2* genes (Simon, Weeden, 1991; Evans et al., 2003). This gene is quite common among wild *Malus* species, but is practically not found in apple cultivars (Patzak et al., 2011; Kozlovskaya et al., 2018; Lyzhin, Saveleva, 2021).

Probably, it is the presence of the *Pl-w* gene that ensures high powdery mildew resistance in G16 and 14-1 accessions, which was revealed during the field resistance assessment (see Table 2). Apparently, this is due to their origin from wild species. The American rootstock G16 is derived from *M. floribunda*, and the hybrid form 14-1 of the Michurinsk SAU is derived from *M. sieboldii*. Previously, using the EM M02 marker, the *Pl-w* gene was detected in *M. floribunda* and *M. sieboldii* (Lyzhin, Saveleva, 2021).

The *Pl-2* gene marker OPU02 SCAR revealed ~1500 bp fragments in two accessions (G16 and 14-1) from the collection (see Fig. 2), while the presence of a dominant *Pl-2* allele was detected by amplification of a 1700 bp fragment (Gardiner et al., 2003). Previously, using the OPU02 SCAR marker, a 1500 bp fragment was detected in the cultivar Favorit of the Crimean Experimental Horticultural Station (Suprun et al., 2015). The presence of unknown alleles of this gene or a new resistance gene was suggested, as the cultivar is powdery mildew resistant. This may also refer to accessions G16 and 14-1, since both are completely resistant to powdery mildew according to the results of the field evaluation, although these accessions also have the *Pl-d* and *Pl-w* genes (see Table 2).

The *Pl-2* gene was first identified in *M. zumi* (Knight, Alston, 1968). However, in the *M. zumi* accession from the VIR collection used as reference for this gene, a fragment of ~2200 bp was obtained, instead of the expected 1700 bp fragment (see Fig. 2). This may be due to the intraspecific polymorphism, previously noted for other powdery mildew resistance genes in wild *Malus* species (Lyzhin, Saveleva, 2021). Alternatively, the presence of a ~2200 bp fragment may be related to the existence of another powdery mildew resistance gene $Pl\text{-}_{MIS}$ in the *M. zumi* accession (Gardiner et al., 2003).

The results obtained using the OPU02 SCAR marker do not allow to identify the *Pl-2* gene by the presence of the 1700 bp fragment, which may be due to a significant distance (8.6 cM) from the marker to the gene (Gardiner et al., 2003). Therefore, in order to assess the presence of the *Pl-2* gene in the collection, other markers can be used, e. g. SNPs (Jänsch et al., 2015; Chagné et al., 2019).

It is known that in breeding for resistance, a common approach is to combine several resistance genes in one genotype, the so-called "pyramiding". The results allowed to identify a number of rootstocks that are promising for breeding. In the studied collection, 30 accessions have two resistance genes, *Pl-1* and *Pl-d*, and accessions G16 and 14-1, highly resistant to powdery mildew, have genes *Pl-d*, *Pl-w*, as well as a specific fragment detected using the *Pl-2* gene marker (see Table 2).

However, not all accessions with powdery mildew resistance genes showed high field resistance. For example, all eight accessions that had the dominant allele of the *Pl-1* gene also had the *Pl-d* gene, however, the level of their field resistance varied from 0.2 points for rootstock 4-2-41 to 2.8 points for rootstock 2-9-94 (see Table 2). At the same time, accessions G16 and 14-1 that had the *Pl-d* and *Pl-w* genes, as well as a specific fragment obtained with the *Pl-2* gene marker, were highly resistant and had no symptoms of infection (0 points).

The discrepancy between field resistance level and the presence of resistance genes can have several causes. For some of the markers, there are conflicting data on the association with the trait. A number of studies do not confirm the correlation of *Pl-1* AT20 SCAR and *Pl-d* EM DM01 markers with resistance (Dunemann et al., 2004; Kellerhals et al., 2008; Urbanovich et al., 2010).

In addition, data on the presence and distribution of different races of *P. leucotricha* and the results of artificial inoculation tests with known pathogen strains are needed for more accurate assessment of the powdery mildew resistance level. Besides, the resistance determined by known *Pl* genes can be overcome, as, for example, *Pl-1* gene resistance (Lesemann, Dunemann, 2006). Furthermore, despite a significant amount of data on the genome sequences of the cultivated apple and other *Malus* species, it is likely that not all powdery mildew resistance genes and alleles of known resistance genes have been identified at the moment.

Conclusion

Field evaluation of powdery mildew resistance of 80 accessions from the Michurinsk SAU apple clonal rootstocks collection was assessed for the first time and accessions with powdery mildew resistance genes were identified. The *Pl-d* and *Pl-1* genes were the most common in the collection. The *Pl-w* gene was found only in two accessions, while the *Pl-2* gene was not detected in the collection. At the same time 32 accessions had a combination of two *Pl* genes. The presence of powdery mildew resistance genes in rootstocks appears to be related to their pedigrees and origin from wild *Malus* species. Comparison of the field resistance assessment results with the molecular analysis data showed that the presence of resistance genes does not always correspond to the level of resistance. Thus, the *Pl-d* and *Pl-1* genes were identified in both resistant and low-resistant accessions, while the *Pl-w* gene was detected only in highly resistant accessions 14-1 and G16. The obtained data can be used for breeding new apple rootstocks with a combination of powdery mildew resistance genes.

References

Amirchakhmaghi N., Yousefzadeh H., Hosseinpour B., Espahbodi K., Aldaghi M., Cornille A. First insight into genetic diversity and population structure of the Caucasian wild apple (*Malus orientalis* Uglitzk.) in the Hyrcanian forest (Iran) and its resistance to apple scab and powdery mildew. *Genet. Resour. Crop Evol*. 2018;65(4):1255-1268. DOI 10.1007/s10722-018-0611-z.

Budagovskiy V.I. Culture of Dwarfing Fruit Trees. Moscow, 1976. (in Russian)

Chagné D., Vanderzande S., Kirk C., Profitt N., Weskett R., Gardiner S.E., Peace C.P., Volz R.K., Bassil N.V. Validation of SNP markers for fruit quality and disease resistance loci in apple (*Malus* × *domestica* Borkh.) using the OpenArray® platform. *Horti. Res*. 2019;6:30. DOI 10.1038/s41438-018-0114-2.

Dayton D.F. Genetic immunity to apple mildew incited by *Podosphaera leucotricha*. *Hortic. Sci*. 1977;12(3):225-226. DOI 10.21273/HORTSCI.12.3.225.

Dunemann F., Urbanietz A., Gardiner S., Bassett S., Bassett H., Legg W., Rusholme R., Bus V., Ranatunga C. Marker assisted selection for *Pl-1* powdery mildew resistance in apple – old markers for a new resistance gene. *Acta Hortic*. 2004;2:757-761. DOI 10.17660/ActaHortic.2004.663.137.

Dunemann F., Peil A., Urbanietz A., Garcia-Libreros T. Mapping of the apple powdery mildew resistance gene *Pl1* and its genetic association with an NBS-LRR candidate resistance gene. *Plant Breed*. 2007;126(5):476-481. DOI 10.1111/j.1439-0523.2007.01415.x.

Evans K., James C. Identification of SCAR markers linked to *Pl-w* mildew resistance in apple. *Theor. Appl. Genet*. 2003;106(7):1178-1183. DOI 10.1007/s00122-002-1147-2.

Gallott J.C., Lamb R.C., Aldwinckle H.S. Resistance to powdery mildew from some small-fruited *Malus* cultivars. *Hortic. Sci.* 1985;20(6):1085-1087. DOI 10.21273/HORTSCI.20.6.1085.

Gardiner S., Murdoch J., Meech S., Rusholme R., Bassett H., Cook M., Bus V., Rikkerink E., Gleave A., Crowhurst R., Ross G., Warrington I. Candidate resistance genes from an EST database prove a rich source of markers for major genes conferring resistance to important apple pests and diseases. *Acta Hortic.* 2003;622:141-151. DOI 10.17660/ActaHortic.2003.622.12.

Holb I.J., Abonyi F., Buurma J., Heijne B. On-farm and on-station evaluations of three orchard management approaches against apple scab and apple powdery mildew. *Crop Prot*. 2017;97:109-118. DOI 10.1016/j.cropro.2016.11.023.

Höfer M., Flachowsky H., Schröpfer S., Peil A. Evaluation of scab and mildew resistance in the Gene Bank collection of apples in Dresden-Pillnitz. *Plants.* 2021;10(6):1227. DOI 10.3390/plants10061227.

James C.M., Clarke J.B., Evans K.M. Identification of molecular markers linked to the mildew resistance gene *Pl-d* in apple. *Theor. Appl. Genet*. 2004;110(1):175-181. DOI 10.1007/s00122-004-1836-0.

Jänsch M., Broggini G.A., Weger J., Bus V.G., Gardiner S.E., Bassett H., Patocchi A. Identification of SNPs linked to eight apple disease resistance loci. *Mol. Breed*. 2015;35:45. DOI 10.1007/s11032-015-0242-4.

Kellerhals M., Patocchi A., Duffy B., Frey J. Modern approaches for breeding high quality apples with durable resistance to scab, powdery mildew and fire blight. In: Ecofruit-13th International Conference on Cultivation Technique and Phytopathological Problems in Organic Fruit-Growing: Proceedings to the Conference from 18th February to 20th February 2008 at Weinsberg/Germany. Weinsberg, 2008;226-231.

Knight R.L., Alston F.H. Sources of field immunity to mildew (*Podosphaera leucotricha*) in apple. *Can. J. Genet. Cytol*. 1968;10(2):294-298. DOI 10.1139/g68-041.

Kozlovskaya Z.A., Kondratsenok Ju.G., Gashenko T.A., Yarmolich S.A. Identification of new complex sources of resistance to apple diseases using molecular methods in Belarus. *Sadovodstvo i Vinogradarstvo = Horticulture and Viticulture.* 2018;5:23-29. DOI 10.31676/0235-2591-2018-5-23-29. (in Russian)

Lesemann S., Dunemann F. Recent findings on the biodiversity of the apple powdery mildew pathogen. *Gesunde Pflanzen.* 2006;58:117-123. DOI 10.1007/s10343-005-0108-5.

Lyzhin A.S., Saveleva N.N. Marker-mediated screening of powdery mildew resistant (*Pl-1* gene) apple genotypes. *Vestnik Mariyskogo Gosudarstvennogo Universiteta. Seriya: Selskohozyaystvennyie Nauki. Ekonomicheskie Nauki = Vestnik of Mari State University. Chapter: Agriculture. Economics.* 2020;6(2):180-186. DOI 10.30914/2411-9687-2020-6-2-180-186. (in Russian)

Lyzhin A.S., Saveleva N.N. Polymorphism of wild species of *Malus* Mill. according to powdery mildew resistance genes. *Izvestiya Natsionalnoy Akademii Nauk Belarusi. Seriya Agrarnykh Nauk = Proceedings of the National Academy of Sciences of Belarus. Agrarian Series.* 2021;59(1):62-70. DOI 10.29235/1817-7204-2021-59-1-62-70. (in Russian)

Markussen T., Krüger J., Schmidt H., Dunemann F. Identification of PCR-based markers linked to the powdery-mildew-resistance gene *Pl1* from *Malus robusta* in cultivated apple. *Plant Breed*. 1995;114(6):530-534. DOI 10.1111/j.1439-0523.1995.tb00850.x.

Patzak J., Paprštein F., Henychová A. Identification of apple scab and powdery mildew resistance genes in Czech apple (*Malus* × *domestica*) genetic resources by PCR molecular markers. *Czech J. Genet. Plant Breed*. 2011;47(4):156-165. DOI 10.17221/140/2011-CJGPB.

Pereira-Lorenzo S., Fischer M., Ramos-Cabrer A.M., Castro I. Apple (*Malus* spp.) breeding: present and future. In: Al-Khayri J., Jain S., Johnson D. (Eds.). Advances in Plant Breeding Strategies: Fruits. Vol. 3. Cham: Springer, 2018;3-29. DOI 10.1007/978-3-319-91944-7_1.

Sedov E.N., Ogoltsova T.P. Program and Methodology of Variety Studies for Fruit, Berry, and Nut Crops. Orel: All-Russia Research Institute for Agricultural Biotechnology, 1999. (in Russian)

Simon C.J., Weeden N.F. Elucidation of crab apple lineage by direct examination of rDNA sequences. *Malus*. 1991;5:4-6.

Solomatin N.M. The gene pool of vegetatively propagated forms of apple trees to improve the assortment of rootstocks, stock and ornamental cultivars in the Central Chernozem Region: Dr. (Agric.) Dissertation. Moscow, 2018. (in Russian)

Suprun I.I., Tokmakov S.V., Risovannaya V.I., Volodin V.A., Sherbatko V.D. Identification of apple tree *Pl1* and *Pl2* genes for resistance to mildew and *Vf* and *Vm* genes for resistance to scab in apple tree varieties from the collection of genetic resources of Crimea. *Nauchnye Trudy SKZNIISiV = Proceedings of the North Caucasian Research Institute for Horticulture and Viticulture*. 2015;7:15-19. (in Russian)

Urbanovich O., Kozlovskaya Z., Kartel N. Occurrence of mildew resistance genes in the collection of apple cultivars and species grown in Belarus. *Molekulyarnaya i Prikladnaya Genetika = Molecular and Applied Genetics*. 2010;11:20-25. (in Russian)

Visser T., Verhaegh J.J. Review of tree fruit breeding carried out at the Institute for Horticultural Plant Breeding at Wageningen from 1951–1976. In: Proceedings of the Eucarpia Meeting of Tree Fruit Breeding. Wageningen, 1976;113-132.

Yakuba G.V. Protection of apple tree from the powdery mildew under conditions of weather stresses. *Zashchita i Karantin Rasteniy = Plant Protection and Quarantine*. 2018;7:44-46. (in Russian)

ORCID ID

I.N. Shamshin orcid.org/0000-0002-4464-1876
M.L. Dubrovsky orcid.org/0000-0003-0883-2867
A.A. Trifonova orcid.org/0000-0001-9618-5932
K.V. Boris orcid.org/0000-0002-8479-4949
A.M. Kudryavtsev orcid.org/0000-0001-6029-0730

Acknowledgements. This research was funded by the Russian Science Foundation project No. 22-26-00326 (https://rscf.ru/project/22-26-00326).

Conflict of interest. The authors declare no conflict of interest.

Received March 27, 2023. Revised May 22, 2023. Accepted May 24, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):582-590
DOI 10.18699/VJGB-23-70

Original Russian text https://vavilovj-icg.ru/

Soft wheat cultivars grown in the Saratov region and their resistance to Septoria blotch

Yu.V. Zeleneva[1] ✉, E.A. Konkova[2]

[1] All-Russian Research Institute of Plant Protection, Pushkin, St. Petersburg, Russia
[2] Federal Center of Agriculture Research of the South-East Region, Saratov, Russia
✉ zelenewa@mail.ru

Abstract. Septoria is one of the harmful diseases of wheat cultivars cultivated in the Saratov region. This infectious disease of fungal etiology limits yield indicators and rapidly progresses in many regions of the Russian Federation. The aim of the research was to assess the resistance of winter and spring wheat cultivars that are referred to as promising and recommended for cultivation in the Low Volga region of the Russian Federation to pathogens of Septoria, to study the populations of *Parastagonospora nodorum* and *P. pseudonodorum* in the territory of the Saratov region in order to detect the presence of effector genes. Using molecular markers, we performed the identification of genes encoding NEs in 220 *Parastagonospora* spp. fungal isolates obtained from 7 cultivars of soft winter wheat, 6 taken from the winter triticale, 5 from soft spring wheat, 3 from durum spring wheat and 1 from spring oats. Among the *P. nodorum* isolates studied, there were both single genes *Tox1*, *Tox3*, and *ToxA*, and combinations of two genes in one genotype. The presence of the *ToxA* gene was not noted in the genotype of *P. pseudonodorum* isolates. During 2020–2022, a collection of winter and spring wheat cultivars was studied to detect resistance to Septoria blotch in field conditions (13 cultivars of winter wheat and 7 cultivars of spring wheat accordingly). The resistance of the cultivars was proven by laboratory evaluation. Three inoculums were used, including the isolates of *Z. tritici*, *P. nodorum* (*ToxA*, *Tox1*, *Tox3*), *P. pseudonodorum* (*ToxA*, *Tox1*, *Tox3*) mainly obtained from Saratov populations of 2022 (except for *P. pseudonodorum* with the *ToxA* gene). The tested cultivars were characterized using the *Xfcp623* molecular marker, diagnostic for *Tsn1*/*tsn1* genes, which controls sensitivity to the fungal toxin of PtrToxA. Of greatest interest are 11 wheat genotypes that showed resistance to one, two and three species which served as causative agents of Septoria blotch (*Zymoseptoria tritici*, *P. nodorum*, *P. pseudonodorum*). These are the soft winter wheat cultivars Gostianum 237 (*tsn1*), Lutescens 230 (*Tsn1*), Guberniya (*Tsn1*), Podruga (*Tsn1*), Anastasia (*Tsn1*), Sosedka (*Tsn1*) and the soft spring wheat cultivars Favorit (*tsn1*), Prokhorovka (*tsn1*), Saratovskaya 70 (*tsn1*), Saratovskaya 73 (*tsn1*), Belyanka (*tsn1*). The results obtained are of interest as they might increase the efficiency of selection based on the elimination of genotypes with dominant *Tsn1* alleles sensitive to PtrToxA. In addition to the economic value of the cultivars studied, it is recommended to use them in breeding for resistance to Septoria blotch.
Key words: effector genes; PCR-diagnosis; wheat selection; Septoria blotch; phytopathogenic fungi; PtrToxA; PtrTox1; PtrTox3.

For citation: Zeleneva Yu.V., Konkova E.A. Soft wheat cultivars grown in the Saratov region and their resistance to Septoria blotch. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):582-590. DOI 10.18699/VJGB-23-70

Устойчивость сортов мягкой пшеницы, возделываемых на территории Саратовской области, к возбудителям септориозных пятнистостей

Ю.В. Зеленева[1] ✉, Э.А. Конькова[2]

[1] Всероссийский научно-исследовательский институт защиты растений, Пушкин, Санкт-Петербург, Россия
[2] Федеральный аграрный научный центр Юго-Востока, Саратов, Россия
✉ zelenewa@mail.ru

Аннотация. Септориоз – одна из вредоносных болезней сортов пшеницы, возделываемых на территории Саратовской области. Это инфекционное заболевание грибной этиологии лимитирует показатели урожайности и быстро прогрессирует во многих регионах Российской Федерации. Целью исследований было оценить устойчивость перспективных и рекомендуемых для возделывания на территории Нижневолжского региона РФ сортов озимой и яровой мягкой пшеницы к возбудителям септориозных пятнистостей и изучить популяции *Parastagonospora nodorum* и *P. pseudonodorum*, распространенных на территории Саратовской области, по наличию генов-эффекторов. С применением молекулярных маркеров проведена идентификация генов, кодирующих

некротрофные эффекторы (NEs), у 220 изолятов гриба *Parastagonospora* spp., полученных с сортообразцов озимой и яровой мягкой пшеницы, яровой твердой пшеницы, озимого тритикале и ярового овса. Среди изученных изолятов *P. nodorum* были как единичные гены *Tox1*, *Tox3* и *ToxA*, так и сочетания из двух генов в одном генотипе. В генотипе изолятов *P. pseudonodorum* не отмечено присутствие гена *ToxA*. Изучено 20 сортов озимой и яровой пшеницы на устойчивость к септориозным пятнистостям в лабораторных условиях и в поле в течение 2020–2022 гг. Было использовано три инокулюма, включающих изоляты *Zymoseptoria tritici*, *P. nodorum* (*ToxA*, *Tox1*, *Tox3*) и *P. pseudonodorum* (*ToxA*, *Tox1*, *Tox3*). Анализируемые сорта были охарактеризованы с помощью молекулярного маркера *Xfcp623*, диагностического для генов *Tsn1*/*tsn1*, контролирующего чувствительность к токсину гриба PtrToxA. Наибольший интерес представляют 11 генотипов пшеницы, которые показали устойчивость к одному, двум и трем видам – возбудителям септориоза (*Z. tritici*, *P. nodorum*, *P. pseudonodorum*). Это сорта озимой мягкой пшеницы: Гостианум 237 (*tsn1*), Лютесценс 230 (*Tsn1*), Губерния (*Tsn1*), Подруга (*Tsn1*), Анастасия (*Tsn1*), Соседка (*Tsn1*) и яровой мягкой пшеницы: Фаворит (*tsn1*), Прохоровка (*tsn1*), Саратовская 70 (*tsn1*), Саратовская 73 (*tsn1*), Белянка (*tsn1*). Полученные результаты важны для повышения эффективности селекции на основе элиминации генотипов с доминантными аллелями *Tsn1*, чувствительными к грибу PtrToxA. Помимо хозяйственной ценности изученных сортов, их рекомендуется использовать в селекции на устойчивость к септориозной пятнистости.
Ключевые слова: гены-эффекторы; ПЦР-диагностика; селекция пшеницы; септориозы; фитопатогенные грибы; PtrToxA; PtrTox1; PtrTox3.

Introduction

The Saratov region is a large administrative district; it includes 37 municipal districts, distributed based on climatic conditions between the right-bank black soil and arid steppe regions on the left bank of Volga. Differences in climatic conditions affect the yield of agricultural crops and the damage caused by diseases, among which the dominant position in the phytopathogenic complex is given to Septoria and pyrenophorous wheat blotch. During the years of epiphytoties caused by Septoria blotch in different regions of Russia, North America, Australia and other parts of the world, crop losses can exceed 30–40 % (Ficke et al., 2018; Sanin et al., 2018).

The annual monitoring implemented shows that in recent years in many regions of Russia, including the Saratov region, the pathogenic complex of wheat Septoria blotch has been dominated by the species of *Zymoseptoria tritici* (Desm.) Quaedvl. et Crous, representing the causative agent of Septoria leaf blotch on wheat, triticale, barley and rye (Zeleneva et al., 2022).

Less notable species are *Parastagonospora nodorum* (Berk.) Quaedvl., Verkley et Crous. They parasitize leaves, stems, glumes, awns of wheat and other cereals (Pakholkova, 2003; Sanin et al., 2018; Zeleneva et al., 2022). It quite often affects the seeds as well. The caryopsis turns puny, defective, its germination rate decreases. In case of strong infection, the coleoptiles get damaged and die off.

Another species, *P. pseudonodorum*, has a strict host specialization. It is a wheat parasite. Until recently, this species was considered a wheat form of *P. avenae* (A.B. Frank) Quaedvl., Verkley et Crous: *P. avenae* f. sp. *triticea*. However, based on the study of morphology using the methods of multilocus phylogeny in modern systematics, the species of *P. pseudonodorum* is one of the seven that have been described so far. In total, 26 species of *Parastagonospora* have been detected by now using phylogenetic analysis (Croll et al., 2021).

The fungi of *P. nodorum* and *P. pseudonodorum* are known for their ability to synthesize necrotrophic effectors (NEs), including host selective toxins (HSTs), that function as pathogenicity factors (Ciuffetti et al., 1997). Sensitivity towards NEs does not always lead to susceptibility of wheat to a Septoria pathogen (van Schie, Takken, 2014; Virdi et al., 2016).

Today, in total there are nine characterized interactions within the pathosystem of wheat – *P. nodorum*: *Tsn1 – SnToxA* (Friesen et al., 2009; Zhang et al., 2009; Faris et al., 2011); *Snn1 – SnTox1* (Shi et al., 2016b); *Snn2 – SnTox267* (Richards et al., 2022); *Snn3-B1 – SnTox3* (Shi et al., 2016a); *Snn3-D1 – SnTox3* (Zhang et al., 2011); *Snn4 – SnTox4* (Abeysekara et al., 2012); *Snn5 – SnTox5* (Sharma, 2019; Kariyawasam et al., 2022); *Snn6 – SnTox267* (Richards et al., 2022); *Snn7 – SnTox267* (Richards et al., 2022). It has been shown that genes encoding SnToxA, SnTox1, SnTox3 proteins are present in the genotype of *P. pseudonodorum* (Hafez et al., 2020; Navathe et al., 2020).

Right now there are three cloned host genes, including *Tsn1* (Faris et al., 2010), *Snn1* (Shi et al., 2016b) and *Snn3-D1* (Zhang et al., 2021). There are also five fungus genes that encode effector proteins: *SnToxA* (Friesen et al., 2009), *SnTox3* (Liu et al., 2009), *SnTox1* (Liu et al., 2012), *SnTox5* (Kariyawasam et al., 2022), and *SnTox267* (Richards et al., 2022).

The increase in the aggression of crop culture fungal diseases was remarked in the Saratov region during the last decade. That is why the continuous search and use of new effective genetic sources and donors was and still remains the prioritized approach in wheat immunity selection in the Low Volga region (Konkova et al., 2022). The present research is annually conducted in the territory of the base provided by the Federal agrarian scientific centre of the South-East (in the city of Saratov). Starting from 2021, the selection of sources and donors expressing susceptibility to Septoria blotch has been conducted using molecular technologies that allow to pick genotypes with specific gene combinations.

The aim of the given research is to assess the resistance of soft winter and spring wheat cultivars recommended for cultivation in the territory of the Lower Volga region of the Russian Federation to Septoria blotch pathogens together with studying the populations of *P. nodorum* and *P. pseudonodorum* spread in the Saratov region territory to detect the presence of effector genes.

Materials and methods

The affected plant samples were collected in 2021–2022 in the Saratov region territory. An infectious sample was understood as plant leaves with well-pronounced symptoms of Septoria

blotch, collected on the observed field along its diagonal plane at equal distances and a certain time period (for example, during the counting process).

To collect the samples of affected plants, crop observations were implemented in areas mentioned in Table 1. All the samples were collected during their maturation phase, at the stage of milky-wax ripeness of plants (75–85 according to the Zadok's scale). The leaves with typical external signs of Septoria blotch disease were picked. The collected material was herbarized and labeled (indicating the place and date of collecting, the phase, plant species and cultivar, information concerning disease symptoms, culture cultivation technologies, information regarding protection measures). Subsequently, the infectious samples of grain crops (leaves) of wheat, triticale and oats were analyzed in laboratory conditions to identify the species composition of Septoria blotch pathogens (Pyzhikova et al., 1989).

Meteorological conditions of 2020–2022 in the region had a beneficial effect on the development of Septoria blotch causative agents of grain crops. According to the Saratov meteorological station data, at the beginning of the vegetation period in May, considering the three-year average, there was a 30.5 mm precipitation fallout. At moderate air temperatures, the hydrothermal coefficient (HTC) was quite high – 1.3. In June, the precipitation amount (34.5 mm) and HTC (0.55) decreased. In the middle of vegetation, in July, the situation improved significantly. Precipitation fallout was 97.2 mm, which is much higher than the norm, and the hydrothermal coefficient was high, showing 1.54. This contributed to the growth and development of agricultural plants and had a positive effect on the development of the phytopathogenic complex. In August, the precipitation level was low – 12.6 mm, on average, within the three-year period. Elevated air temperatures were noted – the number of days with a maximum air temperature above or equal to 30 °C was 17. The hydrothermal coefficient in this month was also extremely low – 0.2, indicating arid conditions.

The degree of damage of the infectious material selected for analysis varied from 30 to 40 %. The species of *Z. tritici* was isolated from all the samples of the infectious material subject to the study. It was possible to obtain monoconidial isolates of fungi of the *Parastagonospora* genus from some samples (see Table 1) (Pyzhikova et al., 1989).

The samples were analyzed in laboratory conditions to identify the species composition of Septoria blotch causative agents. The results of laboratory diagnostics of the specific affiliation of the pathogen were confirmed by the sequencing method using the equipment of the Collective Use Center "Genomic Technologies, Proteomics and Cell Biology" of the All-Russian Research Institute of Agricultural Meteorology.

The analysis included 220 monoconidial isolates of the *Parastagonospora* genus, taken in the amount of 10 from each of the 22 infectious samples. To assess the resistance to Septoria blotch, 13 samples of winter wheat were used (Gostianum 237, Lutescens 230, Saratovskaya 8, Gubernia, Mironovskaya 808, Donskaya bezostaya, Saratovskaya 90, Zhemchuzhina Povolzhya, Saratovskaya 17, Kalach 60, Podruga, Anastasia, Sosedka) and seven samples of spring wheat (Favorit, Prokhorovka, Yugo-Vostochnaya 2, Saratovskaya 70, Saratovskaya 73, Belyanka, Lebyodushka). Screening was carried out on the test fields of the FASC of the South-East in the conditions of a natural infectious background in 2020–2022. A modified and supplemented Saari–Prescott scale was used (Kolomiets et al., 2017). All the cultivars were divided into five groups: RR – highly resistant (damage < 11 %); R – resistant (damage 11–20 %); MS – moderately susceptible (damage 21–40 %); S – susceptible (damage 41–70 %); HS – highly susceptible (damage 71–100 %).

The laboratory evaluation was implemented using isolated leaves as described by G.V. Pyzhikova and E.V. Karaseva (1985). Three inoculums of *Z. tritici*, *P. nodorum* (*ToxA*, *Tox1*, *Tox3*), *P. pseudonodorum* (*ToxA*, *Tox1*, *Tox3*) isolates were used. When inoculating plants under laboratory conditions, fungal isolates of the Saratov population obtained from the infectious material of 2022 were used: *Z. tritici*: 80-22-Z.t – durum spring wheat host, 80-22-Z.t – soft winter wheat, 95-22-Z.t – winter wheat; *P. nodorum*: 80-22-P.n. (*Tox3*) – durum spring wheat host, 101-22-P.n. (*ToxA*, *Tox3*) – soft spring wheat, 88-22-P.n. (*Tox1*, *Tox3*) – winter wheat; *P. pseudonodorum*: 72-22-P.ps. (*Tox1*, *Tox3*) – winter triticale, 89-22-P.ps. (*Tox1*). The presence of *Tox* genes in the genotypes of the *Parastagonospora* spp. fungi isolates used was detected for the first time and the results are presented in this work. An isolate of the Tambov population 82-21-P.ps., the genotype of which contains *ToxA*, was added to the infectious material of the *P. pseudonodorum* species. This isolate was obtained from the soft spring wheat Voronezhskaya 20 in 2021 (Zeleneva et al., 2022).

Fungal genomic DNA was isolated from a pure culture of monoconidial isolates obtained on potato glucose agar (CGA) using the standard CTAB method (Doyle J.J., Doyle J.L., 1990). The same method was chosen to extract DNA from young leaves of 13 winter and 7 soft spring wheat cultivars.

Amplification of genomic DNA was carried out in 25 µl of the reaction mixture (2 µl of genomic DNA (25 ng (permittable from 2 to 50 ng)), 1 µl of each primer (10 pM/µl) (Evrogen, Russia), 0.5 µl of dNTPs mix (10 mM, aqueous solution of dCTP, dGTP, dTTP and dATP) (TransGen, China), 0.55 µl $MgCl_2$ (100 mM), 0.5 µl BioTaq DNA polymerase (5U, 5 U/µl) (Dialat Ltd., Russia), 2.5 µl 10X PCR buffer, 17 µl ddH_2O).

The amplified fragments were separated by electrophoresis in 1.5 % of agarose gel, in 1× TBE buffer (pH 8.2), the gel was stained with ethidium bromide. The DNA marker Step100 plus (Biolabmix, Russia) was used to assess the fragment size.

Screening of isolates of the *Parastagonospora* genus to detect the presence of effector genes: *ToxA*, *Tox1* and *Tox3* was performed using PCR. To obtain statistically valid results, the DNA of ten monoconidial isolates obtained from each infectious sample was analyzed (see Table 1). A total of 220 DNA samples were analyzed. The list of primers for PCR is presented in Table 2.

Screening of genotypes of wheat cultivars aimed at revealing the presence of a dominant or recessive gene (*Tsn1/tsn1*) was put at work according to the method of using PCR with primer pairs *Xfcp623*(*F*)/*Xfcp623*(*R*) (Faris et al., 2010). The presence of the marker amplification product indicates the presence of the dominant allele of the *Tsn1* gene (plant susceptibility to the PtrToxA fungal toxin protein), its absence

Table 1. The origin of the analysed monoconidial isolates of *Parastagonospora* spp. in 2021–2022

No.	Isolate name	The origin of the infectious sample/host plant
1	32-21-P.n.-1…10	Saratov region, test field of FASC of the South-East/ soft winter wheat Anastasia 51°34′28″N, 46°00′20″E
2	33-21-P.ps.-1…10	Saratov region, test field of FASC of the South-East/ durum spring wheat, hybrid line 51°34′38″N, 45°59′51″E
3	35-21-P.ps.-1…10	Saratov region, test field of FASC of the South-East/ winter triticale 51°34′31″N, 46°00′24″E
4	37-63-21-C-P.av.-1…10	Saratov region, test field of FASC of the South-East/spring oats Skakun 51°35′58″N, 46°02′36″E
5	80-22 P.n.-1…10	Saratov region, test field of FASC of the South-East/ durum spring wheat, hybrid line 51°35′58″N, 46°02′36″E
6	82-22 P.n.-1…10	Saratov region, test field of FASC of the South-East/ durum spring wheat, hybrid line 51°34′41″N, 45°59′54″E
7	86-22 P.n.-1…10	Saratov region, test field of FASC of the South-East/ soft spring wheat, introgressive line 51°34′47″N, 45°59′51″E
8	88-22 P.n.-1…10	Saratov region, Baltaysky district/ soft winter wheat Saratovskaya 17 52°28′49″N, 46°33′40″E
9	91-22 P.n.-1…10	Saratov region, Baltaysky district/ soft winter wheat Kalach 60 52°28′21″N, 46°39′51″E
10	92-22 P.n.-1…10	Saratov region, Balashovsky district/ soft winter wheat Levoberezhnaya 1 51°29′25″N, 43°27′13″E
11	93-22 P.n.-1…10	Saratov region, test field of FASC of the South-East/ soft winter wheat, hybrid line 51°34′35″N, 45°59′53″E
12	98-22 P.n.-1…10	Saratov region, Arkadaksky district/ soft winter wheat Kalach 60 51°52′25″N, 43°35′04″E
13	101-22 P.n.-1…10	Saratov region, Yershovsky district/ soft spring wheat Kvartet 51°22′31″N, 48°12′18″E
14	259-22 P.n.-1…10	Saratov region, Yershovsky district/ soft spring wheat Yugo-Vostochnaya 4 51°22′20″N, 48°12′35″E
15	260-22 P.n.-1…10	Saratov region, Arkadaksky district/ soft spring wheat Kvartet 51°52′20″N, 43°33′45″E
16	261-22 P.n.-1…10	Saratov region, Pugachyovsky district/ soft spring wheat Yugo-Vostochnaya 2 52°02′30″N, 49°13′27″E
17	71-22-P.ps.-1…10	Saratov region, Pugachyovsky district/ winter triticale 52°02′01″N, 49°18′05″E
18	72-22-P.ps.-1…10	Saratov region, Balashovsky district/ winter triticale 51°28′43″N, 43°17′04″E
19	73-22-P.ps.-1…10	Saratov region, test field of FASC of the South-East/ winter triticale 51°34′28″N, 46°00′24″E
20	74-22-P.ps.-1…10	Saratov region, test field of FASC of the South-East/ winter triticale 51°34′27″N, 46°00′27″E
21	76-22-P.ps.-1…10	Saratov region, test field of FASC of the South-East/ winter triticale 51°34′31″N, 46°00′28″E
22	89-22-P.ps.-1…10	Saratov region, test field of FASC of the South-East/ soft winter wheat, hybrid line 51°34′28″N, 46°00′20″E

Table 2. List of primers for PCR

Locus	Primer	Sequence 5′–3′	Reference	Amplicon size, bp
Tox1	*SnTox1cF*	ATGAAGCTTACTATGGTCTTGT	Gao et al., 2015	500
	SnTox1cR	TGTGGCAGCTAACTAGCACA		
Tox3	*SnTox3cF*	CTCGAACCACGTGGACCCGGA		600
	SnTox3cR	CTCCCCTCGTGGGATTGCCCCATATG		
ToxA	*TA51F*	GCGTTCTATCCTCGTACTTC	Andrie et al., 2007	573
	TA52R	GCATTCTCCAATTTTCACG		
Tsn1	*Xfcp623F*	CTATTCGTAATCGTGCCTTCCG	Faris et al., 2010	380
	Xfcp623*R*	CCTTCTCTCTCACCGCTATCTCATC		

signifies the presence of the recessive *tsn1* allele (plant resistance to PtrToxA).

Statistical data processing was carried out using the computer program STATISTICA 12. The average damage of the leaf blade by Septoria blotch was calculated during the field assessment within the period of 2020–2022, %; SD – standard deviation (Std. Dev.). The *Q*-Cochran criterion was taken to divide the studied wheat cultivars according to resistance/susceptibility to three pathogens of Septoria blotch. This criterion was used to test a significant difference between phytopathological assessments of wheat cultivars.

Results

As a result of molecular screening, in the studied material (220 DNA probes obtained from 130 monoconidial isolates of *P. nodorum*, 80 *P. pseudonodorum* and 10 *P. avenae* isolates), both single genes encoding NEs and their combinations in one genotype were identified (Fig. 1, Supplementary Material 1)[1].

The *ToxA* gene was identified among monoconidial isolates of the *P. nodorum* species (93-22-P.n.-1…10) obtained from the leaves of a hybrid line of soft winter wheat from the experimental field of the FASC of the South-East and from the infectious material of soft spring wheat Kvartet (110-22-P.n.-1…10) from the Yershovsky district of the Saratov region (see Fig. 1, *a*).

As a result of molecular screening, the *Tox1* gene was identified among the isolates obtained from four cultivars of wheat and five triticale cultivars. The presence of the gene was noted in *P. nodorum* isolates from the soft winter wheat cultivar of Saratovskaya 17 (88-22-P.n.-1…10) and Levoberezhnaya 1 (92-22-P.n.-1…10) from the Balashovsky district. The presence of the *SnTox1* gene was marked in *P. pseudonodorum* monoconidial isolates taken from infectious material from hybrid lines of durum spring wheat (33-21-P.ps.-1…10) and soft winter wheat (89-22-P.ps.-1 ...10) from the experimental field of the FASC of the South-East, as well as winter triticale cultivars from the Pugachyovsky district (71-22-P.ps.-1 ... 10), Balashovsky district (72-22-P.ps.-1 ... 10) and from the experimental field of the FANC of the South-East (73-22-P.ps.-1…10; 74-22-P.ps.-1…10; 76-22-P.ps.-1…10) (see Fig. 1, *b*).

The presence of the *Tox3* gene was detected among isolates of the *P. pseudonodorum* species obtained from plant samples of winter triticale from the Balashovsky district and the experimental field of the FASC of the South-East (72-22-P.ps.-1…10, 73-22-P.ps.-1…10, respectively). The presence of the *Tox3* gene was marked in isolates obtained from winter wheat cultivar Anastasia (32-21-P.n.-1...10), Saratov-

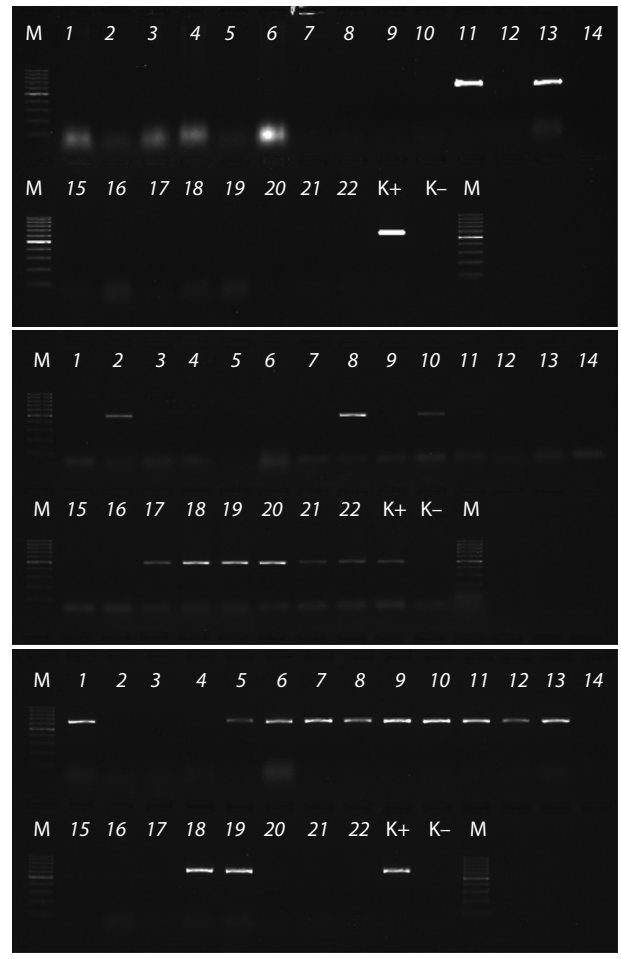

Fig. 1. Electropherogram of amplification products obtained using markers specific to the *ToxA*, *Tox1*, and *Tox3* genes of *Parastagonospora* spp. from the Saratov population.

a – *ToxA*, amplicon size 573 bp; *b* – *Tox1*, amplicon size 500 bp; *c* – *Tox3*, amplicon size 600 bp. M – Step100 plus marker (Biolabmix). The index numbers assigned to the samples correlate with the indexes in Table 1.

[1] Supplementary Materials 1 and 2 are available in the online version of the paper: https://vavilov.elpub.ru/jour/manager/files/Suppl_Zeleneva_Engl_27_6.pdf

Table 3. Damage intensity caused by leaf diseases in soft spring and winter wheat cultivars

No.	Cultivar name	Field phytopathological assessment of cultivars (2020–2022), %		Laboratory assessment on isolated leaves, %					
				Z. tritici		*P. nodorum* (*ToxA*, *Tox1*, *Tox3*)		*P. pseudonodorum* (*ToxA*, *Tox1*, *Tox3*)	
		Mean score ± SD	Phenotype	Mean score ± SD	Phenotype	Mean score ± SD	Phenotype	Mean score ± SD	Phenotype
		Soft winter wheat (*Triticum aestivum* L.)							
1	Gostianum 237 ***(tsn1)***	22 ± 11.6	MS	18 ± 7.6	R	30 ± 0	MS	23 ± 2.7	MS
2	Lutescens 230 (*Tsn1*)	17 ± 10.4	R	16 ± 5.5	R	29 ± 2.2	MS	29 ± 2.2	MS
3	Saratovskaya 8 (*Tsn1*)	18 ± 13.7	R	29 ± 2.2	MS	36 ± 5.5	MS	36 ± 5.5	MS
4	Gubernia* (*Tsn1*)	9 ± 6.1	RR	5 ± 0.9	RR	38 ± 4.5	MS	34 ± 5.5	MS
5	Mironovskaya 808* ***(tsn1)***	15 ± 5	R	47 ± 4.5	S	36 ± 5.5	MS	32 ± 4.5	MS
6	Donskaya bezostaya* (*Tsn1*)	25 ± 5	MS	30 ± 10	MS	36 ± 5.5	MS	40 ± 0	MS
7	Saratovskaya 90* (*Tsn1*)	27 ± 5.8	MS	28 ± 2.7	MS	38 ± 4.5	MS	40 ± 0	MS
8	Zhemchuzhina Povolzhya* (*Tsn1*)	27 ± 5.8	MS	40 ± 0	MS	40 ± 0	MS	40 ± 0	MS
9	Saratovskaya 17* (*Tsn1*)	40 ± 10	S	41 ± 2.2	S	34 ± 5.5	MS	40 ± 0	MS
10	Kalach 60* (*Tsn1*)	16 ± 13.5	R	21 ± 5.5	MS	40 ± 0	MS	22 ± 4.5	MS
11	Podruga* (*Tsn1*)	11 ± 6.9	R	15 ± 7.1	R	30 ± 0	MS	20 ± 0	R
12	Anastasia* (*Tsn1*)	10 ± 8.7	RR	5 ± 0	RR	19 ± 2.2	R	14 ± 5.5	R
13	Sosedka (*Tsn1*)	13 ± 10.4	R	7 ± 2.7	RR	30 ± 0	MS	20 ± 2.7	R
		Soft spring wheat (*Triticum aestivum* L.)							
14	Favorit* ***(tsn1)***	4 ± 1.2	RR	4 ± 1.1	RR	36 ± 5.5	MS	10 ± 0	R
15	Prokhorovka* ***(tsn1)***	4 ± 1.2	RR	18 ± 2.7	R	40 ± 0	MS	28 ± 4.5	MS
16	Yugo-Vostochnaya 2* ***(tsn1)***	52 ± 12.6	S	50 ± 0	S	22 ± 4.5	MS	40 ± 0	MS
17	Saratovskaya 70* ***(tsn1)***	8 ± 2.9	RR	12 ± 2.7	R	23 ± 2.7	MS	22 ± 2.7	MS
18	Saratovskaya 73* ***(tsn1)***	8 ± 5.8	RR	20 ± 10	R	16 ± 2.2	R	18 ± 2.7	R
19	Belyanka* ***(tsn1)***	15 ± 5	R	14 ± 1.1	R	24 ± 5.5	MS	12 ± 2.7	R
20	Lebyodushka* (*Tsn1*)	33 ± 5.8	MS	32 ± 4.5	MS	21 ± 2.2	MS	36 ± 5.5	MS

* The cultivar is admitted to cultivation in the territory of the Lower Volga region of the Russia Federation (region 8).

skaya 17 (88-22-P.n.-1...10), Kalach 60 (91-22-P.n.-1...10; 98-22-P.n.-1…10), Levoberezhnaya 1 (92-22-P.n.-1…10), hybrid line (93-22-P.n.-1…10); from spring wheat of hybrid lines (80-22-P.n.-1…10; 82-22-P.n.-1…10), introgressive line (86-22-P.n.-1…10), from the Kvartet cultivar (101-22-P.n.-1…10) (see Fig. 1, *c*).

In the course of three-year tests on a natural infectious background, cultivars showing resistance or weak susceptibility to Septoria blotch were detected (Table 3).

Genotyping of wheat cultivars using a molecular marker was aimed at identifying carriers of genes that control sensitivity and resistance to the PtrToxA toxin. The *Xfcp623* marker amplified a 380 bp. fragment associated with the *Tsn1* gene sensitive to the PtrToxA toxin in 12 cultivars of soft winter wheat: Lutescens 230, Saratovskaya 8, Gubernia, Donskaya bezostaya, Saratovskaya 90, Zhemchuzhina Povolzhya, Saratovskaya 17, Kalach 60, Podruga, Anastasia, Sosedka and one cultivar of soft spring wheat, Lebyodushka. The genotypes of two cultivars of soft winter wheat: Gostianum 237 and Mironovskaya 808; six cultivars of soft spring whea: Favorit, Prokhorovka, Yugo-Vostochnaya 2, Saratovskaya 70, Saratovskaya 73, and Belyanka represent carriers of the recessive allele of the *tsn1* gene and are protected against PtrToxA at their genetic level (see Fig. 2, Table 3).

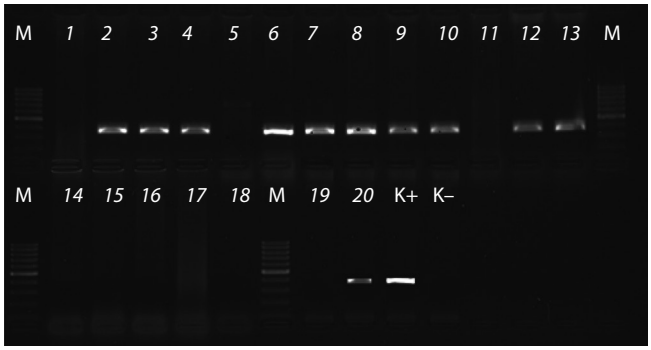

Fig. 2. Electrophoregram of *Tsn1* gene amplification products in soft winter and spring wheat cultivars (amplicon size 380 bp).

The numbers indicated for the samples correspond to the list of cultivars in Table 3 (positive control (K+) – the cultivar of Glenlea, negative control (K–) – line 6B365).

A laboratory test of cultivars concerning the reaction to three pathogens of Septoria blotch, typical for the region (*Z. tritici*, *P. nodorum* (*ToxA*, *Tox1*, *Tox3*), *P. pseudonodorum* (*ToxA*, *Tox1*, *Tox3*)) was carried out. The infectious material of the regional populations of 2022 was used for inoculation. The results are presented in Table 3.

When wheat samples were infected with *Z. tritici*, the following cultivars performed well: Gubernia, Anastasia, Sosedka, Favorit. Their degree of damage did not exceed 7 % on average, they were included in the group of highly resistant cultivars (RR). The degree of damage by *Z. tritici* to Gostianum, Lutescens 230, Podruga, Prokhorovka, Saratovskaya 70, Saratovskaya 73, and Belyanka did not exceed 20 %, which made it possible to classify the cultivars as the members of the resistant group (R).

Two cultivars were confirmed to be resistant to *P. nodorum*: Anastasia and Saratovskaya 73. Six cultivars were confirmed to be resistant to *P. pseudonodorum*: Podruga, Anastasia, Sosedka, Favorit, Saratovskaya 73 and Belyanka.

When using the statistical method of correlation analysis, a weak direct relation was established between the indicators of the presence of the *Tsn1* gene in the wheat cultivar genotype and the intensity of its damage caused by *P. nodorum* and *P. pseudonodorum* species containing the *ToxA* gene in isolates included in the inoculum (the correlation coefficient is 0.3 and 0.2, respectively).

A strong direct correlation was noted between the indicators of the overall degree of leaf blade damage caused by Septoria blotch in the field and the degree of damage to wheat samples caused by *Z. tritici* (0.8) and *P. pseudonodorum* (0.7) in the laboratory.

The indicators of the degree of damage to wheat cultivars of *Z. tritici* had a direct relation with the degree of *P. pseudonodorum* damage (0.77); *P. nodorum* and *P. pseudonodorum* (0.4); *Z. tritici* and *P. nodorum* (0.2).

The *Q*-Cochran criterion made it possible to divide the studied wheat cultivars into four groups based on the criterion of resistance to three pathogens: 1 – lack of resistance to Septoria blotch pathogens; 2 – resistance to one species; 3 – resistance to two species; 4 – resistance to three species of pathogens. The value of the coefficient $Q = 36.35$ with the significance level of p less than 0.009 indicates that the cultivars differed significantly from each other in terms of resistance/susceptibility to Septoria blotch pathogens of *Z. tritici*, *P. nodorum*, *P. pseudonodorum.* The test results are presented in Table 4 and Supplementary Material 2.

Table 4. Non-parametric statistical analysis of indicators of phytopathological evaluation of cultivars to Septoria blotch pathogens

Cultivars	*Q*-Cochran criterion. Resistance to three phytopathogens among 20 wheat cultivars was studied (df = 19); $Q = 36.35$; $p < 0.009$
	Name of the pathogen, to which resistance is shown
Soft winter wheat (*Triticum aestivum* L.)	
Gostianum 237, Lutescens 230, Gubernia	*Z. tritici*
Podruga, Sosedka	*Z. tritici*, *P. pseudonodorum*
Anastasia	*Z. tritici*, *P. nodorum*, *P. pseudonodorum*
Soft spring wheat (*Triticum aestivum* L.)	
Prokhorovka, Saratovskaya 70	*Z. tritici*
Favorit	*Z. tritici*, *P. nodorum*
Belyanka	*Z. tritici*, *P. pseudonodorum*
Saratovskaya 73	*Z. tritici*, *P. nodorum*, *P. pseudonodorum*

Discussion

Septoria blotch represents a dangerous wheat disease; it is one of the most harmful in the fields of the Saratov region. In 2017, a strong epiphytoty of Septoria blotch was recorded on winter wheat crops (the damage was equal to 67 %). In 2018–2019, the intensity of *Z. tritici* damage did not exceed 25 %. Septoria blotch infection, which exceeded the threshold of 40 %, was noted in 2020 – 45 % and in 2021 – 41 % (Konkova et al., 2022).

The proposed study is one of the first in this region. It includes a comprehensive screening of area-specific and promising cultivars of soft winter and spring wheat, as well as molecular analysis that shows the presence of genes encoding NEs in plant pathogen populations and genes in plant genotypes that control disease resistance.

In the course of the study carried out, it was shown that among the genotypes of the studied isolates of *P. nodorum* and *P. pseudonodorum* of the Saratov population, there was a wide representation of the *Tox1* and *Tox3* genes, while the *ToxA* gene was recorded only in isolates 93-22-P.n.-1…10 and 101-22-P.n.-1…10 of the *P. nodorum* species. The results obtained are consistent with foreign publications (Richards et al., 2022) reporting that the prevalence of the *Tox267* and *Tox1* genes is significantly higher than that of *ToxA* in the genotypes of *P. nodorum* populations that are territorially distant from Russian ones.

It is known that *P. nodorum* is a donor of the *ToxA* gene for *Pyrenophora tritici-repentis*; they share a common toxin, PtrToxA (Ciuffetti et al., 1997). Recently, the *ToxA* gene was identified in the fungus that causes wheat blotch – *Bipolaris sorokiniana* (McDonald et al., 2017; Friesen et al., 2018). This means that the cultivars of winter wheat Gostianum 237 (*tsn1*), Mironovskaya 808 (*tsn1*) and spring wheat Favorit (*tsn1*), Prokhorovka (*tsn1*), Yugo-Vostochnaya 2 (*tsn1*), Saratovskaya 70 (*tsn1*), Saratovskaya 73 (*tsn1*), Belyanka (*tsn1*) are protected from the *ToxA* gene toxin of four dangerous phytopathogens at once (*P. tritici-repentis*, *P. nodorum*, *P. pseudonodorum*, *B. sorokiniana*).

In the work of N.M. Kovalenko and colleagues (2022), it is possible to see the results of identification of the *Tsn1/tsn1* allele using the *Xfcp623* molecular marker in 35 cultivars of winter and 31 cultivars of spring wheat, included in the State Register of Breeding Achievements in 2018–2020 for the first time. Out of them, only 9 cultivars of winter and 4 cultivars of spring wheat carried *Tsn1*, which indicates susceptibility to PtrToxA, while the remaining cultivars have protection against the toxin at the genetic level. We consider this a great achievement of national selection. In the work of T.L. Friesen and colleagues (2018), it was mentioned that maintaining *tsn1* in the genotypes of wheat cultivars admitted for selection not only provides a selective advantage over pathogens that currently carry *ToxA*, but may also exert selection pressure on newer or more suitable pathogens that acquire *ToxA* via horizontal transfer.

Conclusion

Thus, using molecular markers, the identification of genes encoding NEs in two species called *P. nodorum* and *P. pseudonodorum* from populations of the Saratov region was carried out. In monoconidial isolates, both single *Tox1*, *Tox3*, and *ToxA* genes, as well as combinations of two genes in one genotype, were noted. The presence of a characteristic amplification product suggests the presence of two NEs genes, *ToxA* and *Tox3*, in *P. nodorum* monoconidial isolates 93-22-P.n.-1…10 and 101-22-P.n.-1…10; *Tox1* and *Tox3* in isolates 88-22-P.n.-1…10, 92-22-P.n.-1…10. 72-22-P.ps.-1…10. One *Tox1* gene in isolate genotypes 33-21-P.n.-1…10, 71-22-P.ps.-1…10, 89-22-P.ps.-1…10. *Tox3* gene in genotypes 32-21-P.n.-1…10; 80-22-P.n.-1…10, 82-22-P.n.-1…10, 86-22-P.n.-1…10; 91-22-P.n.-1…10; 98-22-P.n.-1…10 and 73-22-P.ps.-1…10.

A collection of 20 cultivars (16 area-specific and 4 promising) was studied to detect resistance/susceptibility to Septoria blotch pathogens in the experimental field of the FASC of the South-East in the period of 2020–2022, as well as in laboratory conditions. The cultivars underwent PCR screening showing the presence of a dominant or recessive gene (*Tsn1/tsn1*), which controls sensitivity to the toxin of the fungus PtrToxA. For this reason, 11 cultivars with resistance to one, two or three types of phytopathogens (*Z. tritici*, *P. nodorum*, *P. pseudonodorum*) are of the greatest interest. These are the cultivars of the Saratov selection Anastasia (*Tsn1*), Belyanka (*tsn1*), Gostianum 237 (*tsn1*), Guberniya (*Tsn1*), Lutescens 230 (*Tsn1*), Podruga (*Tsn1*), Prokhorovka (*tsn1*), Saratovskaya 70 (*tsn1*), Saratovskaya 73 (*tsn1*), Sosedka (*Tsn1*) and Favorit (*tsn1*).

References

Abeysekara N.S., Faris J.D., Chao S., McClean P.E., Friesen T.L. Whole-genome QTL analysis of *Stagonospora nodorum* blotch resistance and validation of the SnTox4-*Snn4* interaction in hexaploid wheat. *Phytopathology*. 2012;102(1):94-104. DOI 10.1094/PHYTO-02-11-0040.

Andrie R.M., Pandelova I., Ciuffetti L.M. A combination of phenotypic and genotypic characterization strengthens *Pyrenophora tritici-repentis* race identification. *Phytopathology*. 2007;97(6):694-701. DOI 10.1094/PHYTO-97-6-0694.

Ciuffetti L.M., Tuori R.P., Gaventa J.M. A single gene encodes a selective toxin causal to the development of tan spot of wheat. *Plant Cell*. 1997;9(2):135-144. DOI 10.1105/tpc.9.2.135.

Croll D., Crous P.W., Pereira D., Mordecai E.A., McDonal B.A., Brunner P.C. Genome-scale phylogenies reveal relationships among *Parastagonospora* species infecting domesticated and wild grasses. *Persoonia*. 2021;46:116-128. DOI 10.3767/persoonia.2021.46.04.

Doyle J.J., Doyle J.L. Isolation of plant DNA from fresh tissue. *Focus*. 1990;12(1):13-15.

Faris J.D., Zhang Z., Lu H.J., Lu S.W., Reddy L., Cloutier S., Fellers J.P., Meinhardt S.W., Rasmussen J.B., Xu S.S., Oliver R.P., Simons K.J., Friesen T.L. A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. *Proc. Natl. Acad. Sci. USA*. 2010;107(30):13544-13549. DOI 10.1073/pnas.1004090107.

Faris J.D., Zhang Z., Rasmussen J.B., Friesen T.L. Variable expression of the *Stagonospora nodorum* effector *SnToxA* among isolates is correlated with levels of disease in wheat. *Mol. Plant Microbe Interact*. 2011;24(12):1419-1426. DOI 10.1094/MPMI-04-11-0094.

Ficke A., Cowger C., Bergstrom G., Brodal G. Understanding yield loss and pathogen biology to improve disease management: Septoria nodorum blotch – a case study in wheat. *Plant Dis*. 2018;102(4):696-707. DOI 10.1094/PDIS-09-17-1375-FE.

Friesen T.L., Chu C.G., Liu Z.H., Xu S.S., Halley S., Faris J.D. Host-selective toxins produced by *Stagonospora nodorum* confer disease susceptibility in adult wheat plants under field conditions. *Theor. Appl. Genet*. 2009;118(8):1489-1497. DOI 10.1007/s00122-009-0997-2.

Friesen T.L., Holmes D.J., Bowden R.L., Faris J.D. *ToxA* is present in the U.S. *Bipolaris sorokiniana* population and is a significant virulence factor on wheat harboring *Tsn1*. *Plant Dis*. 2018;102(12):2446-2452. DOI 10.1094/PDIS-03-18-0521-RE.

Gao Y., Faris J.D., Liu Z., Kim Y.M., Syme R.A., Oliver R.P., Xu S.S., Friesen T.L. Identification and characterization of the *SnTox6-Snn6* interaction in the *Parastagonospora nodorum* – wheat pathosystem. *Mol. Plant Microbe Interact*. 2015;28(5):615-625. DOI 10.1094/MPMI-12-14-0396-R.

Hafez M., Gourlie R., Despins T., Turkington T.K., Friesen T.L., Aboukhaddour R. *Parastagonospora nodorum* and related species in Western Canada: genetic variability and effector genes. *Phytopathology*. 2020;110(12):1946-1958. DOI 10.1094/PHYTO-05-20-0207-R.

Kariyawasam G.K., Richards J.K., Wyatt N.A., Running K.L.D., Xu S.S., Liu Z., Borowicz P., Faris J.D., Friesen T.L. The *Parastagonospora nodorum* necrotrophic effector *SnTox5* targets the wheat gene *Snn5* and facilitates entry into the leaf mesophyll. *New Phytol*. 2022;233(1):409-426. DOI 10.1111/nph.17602.

Kolomiets T.M., Pakholkova E.V., Dubovaya L.P. Selection of Starting Material for the Creation of Wheat Cultivars with Long-term Resistance to Septoria. Moscow: Pechatnyy gorod Publ., 2017. (in Russian)

Konkova E.A., Lyashcheva S.V., Sergeeva A.I. Screening of the world winter bread wheat collection for leafstem disease resistance in the Lower Volga Region. *Zernovoe Khozyajstvo Rossii = Grain Economy of Russia*. 2022;14(2):36-40. DOI 10.31367/2079-8725-2022-80-2-36-40. (in Russian)

Kovalenko N.M., Shaydayuk E.L., Gultyaeva E.I. Characterization of commercial common wheat cultivars for resistance to tan spot causative agent. *Biotekhnologiya i Selektsiya Rasteniy = Plant Biotechnology and Breeding*. 2022;5(2):15-24. DOI 10.30901/2658-6266-2022-2-o3. (in Russian)

Liu Z., Faris J.D., Oliver R.P., Tan K.C., Solomon P.S., McDonald M.C., McDonald B.A., Nunez A., Lu S., Rasmussen J.B., Friesen T.L. *SnTox3* acts in effector triggered susceptibility to induce disease on wheat carrying the *Snn3* gene. *PLoS Pathog*. 2009;5(9):e1000581. DOI 10.1371/journal.ppat.1000581.

Liu Z., Zhang Z., Faris J.D., Oliver R.P., Syme R., McDonald M.C., McDonald B.A., Solomon P.S., Lu S., Shelver W.L., Xu S., Friesen T.L. The cysteine rich necrotrophic effector *SnTox1* produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. *PLoS Pathog*. 2012;8(1):e1002467. DOI 10.1371/journal.ppat.1002467.

McDonald M.C., Ahren D., Simpendorfer S., Milgate A., Solomon P.S. The discovery of the virulence gene *ToxA* in the wheat and barley pathogen *Bipolaris sorokiniana*. *Mol. Plant Pathol*. 2017;19:432-439. DOI 10.1111/mpp.12535.

Navathe S., Yadav P.S., Chand R., Mishra V.K., Vasistha N.K., Meher P.K., Joshi A.K., Gupta P.K. *ToxA-Tsn1* interaction for spot blotch susceptibility in Indian wheat: an example of inverse gene-for-gene relationship. *Plant Dis*. 2020;104(1):71-81. DOI 10.1094/PDIS-05-19-1066-RE.

Pakholkova E.V. Septoria disease of grain crops in various regions of the Russian Federation: Cand. Sci. (Biol.). Dissertation. Bol'shiye Vyazemy, 2003. (in Russian)

Pyzhikova G.V., Karaseva E.V. Methods of studying Septoria on isolated wheat leaves. *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology*. 1985;12:112-114. (in Russian)

Pyzhikova G.V., Sanina A.A., Suprun L.M., Kurakhtanova T.I., Gogava T.I., Meparishvili S.U., Antsiferova L.V., Kuznetsov N.S., Ignatov A.N., Kuzmichev A.A. Methods for Assessing the Resistance of Breeding Material and Wheat Cultivars to Septoria. Moscow, 1989. (in Russian)

Richards J.K., Kariyawasam G.K., Seneviratne S., Wyatt N.A., Xu S.S., Liu Z., Faris J.D., Friesen T.L. A triple threat: the *Parastagonospora nodorum* SnTox267 effector exploits three distinct host genetic factors to cause disease in wheat. *New Phytol*. 2022;233(1):427-442. DOI 10.1111/nph.17601.

Sanin S.S., Ibragimov T.Z., Strizhekozin Yu.A. Method for calculating wheat yield losses diseases. *Zashchita i Karantin Rasteniy = Plant Protection and Quarantine*. 2018;1:11-15. (in Russian)

Sharma S. Genetics of Wheat Domestication and *Septoria Nodorum* Blotch Susceptibility in Wheat. Fargo, North Dakota: North Dakota State University, 2019.

Shi G., Zhang Z., Friesen T.L., Bansal U., Cloutier S., Wicker T., Rasmussen J.B., Faris J.D. Marker development, saturation mapping, and high-resolution mapping of the Septoria nodorum blotch susceptibility gene *Snn3-B1* in wheat. *Mol. Genet. Genom*. 2016a;291(1):107-119. DOI 10.1007/s00438-015-1091-x.

Shi G., Zhang Z., Friesen T.L., Raats D., Fahima T., Brueggeman R.S., Lu S., Trick H.N., Liu Z., Chao W., Frenkel Z., Xu S.S., Rasmussen J.B., Faris J.D. The hijacking of a receptor kinase-driven pathway by a wheat fungal pathogen leads to disease. *Sci. Adv*. 2016b;2(10):e1600822. DOI 10.1126/sciadv.1600822.

van Schie C.C., Takken F.L. Susceptibility genes 101: how to be a good host. *Annu. Rev. Phytopathol*. 2014;52:551-581. DOI 10.1146/annurev-phyto-102313-045854.

Virdi S.K., Liu Z., Overlander M.E., Zhang Z., Xu S.S., Friesen T.L., Faris J.D. New insights into the roles of host gene-necrotrophic effector interactions in governing susceptibility of durum wheat to tan spot and Septoria nodorum blotch. *G3 (Bethesda)*. 2016;6(12):4139-4150. DOI 10.1534/g3.116.036525.

Zeleneva Yu.V., Ablova I.B., Sudnikova V.P., Mokhova L.M., Konkova E.A. Species composition of wheat septoria pathogens in the European part of Russia and identifying *SnToxA*, *SnTox1* and *SnTox3* effector genes. *Mikologiya i Fitopatologiya = Mycology and Phytopathology*. 2022;56(6):441-447. DOI 10.31857/S0026364822060113. (in Russian)

Zhang Z., Friesen T.L., Simons K.J., Xu S.S., Faris J.D. Development, identification, and validation of markers for markerassisted selection against the *Stagonospora nodorum* toxin sensitivity genes *Tsn1* and *Snn2* in wheat. *Mol. Breeding*. 2009;23:35-49. DOI 10.1007/s11032-008-9211-5.

Zhang Z., Friesen T.L., Xu S.S., Shi G., Liu Z., Rasmussen J.B., Faris J.D. Two putatively homoeologous wheat genes mediate recognition of *SnTox3* to confer effector-triggered susceptibility to *Stagonospora nodorum*. *Plant J.* 2011;65(1):27-38. DOI 10.1111/j.1365-313X.2010.04407.x.

ORCID ID

Yu.V. Zeleneva orcid.org/0000-0001-9716-288X
E.A. Konkova orcid.org/0000-0001-8607-2301

Acknowledgements. The research was conducted with support from the Russian Science Foundation, project No. 19-76-30005.

Conflict of interest. The authors declare no conflict of interest.

Received April 3, 2023. Revised June 28, 2023. Accepted June 29, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):591-608
DOI 10.18699/VJGB-23-71

Original Russian text https://vavilovj-icg.ru/

Development, results and prospects of the spring durum wheat breeding in Russia (post-Soviet states)

P.N. Malchikov[1, 2] ✉, M.G. Myasnikova[1]

[1] Samara Federal Research Scientific Center of the Russian Academy of Sciences, Samara Scientific Research Agriculture Institute named after N.M. Tulajkov, Bezenchuk, Samara region, Russia
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ sagrs-mal@mail.ru

Abstract. The article outlines a brief historical background on the introduction to cultivation, distribution and breeding of spring durum wheat in the steppe and forest-steppe regions of Eurasia (the countries of the former USSR: Russia, Ukraine, and Kazakhstan). The approaches and methodology for improving durum wheat during certain scientific selection periods are given. The features of the selection program implementation and the breeding scale expansion during the creation of breeding stations at the beginning of the XX century, after the end of the Great Patriotic War, in the second half of the XX century, and at present are considered. A characteristic according to the main features and properties of varieties created in different periods is given. The achievements of the classical breeding method by comparing old and new varieties are analyzed. The efficiency and rate of wheat selection by periods in different regions of Russia is estimated. The results and methods of breeding for yield, resistance to drought, leaf diseases (*Stagonospora nodorum* Berk., *Septoria tritici* (Roeb. et Desm.), *Bipolaris sorokiniana* (Sacc.) Shoemaker, *Pyrenophora tritici repentis* (Died.) Drechs., *Fusarium* sp., *Puccinia titicina* Eriks., *Puccinia graminis* Pers. f. sp. *tritici* Eriks., *Blumeria graminis* (DC.) f. sp. *tritici* Em. Marchal), grain pathogens *Ustilago tritici* (Pers.) Rostr.) and pathogens causing darkening of the corcule and endosperm (*Bipolaris sorokiniana* (Sacc.) Shoemaker, *Alternaria tenuis* (Nees et Fr.), *Alternaria triticina* (Prasada & Prabhu)), pests (*Cephus pygmeus* Lens, *Osinosoma frit* L., *Mayetiola destructor* (Say)), grain quality (protein content, amount of yellow pigments, dough rheology, sprouting resistance) and end products are presented. The prospects for the molecular marker application for a number of traits in breeding in the near future are given.
Key words: durum wheat; Eurasian region; variety; breeding history; breeding rates; crop resistance; pathogen; yield; quality; protein; gluten; pigments; markers.

For citation: Malchikov P.N., Myasnikova M.G. Development, results and prospects of the spring durum wheat breeding in Russia (post-Soviet states). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023; 27(6):591-608. DOI 10.18699/VJGB-23-71

Развитие селекции яровой твердой пшеницы в России (странах бывшего СССР), результаты и перспективы

П.Н. Мальчиков[1, 2] ✉, М.Г. Мясникова[1]

[1] Самарский федеральный исследовательский центр Российской академии наук, Самарский научно-исследовательский институт сельского хозяйства им. Н.М. Тулайкова, п. г. т. Безенчук, Самарская область, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ sagrs-mal@mail.ru

Аннотация. В статье представлен исторический обзор введения в культуру, распространения и селекции яровой твердой пшеницы в степных и лесостепных регионах Евразии (страны бывшего СССР – Россия, Украина, Казахстан). Приведены подходы и методология улучшения твердой пшеницы по периодам научной селекции. Рассматриваются особенности реализации программ и расширения масштабов селекции во время организации селекционных станций в начале XX в., после завершения Великой Отечественной войны, во второй половине XX в. и в настоящее время. Представлена характеристика по основным признакам и свойствам созданных в разные периоды сортов. Анализируются достижения классической селекционной методики путем сравнения старых и новых сортов. Дана оценка эффективности и темпов селекции в разных регионах России. Приведены результаты и методы селекции на урожайность, устойчивость к засухе, болезням листьев (*Stagonospora nodorum* Berk., *Septoria tritici* (Roeb. et Desm.), *Bipolaris sorokiniana* (Sacc.) Shoemaker, *Pyrenophora tritici repentis* (Died.) Drechs., *Fusarium* sp., *Puccinia titicina* Eriks., *Puccinia graminis* Pers. f. sp. *tritici* Eriks., *Blumeria graminis* (DC.) f. sp. *tritici* Em. Marchal.), патогенам зерна *Ustilago tritici* (Pers.) Rostr.) и патогенам, вызывающим потемнение зародыша и эндосперма (*B. sorokiniana* (Sacc.) Shoemaker, *Alternaria tenuis* (Nees et Fr.), *Alternaria triticina* (Prasada & Prabhu)), вредителям (*Cephus pygmeus* Lens, *Osinosoma frit* L., *Mayetiola destructor* (Say)), по качеству зерна (содержание

белка, желтых пигментов, реологические свойства теста, устойчивость к прорастанию) и конечных продуктов. Оцениваются перспективы применения в ближайшем будущем молекулярных маркеров в селекции по ряду признаков.
Ключевые слова: твердая пшеница; Евразийский регион; сорт; история селекции; темпы селекции; устойчивость; патоген; урожайность; качество; белок; клейковина; пигменты; маркеры.

Introduction

Durum wheat is a crop of great importance in the production and consumption of specific food products (pasta, cereals, bread) that have a long shelf life, high nutritional value, and are in demand in almost all regions of the world. The main producers of durum wheat are (in million tons): the European Union – about 9.0 (including Italy – 4.3; France – 1.9; Greece – 1.1; Spain – 1.0); Canada – 5.2; Turkey – 3.7; the USA – 2.3; Kazakhstan, Syria, Algeria – 2.2 each; Morocco – 1.8; Mexico – 1.5; and Tunisia – 1.0 (Eurostat, 2019). In Russia, the domestic consumption of durum wheat grain is about 1.0 million tons per year, 0.2 t of which is imported mainly from Kazakhstan (Groshev, 2019). A significant reduction in planting area (from 2.0 million ha in 1990 to 0.45 million ha in 1994) and in production of durum wheat in Russia took place during the transition from a planned economy to a market economy in the 1990s (Fig. 1).

The durum wheat production decline can be explained by the higher competitiveness of winter crops (winter soft wheat, winter rye) and spring crops (barley, soft spring wheat) in relation to durum wheat which is represented mainly by spring cultivars in Russia. Durum wheat is grown without irrigation in a sharply continental arid climate with an annual rainfall of 250 to 450 mm per year. Biotic stress factors include pests such as aphids, bedbugs, thrips, grass flies, bread sawfly and pathogens such as Helminthosporium and Fusarium leaf spot, Septoria, powdery mildew, stem and in some years brown rust. The effect of pathogen harmfulness in epiphytotic years can be compared with the level of crop losses from severe drought (Vasilchuk, 2001; Rsaliev et, 2020; Tajibayev et al., 2021). The main challenges in spring durum wheat breeding in Russia are associated with the searching and implementation of opportunities to reduce the effect of these limiting environmental factors and improving the quality of grain and end products.

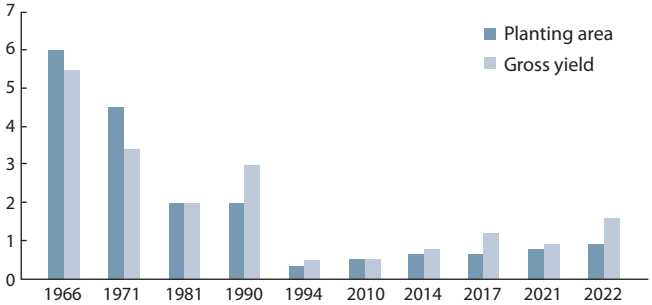

Fig. 1. Dynamics of planting area (million ha) and gross grain yield (million tons) of durum wheat in Russia, 1966–2022.

Prehistory and early history of the durum wheat crop

On the former USSR territory, tetraploid wheat in the form of emmer *T. dicoccum* Schrank ex Schübler was cultivated in the areas of modern regions of Ukraine (Khmelnitsky region) and Azerbaijan in the IV millennium and in the II millennium BC, respectively. The appearance of durum wheat itself (*T. durum* Desf.) was noted in the IX century in Sumy oblast of modern Ukraine (Golik V.S., Golik O.V., 2008). Then, since the XVI–XVIII centuries, it has been widely distributed in the steppe and forest-steppe Russian regions – the central black soil region, the Volga region, the North Caucasus, the Urals, Western Siberia, Ukraine and Western Kazakhstan (Golik V.S., Golik O.V., 2008; Goncharov, 2012). Local cultivars – populations (landraces) under the names of Beloturka, Kubanka, Garnovka, Arnautka, etc. have been cultivated there.

For several centuries, as new territories were being developed, the area of cultivation of landraces increased, regional biotypes were formed simultaneously, which retained (at the time of introduction) the original names (for example, Kubanka), but in the process of long-term reproduction with a change in the dominant variety in the population, acquired regional features. Therefore, in the collection of Russian genetic resources of the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), there are several cultivars Kubanka, Beloturka, etc., with a prefix to the name of the region of origin.

Scientific breeding: development of a network of breeding institutions, general methodology and breeding results from the beginning of the XX century till present

Archival documents of Russian Department of Agriculture (Germantsev, Ilyina, 2019) contain a report on seed production dated 1848, which provides evidence that the peasants of Saratov Trans-Volga region (Novouzensky district) selected the best ears for seeding, i. e. improving mass breeding was carried out. The grain vitreousness and the high protein content in it were the main advantages of Russian durum wheat at that time. At the end of the XIX century, the Italian and French pasta industry worked exclusively on Russian durum wheat (Shekhurdin, 1961).

Scientific approaches to farming and plant improvement have been developed in Russian experimental institutions since the opening of Gorygoretsky Institute of Agriculture with experimental fields in 1840, as well as Riga Experimental Station in 1864, and Petrovsky Agricultural Academy in 1867. By 1917, about 400 experimental institutions including 44 experimental stations were operating in Russia. The study of durum wheat was carried out on experimental stations of the steppe regions of Ukraine, the Volga region, the South Urals, Siberia and Kazakhstan. In 1909, the scientific selection of this crop began at the Krasnokutsk breeding station in Saratov Trans-Volga region that was the centre of production of the most high-quality and demanded durum wheat at that time. Then, selection began at Saratov (1911) and Bezenchuk (1912) experimental stations, also located in the Volga region. At the same time, V.V. Talanov began durum wheat breeding at Yekaterinoslav Experimental Station (Ukraine) and continued it at West Siberian Experimental Station.

One of the main tasks assigned, in particular, to Krasnokutskaya breeding station by the Novouzensky district, was the breeding of more productive cultivars (Germantsev, Ilyina, 2019). There was no problem with the quality of grain grown in the Volga region, since any quantity of it was consumed by the markets of Russia and Europe (Chekhovich, 1924). Breeding work at Krasnokutsk station was led by K.Yu. Chekhovich, who later continued it at Bezenchuk Experimental Station. Under his supervision, the collection and study of durum wheat local samples from Samara, Saratov, Orenburg, Kuban and Ural regions were organized.

The first cultivars in the state variety testing system in 1924 (State Variety Testing Network under the People's Commissariat of Agriculture of the RSFSR) were: Kubanka 5, Gordeiforme 10, Gordeiforme 432, Gordeiforme 189, Melyanopus 69, Gordeiforme 111, Melyanopus 209, Gordeiforme of mass selection by A.I. Nosatovsky, Arnautka of Kochin, Mindum (selection from the Arnautka by Minnesota Experimental Station, the USA), whose originators were eight breeding institutions (Talanov, 1926). In the period from the late 1920s to the mid-1930s, Melyanopus 69, Gordeiforme 189, Gordeiforme 10, Gordeiforme 432, Gordeiforme 27, were approved for use and were most widely used in the USSR (Talanov, 1926). Cultivars Melyanopus 69 and Gordeiforme 189 occupied 2.7 and 0.35 million ha of crops in the USSR, respectively. The third most widespread cultivar was Gordeiforme 10 – 0.25 million ha. Melyanopus 69 and Gordeiforme 189 were created by breeding from local cultivars of the Novouzensky district of Samara province and Ural region, respectively. According to V.V. Talanov (1926), the selection was carried out by K.Yu. Chekhovich in 1911. After many years of research, under the supervision of academician P.N. Konstantinov, they were recommended for farm use.

The morphological and botanical diversity of durum wheat local varieties required an explanation. According to K.Yu. Chekhovich (1924), it was necessary to solve a number of issues in physiology in order for the breeding to have a systematic character. The most successful work was carried out by methods of analytical selection on the source material of local varieties. Despite its effectiveness, collections of cultivars from different ecological and geographical groups were formed in all breeding institutions which were included in intraspecific and interspecific crosses. The breeding efficiency increased after the replenishment of genetic resources due to the expeditions of N.I. Vavilov, his discovery of the centres of origin of cultivated plants, the justification of a systematic approach to the search and study of the source material. However, breeding with the use of hybridization method has been more successful with soft wheat. Hybrid cultivars began to dominate among commercial soft wheat cultivars in the mid-1930s, among durum cultivars – by the early 1970s (Fig. 2).

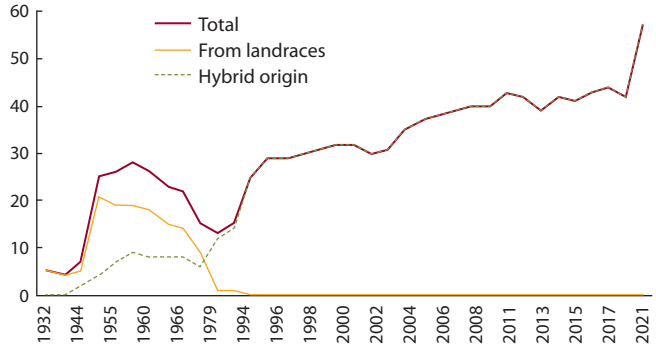

Fig. 2. Dynamics of the number of durum wheat cultivars in the USSR (1932–1991) and in Russia (1992–2021), including those obtained by selection from landraces and hybridization method.

This was because of the desire of government authorities and economic entities to obtain a high result in terms of yield and gross grain harvest that was more successfully achieved when cultivating soft wheat and concentrated the attention of breeders on it. The dominance in the USSR of the theory of T.D. Lysenko also reduced the efficiency of breeding including durum wheat in the 1950s. In the years of war 1941–1945, many breeding institutions in the occupied regions either stopped their activities or were evacuated, while part of the breeding material was lost (Golik V.S., Golik O.V., 2008). Institutions located in the rear, because of lack of funds and personnel, reduced research and breeding work. Nevertheless, in 1954, 25 cultivars were included in the catalog of commercial cultivars of spring durum wheat of the USSR (Catalogs…, 1954–1992). Selections from landraces (21 cultivars) dominated among them. The most widespread, as well as in the 1930s, were the cultivars of the Krasnokutskaya Breeding Station – Melyanopus 1932, Melyanopus 69, Gordeiforme 189, when Melyanopus 1932 was created from crossing landraces, Melyanopus 26 – from crossing Melyanopus 69/Melyanopus 1932.

For 40 years, cultivars of Krasnokutskaya Breeding Station held a monopoly, occupying in some years 86.0 %

of durum wheat crops in the USSR (Germantsev, Ilyina, 2019). Melanopus 69 had been the standard of grain quality on the world market for almost 30 years (Shekhurdin, 1961; Dragavtsev et al., 1994). Cultivars Gordeiforme 432, Gordeiforme 5695 (Institute of the South-East, Saratov), Leukurum 33 (Bezenchuk Experimental Station) were widely distributed.

In Kazakhstan, cultivars of local breeding Akmolinka 5 with high quality and resistance to pathogens that cause black germ (Dorofeev et al., 1987) and Kustanaiskaya 14 were released. In the North Caucasus, the cultivar Krasnodarskaya 362 (Krasnodar Experimental Station) resistant to Swedish fly was introduced, obtained from crossing Gordeiforme 10 with a sample from Algeria. In Ukraine, the cultivar Narodnaya – selection from a local variety at the Kharkov Breeding Station had a significant distribution (0.9 million ha) in the 1950–1960s.

In the mid-1960s, 23 durum wheat cultivars were released in the USSR (Catalogs…, 1954–1992), among which 15 wear the selections from landraces while 8 were the selections from hybrid populations. The main areas of commercial crops were occupied by varieties Melyanopus 1932 and Melyanopus 26 of the Krasnokutskaya Breeding Station. In 1957, Kharkovskaya 46 cultivar was released, created by the Ukrainian Research Institute of Plant Breeding and Genetics from crossing line 34-5129 (interspecific hybrid *T. turgidum*/*T. dicoccum*) presumably either with Algerian durum wheat from the All-Union Research Institute of Plant Breeding collection, or with the line of Kharkov Station. Both versions could not be confirmed because of the loss of documentation during the Great Patriotic War (1941–1945) (Golik V.S., Golik O.V., 2008). That cultivar was distinguished by productivity, responsiveness to favourable environmental conditions, quite resistant to drought, and it was stood out in terms of protein and gluten content. By 1969, it occupied the main crop areas under durum wheat – 4.9 million ha in the USSR (Golik V.S., Golik O.V., 2008).

In the Middle Volga region, cultivars of the Kuibyshev (Bezenchukskaya) Breeding Station were released – Bezenchukskaya 102 (1962) and Bezenchukskaya 105 (1965). In the breeding record of both varieties there was the line Leukurum B-40, obtained by crossing Melyanopus 212 and Gordeiforme 1717 – an interspecific hybrid of landraces of soft and durum wheat. The cultivar Bezenchukskaya 105 (Blagonadezhdina, 1968) had the greatest practical value. It showed good combinational ability, crossings with Kharkovskaya 46 were especially promising, as they gave transgressions in terms of adaptability in the region of the Middle Volga and the Urals (cultivars Orenburgskaya 2, Orenburgskaya 10, Bezenchukskaya 182, Bezenchuksky yantar and their descendants) with great frequency.

In 1962, the cultivar of the Krasnoyarsk Research Institute of Agriculture – Raketa, obtained from interspecific hybridization – Gordeiforme 27*2/Zabaikalskaya emmer wheat was released. That cultivar was a good component for hybridization. It is included in the breeding record of two commercial cultivars with a high content of yellow pigments in the grain – Saratovskaya zolotistaya and Svetlana (Vasilchuk, 2001; Malchikov, Myasnikova, 2020).

In the next decade (1969–1979), the number of released durum wheat varieties in the USSR decreased to 13 (Catalogs…, 1954–1992). Among them there were only two cultivars of non-hybrid origin, one of them was Narodnaya in Ukraine, and the other was local Shavpkha in Georgia. The reduction in the total number of cultivars can be explained by the high level of productivity and competitiveness of Kharkovskaya 46 in all the regions. Cultivars of local importance were: Zernogradskaya 39, Krasnodarskaya 362 and Melyanopus 7 in the North Caucasus; Krasnokutka 6, Melyanopus 26, Leukurum 43, Saratovskaya 40, Saratovskaya 41 in the Volga region; Narodnaya and Nakat in Ukraine.

The catalog of zoned cultivars of the USSR in 1985 included 15 cultivars of durum wheat, 5 of which were from 1979: Almaz (Siberian Research Institute of Agriculture), Altaika (Altai Research Institute of Agriculture and Breeding), Bezenchukskaya 139 (Kuibyshev Research Institute of Agriculture), Orenburgskaya 2 (Orenburg Research Institute of Agriculture), Kharkovskaya 3 (Ukrainian Research Institute of Crop Production, Breeding and Genetics). In the same period, the cultivar Narodnaya, the last one obtained by selection from landraces, lost its commercial significance. Bezenchukskaya 139 was the most popular cultivar; in the years of maximum distribution, it occupied 1.5 million ha. In Siberia, the cultivar Almaz which was distinguished by its high protein content and the quality of pasta, was grown on 0.25–0.35 million ha. The drought-resistant and productive cultivar Orenburgskaya 2 spread in the Ural region (Dolgalev, Tikhonov, 2005). Kharkovskaya 3 was cultivated on irrigation in the Lower Volga region.

In the next decade (1986–1995), 18 cultivars were created that was 62.1 % of the total number of cultivars approved for use in the Russian Federation. Their originators were 10 institutions – 8 from Russia, one each from Ukraine and Kazakhstan (Catalogs…, 1954–1992; State Registers of Selection Achievements…, 1993–2022). The cultivars most widely distributed in terms of the number of tolerance regions were: Bezenchukskaya 182 (Samara Scientific Research Institute of Agriculture) – 5 regions, Krasnokutka 10 (Krasnokutskaya Breeding Station) – 5, Kharkovskaya 23 (Ukrainian Research Institute of Crop Production, Breeding and Genetics) – 4, Svetlana (Research Institute of Agriculture of the Central Chernozem Region) – 4, Orenburgskaya 10 (Orenburg Research Institute of Agriculture) – 4, Saratovskaya zolotistaya (Research Institute of Agriculture of the South-East) – 3, Voronezhskaya 7 (Research Institute of Agriculture of the Central Chernozem Region) – 2, Omsky rubin (Siberian Research Institute of Agriculture) – 2 regions. Cultivars of a wide range include Bezenchukskaya 182, Orenburgskaya 10, Krasnokutka 10, Kharkovskaya 23, Svetlana, Saratovskaya zolotistaya. The first two cultivars are well adapted to the conditions of the Middle Volga region, the Urals and the Central Chernozem region. Krasnokutka 10 with a high accumulation of protein and gluten in the grain is highly drought-resistant in the

Lower Volga region. Kharkovskaya 23 was widely used in the Central Chernozem region and the Urals.

During that period, in the Research Institute of Agriculture of the South-East under the supervision of N.S. Vasilchuk (2001) a program for breeding high-quality cultivars was developed, methods for assessing the yellowness index of semolina, pasta, assessing the rheological properties of dough using a mixograph, farinograph, SDS sedimentation and culinary properties of pasta were introduced and improved. In the same institution, the cultivar Saratovskaya zolotistaya was created which significantly exceeded all previous cultivars in the concentration of yellow pigments in the grain.

From 1996 to 2006, 18 new cultivars of spring durum wheat were approved for use in Russia (State Registers of Selection Achievements…, 1993–2022). These cultivars were created at the Siberian Research Institute of Agriculture – 4, Altai Research Institute of Agriculture – 3, Research Institute of Agriculture of the South-East – 3, Samara Scientific Research Institute of Agriculture – 2, Bashkir Research Institute of Agriculture – 1, North-Donetsk Experimental Station – 1, Krasnodar Research Institute of Agriculture – 1. Two cultivars, Bezenchukskaya stepnaya and Steppe 3, were approved for use in three regions. Despite the reduction in the durum wheat crop area, the Bezenchukskaya stepnaya variety was actively used in the country during this period, occupying 120, 000 ha in some years.

With the development of regional selection, specialization of cultivars in ecological and geographical zones began to appear (Rozova et al., 2017). The conditional border of biotype specialization in Russia was marked on the territory of the Ural region, where cultivars originating from all six large agroecological zones (the North Caucasus, the Central Chernozem Region, the Lower Volga Region, the Middle Volga Region, the Urals, and Siberia) are competitive to one degree or another. The number of cultivars with a high content of yellow pigments in grains and with high gluten quality has increased. During 2007–2016 in all regions of Russia, 22 new cultivars were approved for use, i. e. the breeding rates averaged 2.2 cultivars per year that significantly exceeded all previous periods of durum wheat breeding. In 2016, the State register (State Registers of Selection Achievements…, 1993–2022) increased the total number of cultivars of spring durum wheat to 42.

The process of increasing the number of cultivars used can be seen as a movement towards the creation and diversification of cultivar systems. This is confirmed by both the number of successful breeding institutions (11) and their geography – from Irkutsk to Krasnodar. Along with the traditionally tall cultivars, short-stemmed ones Bezenchukskaya 209 and Rusticano, and medium-sized ones Bezenchukskaya 210, Bezenchukskaya zolotistaya, Lilek, Omskaya yantarnaya were registered. In terms of the duration of the growing season, the differences between early-ripening biotypes (Krasnokutka 13, Nikolasha) and late-ripening ones (Omsky izumrud, Omsky corund) in terms of heading date amounted to 10–12 days when tested in the Middle Volga region.

The number of cultivars with a high content of yellow pigments in the grain (at the level of Saratovskaya zolotistaya) continued to increase. Cultivar Bezenchukskaya zolotistaya exceeded this level by 25.0 %, reaching values of 8.5–9.0 ppm (Malchikov, Myasnikova, 2020). New cultivars were created that consistently form high-quality gluten: Bezenchukskaya 209, Bezenchukskaya niva, Bezenchukskaya zolotistaya, Luch 25, Annushka, Krassar, Lilek, Nikolasha. Breeders of the Altai Research Institute of Agriculture have created a cultivar Solnechnaya 573 which combines two properties that are difficult to combine – high yield in the tolerance regions (Western and Eastern Siberia) and high protein content. In the European part of the country, seven cultivars (Bezenchukskaya 205, Bezenchukskaya 210, Bezenchukskaya niva, Bezenchukskaya zolotistaya, Donskaya elegiya, Marina, Melodiya Dona) have been created which are approved for use in the Ural region. In Siberia, cultivars of only local selection have been released.

Over the past six years (2017–2022), 20 new cultivars have been created (State Registers of Selection Achievements..., 1993–2022) – 3.3 per year, so breeding rates increased compared to the previous period. Among them, cultivars recommended for use in one region prevailed – 18 to 90.0 %, which is a continuation of the trend that was determined at the previous stages – a gradual increase in the share of cultivars of local importance and regional diversification of cultivar systems. The specialization of cultivar systems at this stage is confirmed by a sharp increase in the number of short-stemmed cultivars carrying the *RhtB1b* gene: Triada, SI ATLANT, SI NILO and Tessadur that are recommended for the Central Chernozem region, Bourbon and Nikola that are for the Ural region. These cultivar are resistant to lodging, have a high productivity potential (it is 8.96 t/ha for Triada, 8.98 t/ha – for SI NILO).

Bezenchukskaya krepost, Kremen, and Melyana have middle-sized stems. The first one is included in the register for the Middle Volga and the Ural regions; both of the latter cultivars are only for the Ural region. Bezenchukskaya krepost is resistant to powdery mildew, brown rust, hard smut and accumulates in the grain almost the same amount of yellow pigments as the Bezenchukskaya zolotistaya cultivar. Obtained by LLC “Agroliga – Center for Plant Breeding” from the basic genotype of Samara Scientific Research Institute of Agriculture (1469D-59) by backcrossing to a cultivar from the USA Kofa (a donor of high quality gluten) and selection using molecular markers cultivar Taganrog is resistant to leaf spots, powdery mildew, grain sawfly and distinguished by a high content of yellow pigments and gluten quality. Cultivar Shukshinka is zoned in the Urals, Eastern, and Western Siberia, cultivar Oasis – in Western and Eastern Siberia and cultivar Omsky coral – in Western Siberia.

Cultivars Oasis and Omsky coral are middle-late, tall, but resistant to lodging with a realized grain yield potential of 5.7–6.2 t/ha. The cultivar Shukshinka is middle-late, medium-sized, resistant to lodging, has good and excellent pasta qualities, the maximum yield is 7.28 t/ha. Two

cultivars, Yasenka and Yarina, were approved for use in the North Caucasus region. They are resistant to drought and lodging, loose smut, and are distinguished by large grains and good quality pasta. They are quite competitive in the Volga region and the Urals.

Source material and methods for creating genetic variation

At the beginning of scientific breeding, local cultivars were the main source material. They consisted of mixtures of cultivars and biotypes. Their study was carried out in accordance with the intraspecific classification of cultivated plants of F.K. Körnicke (1873) and J. Percival (1921) based on well-distinguishable features of the ear and grain. Subsequently, N.I. Vavilov (1940) proposed an ecological-geographical principle that combines the classification into cultivars according to F.K. Körnicke (1873) with the division of the entire diversity into ecological-geographical groups. The adequacy of such a division of landraces and ancient hexaploid wheat varieties was confirmed by the results of their clustering based on SSRs and RAPDs markers (Mitrofanova, 2012).

Separation of Eurasian durum wheat cultivars (Russia, Ukraine, Kazakhstan) of the steppe and forest-steppe ecotypes from groups of durum wheat from the Mediterranean and the Middle East is confirmed by the spread of various groups of alleles of gliadin-coding loci among the landraces of these regions (Kudryavtsev et al., 2014). Selection from landraces that was used in all breeding institutions proved to be effective. At the beginning, interspecific crossings of durum wheat (*T. durum* Desf.) with soft wheat (*T. aestivum* L.) and emmer wheat (*T. dicoccum* Schuebl.) prevailed in hybridization, after that – intraspecific crossings with the involvement of initial material from different ecological and geographical groups and from the other countries.

Landraces Beloturka and Sivouska were included in the breeding records of 53 and 41 % of released cultivars, respectively, which indicates the evolutionary nature of breeding with improving genetic systems of adaptability. Derivatives obtained from crossing durum wheat with emmer wheat were of great importance. Variety Kharkovskaya 46 created as a result of interspecific hybridization *T. dicoccum*, *T. turgidum*, and *T. durum* was included in the breeding record of 85 % of the cultivars included in the Russian register in 2004 (Martynov et al., 2005). Cultivar Raketa obtained using the sample *T. dicoccum* from Trans-Baikal territory through cultivars Saratovskaya zolotistaya, Svetlana is currently included in the breeding record of 36.0 % of commercial cultivars in Russia. Sample k-46995 *T. dicoccum* (All-Union Research Institute of Plant Breeding) through the cultivar Pamyati Chekhovicha participated in the origin of seven cultivars included in the register over the past eight years (State Registers of Selection Achievements…, 1993–2022). At all stages of breeding, soft wheat was used for hybridization. At present, 32.7 % of commercial cultivars have soft wheat among their ancestors.

Triticum timopheevii (Zhuk.) cultivar was involved in hybridization as a source of resistance to pathogens. In the register of cultivars protected for 2022, there were three cultivars carrying a translocation on the 6B chromosome from *T. timopheevii*, which provides resistance to powdery mildew (Malchikov et al., 2015). In the 1970s, the commercial variety Melanopus 7 was obtained in Krasnodar with the involvement of *T. timopheevii*.

Since the 1930s, along with interspecific and intraspecific crosses, foreign cultivars have been involved in the creation of hybrid populations. The cultivars WSMP-13 (USA), Leucurum 983 (Italy) were widely included in crosses. WSMP-13 cultivar through the breeding line of the Samara Scientific Research Institute of Agriculture Gordeiforme 740 is included in the breeding record of 12 modern cultivars. The Research Institute of Agriculture of the South-East and the Federal Scientific Centre for Grain named after P.P. Lukyanenko use the gene pool from the international centre ICARDA, the USA and Canada – the cultivars Annushka and Krassar were obtained using the American cultivar Medora, the genes of the American line AWll/Sbl 4 were used for the cultivar Lilek. Omsky corund, Omsky coral were obtained from crossing with genotypes: k-47117, T 1004 = POD 11/Yazi 1 (CIMMYT). When creating short-stemmed varieties Bezenchukskaya 209 and Triada, the donors of the *RhtB1b* gene were Coccorit 71 and Anser 10 (CIMMYT). At present, the cultivar Pamyati Chekhovicha carrying the plant height reduction gene from the cultivar Ahninga (CIMMYT) is widely used for hybridization. In Federal Rostov Agricultural Research Centre, chemical mutagenesis and hybridization with foreign cultivars (Wells, Wascana) are used to induce genetic variability. As a result, cultivars Novodonskaya, Volnodonskaya, Donskaya elegy were created.

The main method of induction of genetic variability in durum wheat in Russia and in the countries of the former USSR is hybridization within and between species. Some scientists assign a crucial role to hybridization and estimate its contribution to the effectiveness of the breeding process up to 60.0 % (Vedder, 1992). The approaches used in the selection of parental components for hybridization correspond to the three principles proposed by S. Boroevich (1984): cultivar, trait, gene.

The principles of cultivar and trait in the domestic literature are usually not separated. In this case, cultivars selected for hybridization are characterized by the entire range of breeding traits (Vasilchuk, 2001; Evdokimov et al., 2022). At the same time, the stage of prebreeding selection, or the purposeful creation of intermediate forms for the stepwise hybridization, is singled out (Shekhurdin, 1961; Vasilchuk, 2001). This principle of parental component selection is widely used in breeding for quantitative traits, resistance to drought and high temperatures. Methods based on the gene principle are used for: backcrossing – transferring genes to a specific gene pool, accumulation – combining genes in one genotype that determine different traits, "pyramidization" – combining two or more genes that determine one trait. The gene principle is used in breeding

Nei's indices of spring durum wheat genetic diversity from Russia and the former USSR for four gliadin-coding loci in a historical context (according to Kudryavtsev, et al., 2014)

Locus	Indices of genetic diversity by periods of variety regionalization			
	1929–1950	1951–1980	1980–2000	2001–2012
Gli-A1		0.64	0.12	0.09
Gli-B1	0.55	0.58	0.69	0.68
Gli-A2	0.58	0.77	0.63	0.58
Gli-B2	0.82	0.80	0.86	0.74
Average	0.68	0.70	0.58	0.52

for resistance to pathogens, culm length and completeness, enzyme activity that are signs closely linked to biochemical or DNA markers.

Genetic diversity of commercial cultivars

The process of increasing uniformity in varietal populations is undesirable, as it increases the likelihood of rapid development of epiphytoties, the spread of pests, and the vulnerability of the varieties to the effects of other extreme environmental factors over a large area (Jacques et al., 2014). Genealogical analysis of the varieties released on the territory of Russia in 1929–2004 based on relatedness coefficients showed an increase in genetic diversity. At the same time, genetic erosion of local material was recorded – the number of Russian original ancestors of modern cultivars decreased by 20 %. In general, during this period, the lower genetic diversity threshold in all regional breeding centres did not reach a critical level corresponding to the similarity of half-sibs (*r*-coefficient of relatedness varied from 0.18 to 0.23 for breeding centres; *r*-coefficient for half-sibs – 0.25) (Martynov et al., 2005). Unambiguous trends in the change in genetic diversity for alleles of gliadin-coding loci during four historical (in time) evolution stages of the varietal population of spring durum wheat were not found (see the Table).

The constancy of the allelic composition at the *Gli-A1*, *Gli-B1*, *Gli-B2* loci was established for the first (1929–1950) and second (1951–1980) stages. In cultivars of the second stage, the number of alleles at the *Gli-A2* locus increased significantly that led to some increase in the coefficient of genetic diversity in general for 4 loci from 0.68 to 0.70. At the next stages, this trend was reversed. A significant narrowing of the diversity occurred at the third (*Gli-A1*, *Gli-A2* loci) and especially the fourth stage, most strongly at the *Gli-A1* locus. A significant decrease in the diversity coefficient at the fourth stage also occurred at the *Gli-A2* and *Gli-B2* loci, which was constant at throughout the XX century. *Gli-B1* which is stable in composition of alleles includes allele *c* with a frequency of 0.46 which contains the electrophoretic component γ-45 that is a marker of genes determining the formation of high-quality gluten.

It is obvious that the current intensive selection of high-quality cultivars can lead to a monoallelic state of the locus and a narrowing of the genetic diversity of the created cultivars. This negative trend can be overcome by attracting a genetically diverse source material, identifying new alleles and intensifying the variety creation in regional breeding centres (Kudryavtsev et al., 2014).

Yield selection results

There is information in the scientific papers about the results of studying the yield trend in breeding institutions in the Volga region (Research Institute of Agriculture of the South-East, Samara Scientific Research Institute of Agriculture) and Western Siberia (Omsk Agricultural Research Centre, Federal Altai Scientific Centre for Agrobiotechnology). At the Samara Research Scientific Institute of Agriculture, the genetic component of yield has been determined quite accurately since the regular testing of Leukurum 33 cultivar in the 1930s. This cultivar of the second stage (Malchikov, Myasnikova, 2015), until the completion of testing the cultivar of the first stage – Gordeiforme 189 in 1958, exceeded it by 23.6 %. If we take the value of 5.0 % as the minimum difference between the yields of Gordeiforme 189 and landraces, then the breeding contribution to the increase in yield in the Middle Volga region when creating the Leukurum 33 cultivar is 28.5 % or 0.8 % per year.

The breeding contribution to the yield value during the creation of the cultivar Bezenchukskaya 105 (1965) and Kharkovskaya 46 (1968) increased by 5 %, but, at the same time, the growth rate due to the breeding decreased to 0.6 % per year from 1912, in the period from 1948 by 1965 – to 0.26 %. The main contribution to the increase in the yield of cultivars at the 2nd–3rd stage of breeding was associated with an improvement in the survival of plants at the time of maturation.

At the next stages – the creation of Bezenchukskaya 139 (stage 4, 1980), Bezenchukskaya 182 (stage 5, 1993), the genetic yield trend related to the variety of the 3rd stage – Kharkovskaya 46 was 10.1 %, related to the variety of the stage 4 – Bezenchukskaya 139 was 15.7 %, with breeding rates of 0.84 and 1.12 % per year, respectively (Malchikov, Myasnikova, 2015).

Most often, the genotypic dispersion of the yield of cultivars of stages 3– 6 was associated with the variability of ear productivity traits and morphophysiological traits, such as Plant height, Harvest index – 65.5 and 31.1 % of

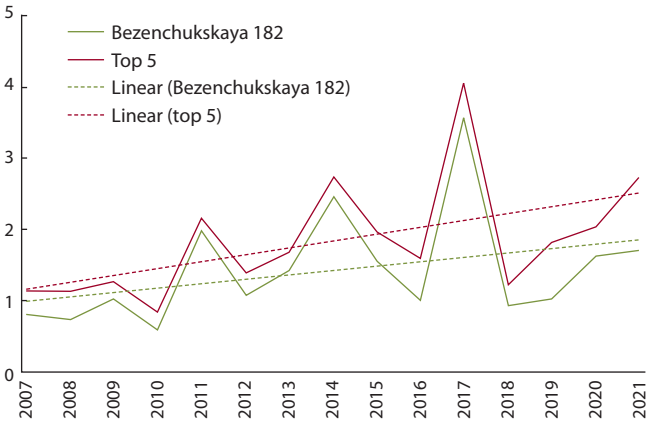

Fig. 3. General and genetic trend (the ratio of the indicators of the five best breeding lines to the indicators of the cultivar of the 5th stage of breeding Bezenchukskaya 182) of durum wheat yield over the past 15 years (2007–2021).

cases, respectively. During the last 15 years (2007–2021), against the background of a general increase in the yield of durum wheat in competitive variety trials, an increase in selection improvement rate has been observed (Fig. 3). The trend in the increase in yield of Bezenchukskaya 182 variety from 1993 to 2021 was 28.5 % or 0.95 % per year.

Thus, over 110 years of durum wheat breeding at the Samara Scientific Research Institute of Agriculture, the genetic yield progress was 0.80 % per year or 87.7 % for the entire period, including the stages of breeding (varietal change): stages 1–2 – 28.5 %, stage 3 – 5 %, stage 4 – 10.0 %, stage 5 – 15.7 %, stage 6 – 28.5 %.

In the Research Institute of Agriculture of the South-East in the XX century (1929–1999), from the moment of the variety Gordeiforme 432 creation (selection from the local Beloturka), the genetic yield trend was 54 % (Vasilchuk, 2001). In the XXI century (2000–2017) it increased by 22 % and amounted to 73.0 % or 0.79 % per year in general for the entire breeding period (Gaponov et al., 2017). Breeding progress in terms of yield was associated primarily with an increase in grain size, the number of productive stems and Harvest index (HI). It was not possible to increase the number of spikelets per spike, the number of grains per spike and spikelet, with the exception of some cultivars. It is predicted that no cardinal changes in plant habitus (reduction in plant height) are expected (Vasilchuk, 2001; Gaponov et al., 2017).

The first cultivar of spring durum wheat created at the West Siberian Station (Omsk Agricultural Research Centre) was Gordeiforme 10 (Talanov, 1926). It was released from 1929 until 1960 and remained the main variety in the region. It was replaced by the Kharkovskaya 46 cultivar which surpassed Gordeiforme 10 by 10 % in the forest-steppe zone (Savitskaya et al., 1980). In 1979, the Almaz variety was released which exceeded Kharkovskaya 46 in yield by 0.23 t/ha or 9.7 % in Siberia and Kazakhstan. Over the next 20 years, Omsk rubin (1991), Angel (1997), and Omskaya yantarnaya (1999) were created. The yield trend for this period (from the Almaz cultivar to Omskaya yantarnaya) amounted to 0.6 t/ha or 22.5 %, i. e. 1.07 % per year based on the results of a long-term study at the Omsk Agricultural Research Centre (Evdokimov et al., 2021).

Cultivars of the last breeding period of durum wheat in the Omsk Agricultural Research Centre (2000–2021) – Omsky izumrud and Omsky coral in terms of grain yield exceeded Omskaya yantarnaya by 18.4 and 19.2 %, i. e. selection rates were about 0.9 % per year (Evdokimov et al., 2020). The general trend of yield improvement breeding if we count from the Gordeiforme 10 variety is 61 % or 0.67 % per year. The cultivar productivity increase in the process of breeding at the Omsk Agricultural Research Centre followed the path of improving the number of productive stems per 1 m^2 and the number of grains per spike. The caryopsis weight changed insignificantly (Evdokimov et al., 2021).

In the Altai Research Institute of Agriculture (Federal Altai Scientific Centre for Agrobiotechnology), the selection of spring durum wheat began in 1929 after the organization of the Barnaul Experimental Station, but systematic and large-scale work has been carried out since 1970. For 50 years, 10 commercial cultivars have been created. The yield trend from Kharkovskaya 46, the main standard in the 1970s, to modern cultivars was 0.39–0.79 t/ha or 14–29 %, the breeding rate for yield was 0.38–0.58 % per year (Rozova et al., 2017). A general trend in the change in yield elements was not found with the exception of the tendency to increase the weight of the grain at all stages. At the same time, most cultivars of all breeding stages formed larger grains. Some varieties surpassed Kharkovskaya 46 in terms of the number of grains per ear and Harvest index.

Thus, the selection of spring durum wheat for yield in the main breeding centres of Russia is carried out quite effectively. The rate of its increase depending on the breeding centre is 0.58–0.80 % per year.

Selection of cultivars resistant to pathogens and pests

At the beginning of breeding work with spring durum wheat in the USSR (Russia), much attention was paid to the creation of cultivars resistant to dust brand (*Ustilago tritici* (Pers.) Rostr.) that is explained by its distribution in the regions of cultivation, increased durum wheat species susceptibility to the pathogen and deterioration in the quality of products. In the 1920–1940s, Gordeiforme 10, Melyanopus 69, Raketa, Gordeiforme 27 were resistant against an infectious background (Shestakova, Vyushkov, 1975). On their basis, highly resistant varieties Bezenchukskaya 121, Bezenchukskaya 139, Svetlana, Valentina, Bezenchukskaya 205, Triada, etc. were obtained. Currently, among the cultivars included in the State Register of Russia there are about 40 % resistant ones which in combination with the use of systemic seed treaters, effectively restrains the development of this pathogen (Vyushkov, 2004).

The most harmful diseases parasitizing the Eurasian durum wheat are leaf spots: *Stagonospora nodorum* (Berk.), *Septoria tritici* (Roeb. et Desm.), *Bipolaris soro-*

kiniana (Sacc.), *Pyrenophora tritici-repentis*, *Fusarium* sp. (Koishibaev, 2018). Breeding for resistance to these pathogens is regional. There is no complete resistance or immunity in durum and soft wheat cultivars to leaf spots (Lamari et al., 1989). In a study by C. Chu et al. (2008), only 25 of 132 durum wheat accessions had high or partial resistance to *P. tritici-repentis* and *S. nodorum* (Berk.). A high level of partial resistance to *P. tritici-repentis* (Tan spot) and *S. nodorum* (Berk.) has been found in synthetic hexaploid wheat (SHW) and soft wheat (Xu et al., 2004; Singh P.K. et al., 2006). Sustainable sources of SHW and wild relatives can potentially be used to improve durum wheat (Singh P.K. et al., 2006).

In the system of the Kazakh-Siberian program (KASIB), the following samples of durum wheat were identified (highly resistant to leaf spot pathogens in the field): 1693d-71, 2021d-1, Gordeiforme 1591-21 (Samara Research Institute of Agriculture), D-2165 (Research Institute of Agriculture of the South-East), Gordeiforme 08-107-5, Gordeiforme 178-05-02, Gordeiforme 05-42-12 (Omsk Agricultural Research Centre) (Gultyaeva et al., 2020; Rsaliyev et al., 2020). A significant part of durum wheat cultivars included in the state register of Russia according to the description of the authors and the results of the study in the state variety network show resistance (R), medium resistance (MR) or medium susceptibility (MS) to these pathogens. Sufficiently effective selection for resistance is confirmed by the low frequency of occurrence of the dominant allele of the susceptibility gene to *P. tritici-repentis – Tsn1*. Among the 43 studied cultivars, it was identified only in two – Soyana and Gordeiforme 08-25-2 (Rsaliyev et al., 2020).

The cultivars of durum wheat from Russia and Kazakhstan are sufficiently resistant to leaf rust (*P. triticina* Eriks.) which confirms the thesis that resistance to this pathogen is higher in durum wheat than in soft wheat (Ordoñez, Kolmer, 2007). Most durum wheat varieties included in the Russian Register have field resistance to leaf rust. During the years of its epiphytoties, cultivars with this type of resistance reduce the yield, the weight of the grain and its fillness, but much less than the susceptible ones. It is advisable to have cultivars with field resistance in the steppe regions with a dry climate where the harmfulness of brown rust is low (Krupnov, 2016). For regions with an increased level of annual precipitation (> 450 mm), it is advisable to use immune cultivars (immunity type 0-1). Donors of resistance genes (immunity with type 0-1) are varieties of the Samara Research Institute of Agriculture Marina and Leukurum 1750. These lines were isolated according to the field assessment in the KASIB seed-plot against a natural infectious background in two ecopoints (Omsk, South Kazakhstan) – Kargala 223, Gordeiforme 178-05-2, Gordeiforme 05-42-12 and Triada.

The results of the multipathogenic test showed the absence of the *Lr2a*, *Lr2b*, *Lr2c*, *Lr9*, *Lr15*, *Lr16*, *Lr17*, *Lr19*, *Lr20*, *Lr24*, *Lr26* genes in the Russian and Kazakh breeding material. The use of molecular markers did not reveal the genes *Lr1*, *Lr3a*, *Lr9*, *Lr10*, *Lr19*, *Lr20*, *Lr24*, *Lr26*, *Lr34*, *Lr37* (Gultayeva et al., 2020). In Spain, they did not identify any known *Lr*-gene in the collection of durum wheat cultivars either (Martínez et al., 2007). Of all known leaf rust resistance genes, only *Lr14*, *Lr23*, *Lr61* and *Lr 79* are derived from durum wheat and spelt (McIntosh et al., 2013). In this regard, it is not clear whether durum wheat has the same resistance genes in the A and B genomes that are identified in soft wheat, or whether these genes are completely different (Gultayeva et al., 2020).

Currently, due to climate change, there is a danger of the spread and intensification of the harmfulness of stem rust (*P. graminis* f. sp. *tritici*) in the steppe regions of the Urals, the Volga region and Siberia (Gultyeva et al., 2020; Evdokimov et al., 2022). An analysis of the periodicity of the pathogen spread shows that durum and soft wheat are very vulnerable to outbreaks of stem rust in the world. The emergence and spread of the Ug99 race and its highly virulent strain TTKSK which overcomes most resistance genes, requires special attention of breeders, geneticists, and phytopathologists (Singh R.P. et al., 2015).

In Russia, cultivars of durum wheat are used that exhibit to some extent resistance to stem rust in the field. According to the State Commission for Testing and Protection of Breeding Achievements, out of 63 cultivars included in the Russian register 14 cultivars showed resistance from medium (MR) to high (R) against a natural infectious background. Highly resistant cultivars include commercial cultivars: Triada, Lilek, Omskaya stepnaya, Omsky coral, Omsky izumrud. The following were moderately stable: Nikolasha, Nikola, Tselinnitsa, Taganrog, Bezenchukskaya 210, Bezenchukskaya krepost, Bezenchukskaya jubileynaya, Bezenchukskaya 205, Tessadur. Sources of resistance are intraspecific variability and introgression of genetic material from other species, primarily *T. dicoccum* Schuebl., *T. timopheevii* Zhuk., *T. aestivum* L. In particular, the lines NT-7, NT-10, and NT-12 that are highly resistant to stem rust, obtained by selection in F3BC1 Shortandinskaya 71/Orenburgskaya 2//*T. timopheevii* k-38555/3/Shortandinskaya 71 are used as initial material (Kozlovskaya et al., 1990). The resistance of the obtained lines is well inherited and controlled by groups of 3–4 genes with the manifestation of complete dominance, incomplete dominance, recessive control of resistance (Khlebova, Barysheva, 2016).

In Western Siberia, on the genetic material of soft wheat from CIMMYT on a natural infectious background during 2018–2019, a high resistance to the stem rust population of the genes *Sr23*, *Sr31*, *Sr38*, *Sr39*, *Sr40* and combinations of the genes *Sr6*, *Sr24*, *Sr36* and *1RS-Am*, *Sr21*, *Sr31* (Evdokimov et al., 2022). Therefore, there are no corresponding virulence alleles in the stem rust population.

Also in the same studies, the high resistance of the cultivars Triada, Omsky coral, Odisseo was determined. The average susceptibility was noted in Omsky izumrud and Luch 25. The vast majority of commercial varieties from Russia and Kazakhstan are very sensitive to race Ug99 which was determined in tests against a natural background in Kenya (Shamanin et al., 2016). At the same time, it was possible to identify genotypes resistant to Ug99 among the

breeding material of which three samples (Gordeiforme 178-05-02, Gordeiforme 05-42-12, and Triada) were resistant or moderately resistant in Kazakhstan and Russia. Stem rust reactions in Kenya and Kazakhstan were similar. In Western Siberia (Omsk, Barnaul, 2017–2018), the degree of damage was higher. Scientists of the Global Rust Reference Centre found that *P. graminis* races from Omsk have unusual virulence patterns compared to Ug99 and races from other regions (Hovmøller et al., 2017).

In the Eurasian regions with a hot climate where spring durum wheat is mainly cultivated, epiphytotics of powdery mildew (*Blumeria graminis*) take place. This can be explained by a short incubation period which at an average daily temperature of 20 to 24 °C ranges from 2.8–3.5 days (Fissyura et al., 1987) that in the presence of an infection ensures its rapid spread. Powdery mildew epiphytoties negatively affect grain quality reducing the content of protein, gluten, weight and grain size (Dolgalev, Tikhonov, 2005). In this regard, resistant cultivars are included in the Register of Russia and Kazakhstan, including steppe regions with a hot climate.

18 spring durum wheat cultivars of 61 ones included in the State Register of Russia show high resistance (R, RMR) to powdery mildew. Analysis of the breeding records of these cultivars shows various sources of resistance, including those based on the genetic variability of *T. durum*, *T. dicoccum*, *T. timopheevii*. In particular, the cultivars Bezenchukskaya krepost and Taganrog which are resistant to powdery mildew carry a translocation from *T. timopheevii* on the 6B chromosome range of microsatellite markers *Xgwm518* and *Xgwm1076* (Malchikov et al., 2015). Lines 1438D-13 (resistance donors – *T. timopheevii*, *T. durum*), 1389DA-1, 1477D-4 (resistance donors – *T. durum*, *T. dicoccum*) with immunity to powdery mildew were obtained at the Samara Research Institute of Agriculture (Malchikov et al., 2015).

A grain with a "black germ" is the main reason for the presence of spex (dark inclusions) in the grains that reduces the colour and nutritional quality of pasta in general (Vasilchuk, 2001). The black germ appears as a result of grain infection with pathogens *B. sorokiniana* (Sacc.) Shoemaker, *Alternaria tenuis* (Fr.), *A. tritici* (Pers.) during the filling period (Conner, 1987).

The assumption that large-grain cultivars are more susceptible was not confirmed – in the process of breeding, resistant cultivars with different grain weights were obtained. Under the conditions of the Altai Territory, the following cultivars are classified as resistant: Salut Altaya, Pamyati Yanchenko, Altaisky yantar, Solnechnaya 573, Angel, Omsky izumrud, 1480d-2, Luch 25, Kharkovskaya 46, Donskaya elegiya, Orenburgskaya 10 (Barysheva et al., 2016). The Samara Research Institute of Agriculture (Malchikov et al., 2022) also identified resistant cultivars: Kharkovskaya 46, Bezenchukskaya 139, Bezenchukskaya 182, Marina, Taganrog, 1963D-71, 2021D-1, Gordeiforme 910, Gordeiforme 08-25-2, Gordeiforme 08-107-5, Melyana. Highly resistant durum wheat genotypes from Italy – ISD19, ISD20, ISD22, Achille, Grecalle, Odisseo and Austria – Duroflaus and Duromax were proposed as initial material.

In Eurasia, significant damage to the yield and quality of durum wheat grain is caused by pests. The leech, bread beetles, thrips, turtle bug have a focal distribution pattern with a rare manifestation of signs of epizootic in some regions. Reducing the harmfulness of these pests is provided by agrotechnical methods. Selection for resistance to Hessian flies, Swedish flies, and the grain sawfly is quite effective (Vyushkov, 2004).

Hessian fly resistance is controlled by a block of dominant H_1–H_{24} genes and several recessive genes (McIntosh et al., 2013). The varietal population of durum wheat in Russia has a sufficient concentration of these genes. The genetics of resistance to the Swedish fly is less studied, but modern cultivars of durum wheat show relative resistance to the pest – damage rarely reaches 11–15 % (Blagonadezhdina, 1968). Kharkovskaya 46 and Bezenchukskaya 139 are evaluated as genetic donors of resistance to the Swedish fly (Vyushkov, 2004).

The control system of breeding material in seed-plots and statistical analysis of varietal variability makes it possible to create slightly damaged cultivars. Resistance to the grain sawfly ensures that the stem is solidness with parenchymal tissue. The high heritability of the trait and the control of the system of dominant genes in interaction with genes inhibitors and anti-inhibitors of stem core formation (Malchikov, Myasnikova, 2008) determine the efficiency of selection of cultivars with a fully or partially completed culm. The absence of negative effects of the genes that control straw solidness on the production process and grain quality allows them to be widely used in breeding for all regions (Malchikov, Myasnikova, 2008). The use of cultivars with completed straw is most effective in the southern regions where the grain sawfly causes maximum damage to the yield and grain quality (Kryuchkov, 2006). Among the cultivars of durum wheat approved for use in 2022, 4 have a fully solidness culm, 36 – medium and 21 cultivars – hollow straw (State Registers of Selection Achievements…, 1993–2022).

Selection for drought tolerance

N.I. Vavilov (1935) attached great importance to the creation of drought-resistant cultivars. The property itself of drought resistance was considered by him as extremely dynamic, depending on time, duration of stress and the period of plant ontogenesis. In his opinion, drought resistance is not a specific trait permanently and invariably inherent in one or another cultivar (Vavilov, 1935). Nevertheless, P.N. Konstantinov (1923) singled out the most significant features which determine drought resistance – the development of the root system and precocity. Subsequently, P.A. Genkel (1982) identified the following factors of drought resistance: 1) resistance of the cytoplasm to dehydration and overheating; 2) rhythm of development; 3) development of the root system; 4) potential productivity.

According to V.A. Kumakov (1985), the weak cytoplasm resistance reduces agronomic drought resistance. Physiological (cytoplasmic, cellular) resistance is considered by breeders as the basis for root growth in soil with low humidity (Kumakov, 1985; Vasilchuk, 2001). Stronger inhibition by drought of cell division in the zones of durum wheat apical meristems than in soft wheat is the main reason for the decrease in its fertility and tillering (Kumakov, 1985).

Selection of spring durum wheat in Russia and the CIS countries for a long period (50–112 years) is carried out in the steppe regions under conditions of significant drought pressure. During this time, tens of thousands of hybrid combinations and millions of breeding lines have gone through cycles of natural and artificial selection. Modern cultivars in drought conditions exceed grain landraces by 2 times in productivity, the first breeding cultivars – by 1.5–1.8 times (Gaponov et al., 2017).

Drought resistance of cultivars is determined by the degree of adaptation to the regional dynamics of growing season meteorological factors. Krasnokutskaya Breeding Station selects the most early-ripening cultivars in the Commonwealth of Independent States (CIS). This is due to the high probability of developing a spring-summer drought with high temperatures in this zone. Early maturing cultivars which form an acceptable yield due to autumn-winter precipitation are an expedient agroecotype here. In this region, early maturation must be combined with a strong root system (Konstantinov, 1923) which due to the reduction in the total growth period is a difficult task for breeders. It is necessary to take into account the decrease in the root system size caused by the negative effect of the dominant *Vrn-A1* gene allele (Smirnova, Pshenichnikova, 2021).

The cultivars bred by the Research Institute of Agriculture of the South-East also mainly belong to the early-ripening biotype. Like the Krasnokut ones, they have a high field drought resistance and are well adapted to the conditions of the Lower Volga region. Cultivars Krasnokutka 13, Nikolasha and Saratovskaya zolotistaya should be considered a significant success in breeding drought-resistant cultivars of these institutions. The first two combine this property with precocity. The third one has a high heat resistance of the ear and has a good cultivar-forming ability in the Middle Volga region and in the Urals (Malchikov, Myasnikova, 2015). Under drought conditions, when studying a set of contrasting cultivars of the Volga region by the method of principal components, a cluster of traits was identified that is closely and positively related to productivity: the number of grains per spike, the nitrogen harvesting index, Harvest index, the growth function of the spike during flowering, the removal of nitrogen and phosphorus during the period from flowering to maturation, the leaf area of the main shoot in tillering (Malchikov, Myasnikova, 2015).

The positive relationship between the removal of macronutrients and, first of all, phosphorus, with the yield during the drought can be interpreted as a result of more vigorous growth of the root system (formation of root hairs) and its activity in drought-resistant varieties (Reynolds et al., 2012). Based on the results of many years of research, drought-resistant cultivars have been identified in Samara Research Institute of Agriculture which are widely used in breeding as initial material: Pamyati Chekhovicha, Bezenchukskaya 205, Bezenchukskaya zolotistaya, Marina, Bezenchukskaya 207, 653d-53, 1368d-18, 2034d-41 (Samara Research Institute of Agriculture), k-16441 (Saada – Morocco), D2017/Karasau//D2043 (Research Institute of Agriculture of the South-East). In the KASIB system, 34 drought-resistant samples were identified out of 154 genotypes during 2000–2015; some of them were later released in arid regions (Evdokimov et al., 2017).

In the Rostov Agricultural Research Centre, successful breeding for drought resistance is carried out using chemical mutagenesis combined with hybridization (Kadushkina et al., 2016). The cultivars of this institution, Donskaya elegiya, Melodiya Dona and Donella M, have high drought resistance in the southern region and in the Volga region. The high drought resistance of cultivars of the National Grain Centre named after P.P. Lukyanenko – Yasenka, Yadritsa and Yarina, and Altai Scientific Centre of Agrobiotechnologies – ATP Prima, ATP Partner, Shukshinka was manifested in the Middle Volga region in the system of ecological testing. The Orenburg Research Institute of Agriculture found that drought-resistant durum wheat cultivars are an effective component of conservation technologies that compensate their negative effects associated in some cases with soil compaction and reduced fertility (Besaliev, Kryuchkov, 2014).

M. Reynolds (2012) with co-authors identified as the main components of the complex property "drought resistance": CTV/CTG – leaf surface temperature at the stages of vegetative growth and grain filling; GC – soil cover during the formation of the crop; ANT – the number of days before flowering; CAR – concentration of carotenoids in leaves; TE – transpiration efficiency based on carbon isotope discrimination; WSC – concentration of sugars in the stem immediately after flowering; HI – Harvest index.

Parameters CTV/CTG, TE are largely determined by the depth of penetration and activity of the root system, xylem diameter and stomatal density. This complex allows you to extract water from the soil, maintain normal transpiration and photosynthesis. The genes *TaMOR*, *TaERs* have been described on soft wheat. The first one is a transcription factor activated by auxin affects the growth and number of roots. The second one (includes *TaER-1*, *TaER-2*) increases the density of stomata and their conductivity, reduces the size of epidermal cells and TE that has a positive effect on photosynthesis and the accumulation of plant biomass in drought conditions.

The study of genes and transcription factors which are orthologous to the structures of other species (arabidopsis, rice, maize) will make it possible to understand the mechanisms of drought and heat resistance in wheat (Kulkarni et al., 2017). V.A. Dragavtsev et al. (2017) suggested using approaches based on the ecological-genetic theory of the organization of quantitative traits in the selection of drought-resistant cultivars. According to their ideas, drought resistance is included in the genetic and physio-

logical system of adaptability and contains 22 component traits. The acceleration of breeding for drought resistance involves the identification of the physiological characteristics of cultivars created in breeding centers located in regions that differ in types of dominant droughts and the dynamics of meteorological factors in plant ontogenesis. Knowledge of specific physiological, epigenetic, biochemical components of drought resistance in the future will make it possible not only to obtain transgressive forms in the process of recombinant breeding, but also to mark the corresponding QTLs for the formation of marker-associated breeding technology. For each agroecological zone, it is necessary to determine the parameters of morphotypes that are complementary to the main types of drought.

Breeding of short-stemmed cultivars

The breeding of short stem varieties of spring durum wheat carried out in the former USSR in the 1970s, was not successful (Vyushkov, 2004; Golik V.S., Golik O.V., 2008). The main reason for the failure was the low adaptability of the short-stemmed donors from Mexico, Chile, Italy, Australia, and Canada. Resistance to lodging in arid zones is expected to be improved by increasing stem strength. (Vasilchuk, 2001). However, the benefit of reducing plant height is not limited to increased resistance to lodging. Low-growing cultivars have a high grain yield from total biomass and productivity potential which is confirmed by the history of wheat breeding in the XX century in many countries.

In Russia, plant height reduction genes are used in the selection of soft and durum winter wheat. Much more difficult is their use in the breeding of spring wheat that is more severely affected by drought than winter wheat. The creation of short-stemmed analogues of Russian cultivars and the gene study on the material of isogenic lines turned out to be promising (Gurkin, 1984; Vyushkov, 2004). By expression at plant height in the Volga region, the studied genes are distributed in the following order: *Rht 14* > *RhtB1b* > *RhtAz* > *RhtAhn.* The *Rht 14* gene was recognized as unpromising because of strong negative effects on adaptability. The *RhtB1b* gene reduced the height of plants by 40 %, increased the grain size of the ear by 17 %, the total tillering by 15 %, Harvest index by 11 %, reduced the weight of 1000 grains by 9 %. The *RhtAz* and *RhtAhn* genes reduced plant height by 17 and 12 %, respectively. The effects of the latter genes on productivity elements were insignificant (Gurkin, 1984; Vyushkov, 2004).

A.A. Alderov (2001) at the Dagestan Experimental Station of All-Russian Institute of Plant Genetic Resources (VIR) introgressed into *T. durum* the genes controlling short stature from the diploid species *T. sinskajae* (*SIS2*) from the hexaploid species *T. aestivum* – Tom Pouce (*Rht 3*) and from *T. dicoccum* – k-25459 (*rhtx 1*, *rhtx 2*). The use of these genes was recommended in the North Caucasus. In the Samara Research Institute of Agriculture, on the basis of an analogue of Kharkovskaya 46 (*RhtAhn* gene), a drought-resistant cultivar Pamyati Chekhovicha was obtained which transferred its properties, including the gene for reducing plant height, to cultivars: Bezenchukskaya 210, Bezenchukskaya zolotistaya, 1368D-18, Bezenchuksky podarok (Samara Research Institute of Agriculture), Shukshinka, ATP Prima (Federal Altai Scientific Centre for Agrobiotechnology), Kremen (Federal Scientific Center for Biological Systems and Agrotechnologies of the RAS). The *RhtB1b* gene was introduced into the commercial varieties Bezenchukskaya 209 and Triada.

In the same period, short-stemmed foreign cultivars Sea Nillo, Sea Atlanta, Tessadour were released. The cultivars Bourbon and Nikola created on genetic material from Italy have a shortness phenotype. The cultivars Bezenchukskaya 209 and Triada have a sufficient level of drought resistance and can be used with intensive tillage technologies in the steppe regions of Russia with an arid climate. Thus, purposeful and long-term work with carriers of plant height reduction genes made it possible to create competitive low-growing/medium-sized varieties and move on to the formation of new morphophysiological types of spring durum wheat in various regions of Russia.

Breeding for grain quality traits

At the end of the XIX century and the beginning of the XX century, the quality of Russian durum wheat on the European market was beyond competition (Chekhovich, 1924). At that time, the main criteria for grain quality were: vitreousness, protein content, completion, grain weight and colour. In 1929 in Saratov (Research Institute of Agriculture of the South-East) A.I. Marushev (1968) began to study in addition to these features in a specialized laboratory: baking properties, strength of flour (gluten), breaking strength macaroni, pasta colour, digestibility, loss of solids when cooking pasta. These parameters formed the basis for evaluating varieties in other breeding centres, in the state commission and formed the basis of GOST for durum wheat quality classification. At present, classiness is determined by vitreousness, grain nature, the presence of grains with a black germ, the quantity and quality of gluten (using the GDI-1-Gluten deformation index device), and the falling number.

In the late 1980s, N.S. Vasilchuk (2001) proposed to evaluate the quality of durum wheat varieties by groups of traits: 1) determined on the grain (glassiness, weight of 1000 grains, nature, colour, ash content, falling number, protein content); 2) determined on meal: the number of spices, colour, content of yellow pigments and activity of oxidizing enzymes; 3) rheological properties of the dough – SDS sedimentation, parameters determined on a mixograph, farinograph, alveograph, glutomatic, glutograph, electrophoresis of gliadin and gutenin blocks, DNA markers); 4) culinary properties of pasta (colour, strength of dry and cooked products, boilability, amount of solids in cooking water).

The most difficult for breeding are signs that negatively correlate with yield, weight of 1000 grains and grain volume weight. They include the protein content in the grain and the amount of gluten. In the process of long-term breeding, these contradictions have become aggravated. Attempts have been made to overcome or significantly

reduce the effects of this negative dependence. The discovery of the wild emmer wheat *T. dicoccoides* (FA-15-3) with large grains and high protein content in Israel made it possible to mark the corresponding locus *QGpc.ndsu-6Bb* on the short arm of the 6B chromosome in the region *Xabg387-6B* and *Xmwg79-6B* where 11 markers are located (Joppa et al., 1997). This made it possible to identify high *Gpc* protein loci in collections of wild species, landraces, and breeding cultivars. The presence of translocation in the group of modern cultivars shows that joint selection for protein content and productivity is possible. In Russia, many cultivars of durum wheat have Kharkovskaya 46 in their breeding record obtained with the participation of the *T. dicoccum* and its sister line Kharkovskaya 51 and other cultivars whose ancestors are *T. dicoccum*, *T. timopheevii*. That suggests a certain probability of their having *Gpc* genes.

The study of 38 durum wheat cultivars created in different ecological and geographical zones (Kharkov, Rostov, Saratov, Bezenchuk, Omsk, Altai) for three years at four points (Bezenchuk, Kurgan, Barnaul, Aktyubinsk) made it possible to determine the following evolutionary grain yield and protein content trends: 1) an increase in the protein content in the grain while maintaining the intensity and adaptability of the production process at the level of the previous breeding stages (cultivar Solnechnaya 573); 2) a significant and stable improvement in yield properties during the breeding process is not accompanied by a decrease in the protein content in the grain (cultivar Bezenchukskaya krepost); 3) a significant and stable improvement in yield properties during the selection process is accompanied by a significant decrease in the protein content in grain (cultivars Bezenchukskaya niva, Bezenchukskaya 210) (Myasnikova et al., 2019).

In addition to protein and gluten content, significant efforts by breeders have been directed towards improving the gluten quality. A.I. Marushev (1968) established a connection between the strength and cooking properties of pasta and the baking properties and strength of flour. The strength of the flour did not always determine the baking qualities of durum wheat, but it was also more closely related to the strength and cooking properties of pasta. The possibility of combining good pasta and baking qualities in durum wheat was also established. Gordeiforme 432, Melyanopus 26, Saratovskaya 34 cultivars were assigned to such ones at the Research Institute of Agriculture of the South-East. Later, V.S. Golik and O.V. Golik (2008) showed the possibility of creating durum wheat cultivars with good baking properties. Canadian breeders came to a similar conclusion where the expediency of breeding durum wheat cultivars of dual use was substantiated.

The discovery of R. Damidaux et al. (1978) two components of γ-gliadin designated γ-42 and γ-45 which were found to be markers of weak and strong gluten, respectively, was important for breeding durum wheat cultivars with high gluten quality. The strong gluten of the γ-45 genotypes (allele) is now known to be functionally mediated by a specific group of low molecular weight (LMW-GS) glutenin subunits designated LMW-2, closely linked to *Gli-B1^{d}c* (γ-45) (Pogna et al., 1988). *GliB1^{d}c* (γ-45)/LMW-2 white spike genotypes were identified early on. Subsequently, A.M. Kudryavtsev (1994) identified two biotypes, γ-45 and γ-42, in the Kharkovskaya 3 cultivar with red ear. The LMW-2 durum wheat genotypes have a wide range of gluten strengths, but they almost always outperform the LMW-1 (γ-gliadin 42) genotypes in baking quality (Kosmolak et al., 1980). There is no evidence that stronger LMW-2 genotypes are better in pasta quality than weaker LMW-1 genotypes (Marchylo et al., 2001). However, strong gluten with its raw content of 28.0–35.0 % has high technological properties in the manufacture of pasta as it gives a dense, viscous dough, well molded, elastic, not wrinkled, not sticky when extruding pasta (Marushev, 1968; Savitskaya et al., 1980).

It has now been established that the quality of durum wheat gluten is determined mainly by five loci, two of them, *Glu-A1* and *Glu-B1*, control the synthesis of high molecular weight glutenins (HMW-GS), three, *Glu-A3*, *Glu-B2*, *Glu-B3*, control the low molecular weight glutenins (LMW-GS). In the *Glu-B1* locus, a positive effect on the quality of gluten (according to the SDS test) was found for alleles *b* (*7+8*), *d* (*6+8*), *z* (*7+15*), *ch* (*7+12*), in the *Glu-B3* for allele *a* (*2+4+15+19*), in the *Glu-A3* locus – for alleles *a* (*6*), *c* (*6+10*), *d* (*6+11*), *e* (*11*). The combination of different alleles at the loci forms more than 40 haplotypes (Ronchallo et al., 2021). Obviously, the study of new varietal collections will make it possible to identify new alleles and determine their effect on the quality of gluten. In Russia, the polymorphism of glyadin-coding loci was studied on the material of collections of historical and modern varieties. In 46.0 % of the cultivars included in the Russian register for 2014, the block *Gli-B1^{d}c* (γ-45)/LMW-2 was found, which implies a high efficiency of breeding high-quality varieties.

Judging by the genealogy of cultivars approved for use in Russia in 2022, it can be assumed that the number of cultivars with high gluten quality (presumably having LMW-2) has increased. Since the end of the 1980s, selection for the quality of gluten has been carried out according to the parameters of GDI (gluten deformation index), SDS sedimentation, mixograph, farinograph. In the Research Institute of Agriculture of the South-East, they improved the assessment of cultivars according to SDS sedimentation (microsedimentation), proposed to expand the 8-point scale of the mixogram at the beginning to 9-point scale (Vasilchuk, 2001) and then to 10-point scale (Gaponov et al., 2020) which was caused by the creation of high-quality cultivars that, under local conditions, form heavy-duty gluten that does not fit into the parameters of the 8-point scale. Currently, the quality of gluten is assessed additionally using glutomatic and glutograf devices (Gaponov et al., 2020). LLC "Agroliga Center for Plant Breeding" and Samara Research Institute of Agriculture using the biochemical marker *Glu-B1* (*7+8*) and molecular markers of the microsatellite group *SSR – Single sequence Repeat* (*Barc148*) *and SNP – Single Nucleotid Polimorphism* (*BM140362*) linked to the genes of high-molecular glutenins

on the 1A chromosome created two cultivars – Taganrog and Alazar (Shevchenko et al., 2019).

In the future, the widespread use of marker-associated breeding technology for these traits based on extensive studies on the phenotyping and genotyping of the gluten quality of collections and commercial durum wheat cultivars adapted to environmental conditions in various agroecological zones of Russia is expected.

Significant progress has been made in breeding for the content of yellow pigments in the grain. This trait is quantitative and under the control of genes with strong additive effects. The corresponding QTLs are distributed over all chromosomes of the durum wheat genome. The trait variation is 60 % determined by two QTLs located on chromosomes 7AL and 7BL (Elouafi et al., 2001; N'Diaye et al., 2017). The first breeding cultivars (created in the 1920–1940s) – Gordeiforme 432, Melyanopus 69, Gordeiforme 189, Gordeiforme 675, Melyanopus 26, accumulate 3.6–5.0 ppm of yellow pigments in the grain. The same level or slightly higher had cultivars of 1960–1980s Kharkovskaya 46, Bezenchukskaya 105, Bezenchukskaya 139. Cultivars Svetlana (1987) and Saratovskaya zolotistaya (1993) accumulated in grain 6–6.5 ppm and 7–7.5 ppm, respectively, that exceeds the level of the first grade of scientific breeding – Gordeiforme 432 by 25–55 %.

Among the cultivars released since 2016, Bezenchukskaya zolotistaya (8.5–9.0 ppm), Bezenchukskaya krepost and Tamara (7.5–8.5 ppm) stand out noticeably. All genotypes of foreign origin were inferior to these cultivars in terms of the trait size when studied in the breeding centres of Russia. The concentration of yellow pigments in the grain of these cultivars exceeds Gordeiforme 432 by 65–85 %. The cultivars Bezenchukskaya zolotistaya, Bezenchukskaya krepost, Bezenchukskaya 210 and Saratovskaya zolotistaya with the optimal combination of the amount of pigments, adaptability, stability and responsiveness of their accumulation in grain were identified. These genotypes are most appropriate to use for the formation of populations and the creation of recombinant inbred lines followed by mapping of the corresponding QTLs and the creation of marker-associated selection technology (Malchikov, Myasnikova, 2020).

Thus, most modern cultivars of spring durum wheat surpass the cultivars of the first breeding stages in terms of the content of yellow pigments in the grain, the rheological properties of the dough and the culinary quality of pasta – primarily the colour and the strength of boiled pasta.

Conclusion

The history of durum wheat growing in Russia and in the former USSR covers many centuries. It is associated with the spread of agriculture in the steppe regions of the Kuban, the Volga region, Siberia and Kazakhstan. At present, about 0.8 million hectares are sown annually in Russia. That is much less than during the period of the planned economy. Scientific breeding of durum wheat has been carried out since 1909. Its yield trend results, depending on the period and region, are 0.26–1.0 % per year, quite comparable with similar results in other countries with a long history of durum wheat breeding. The genetic core of modern varieties was formed on the basis of local varieties of durum wheat, their hybridization with *T. aestivum* L. and *T. dicoccum* Shuebl. The attraction of source material from other countries that has recently increased has had a positive effect.

Despite a significant number of durum wheat breeding laboratories in Russia and a variety of source material, there is a decrease in diversity for alleles of gliadin-coding loci and erosion of original Russian ancestors in modern varieties. Significant advances have been made in breeding for resistance to dust brand (currently 40 % of commercial varieties are resistant) and the most harmful pathogens causing leaf spot in Eurasia (*Stagonospora nodorum* Berk., *Septoria tritici* (Roeb. et Desm.), *Bipolaris sorokiniana* (Sacc.) Shoemaker, *Pyrenophora tritici-repentis*, *Fusarium* sp.). Varieties immune to leaf rust (*Puccinia triticina*) (immunity type 0-1) along with field resistance created in Russia and Kazakhstan are recommended for regions with an increased level of annual precipitation. According to the State Commission for Testing and Protection of Breeding Achievements, 14 varieties out of 63 ones included in the Register of Russia showed resistance to *P. graminis* f. sp. *tritici* from medium to high against a natural infectious background.

Breeding of the cultivars resistant to *B. graminis* (DC.) f. sp. *tritici* Em. Marchal. is being carried out effectively – 30 % of the cultivars show high resistance to this pathogen and pathogens that cause blackening of the corcule and the endosperm. High performance was achieved in breeding for resistance to Hessian, Swedish flies and grain sawfly. The varietal population of durum wheat in Russia has a sufficient concentration of dominant (H_1–H_{24}) and recessive genes that determine resistance to the Hessian fly. Most Russian durum wheat cultivars show relative resistance to Swedish fly – damage rarely reaches 11–15 %.

Selection for drought resistance is regional in nature. It means that drought-resistant biotypes are formed depending on the amount of precipitation, temperature and their dynamics in the regions. Drought-resistant cultivars with the *Rht Ahn* plant height reduction gene (Bezenchukskaya zolotistaya, Bezenchukskaya 210, Bezenchuksky podarok, Shukshinka, ATP Prima) and cultivars Bezenchukskaya 209 and Triada carrying the *RhtB1b* gene adapted to the steppe arid zones were created.

Improvement of grain quality in addition to physical properties is carried out in terms of protein content, amount of yellow pigments and gluten quality. A pair of traits "protein (gluten) content – yield" in the breeding process evolved in the following degree of conjugation: 1) an increase in protein concentration while maintaining yield; 2) decrease in protein concentration with a significant increase in yield; 3) increase in yield while maintaining protein concentration. The concentration of yellow pigments in the grain of modern cultivars exceeds the indicator of the first breeding cultivar Gordeiforme 432 by 65–85 % and reached 8–9 ppm in a number of cultivars. A significant

improvement in the quality of gluten during the selection process occurred in terms of GDI (gluten deformation index), sedimentation (SDS), mixograph, farinograph parameters, and gluten index. According to the results of electrophoresis of the gliadin fraction of storage proteins, the presence of a low molecular weight component of glutenin of the second type (LMW-2) functionally associated with the formation of high-quality gluten was revealed in 50 % of commercial cultivars.

In the short and medium term, classical breeding approaches will continue to play an important role in the durum wheat improvement. Advances in DNA sequencing and other technologies such as bioinformatics, statistics, etc. can help breeders improve the efficiency and speed of the breeding process. Finally, the use of new molecular biology technologies is essential, but their applica-tion must be combined with reliable and extensive field testing.

References

Alderov A.A. Genetics of Short-stem Tetraploid Wheats. St. Petersburg: VIR Publ., 2001. (in Russian)

Barysheva N.V., Rozova M.A., Ziborov A.I., Khlebova L.P., Kraynov A.P. Durum wheat resistance to black point. *Acta Biologica Sibirica*. 2016;2(4):45-51. (in Russian)

Besaliev I.N., Kryuchkov A.G. To the assessment of ecological plasticity spring durum wheat cultivars in hosting bases tillage. *Byulleten' Orenburgskogo Nauchnogo Centra UrO RAN = Bulletin of the Orenburg Scientific Center of the Ural Branch of the Russian Academy of Sciences.* 2014;4:3. (in Russian)

Blagonadezhdina O.A. The results of spring durum wheat breeding under the conditions of dry farming. In: Collection of Research Works of the Kuybyshev (Bezenchuk) Agricultural Station. Kuybyshev, 1968;24-30. (in Russian)

Boroevich S. Principles and Methods of Plant Breeding. Moscow: Kolos Publ., 1984. (in Russian)

Catalogs of Released Varieties in the USSR and Russia. Moscow: Kolos Publ., 1954–1992. (in Russian)

Chekhovich K.Yu. The breeding department of the Bezenchuk Experimental Station. In: Breeding and Seed Production in the USSR. Moscow, 1924;86-96. (in Russian)

Chu C.G., Friesen T.L., Faris J.D., Xu S.S. Evaluation of seedling resistance to tan spot and *Stagonospora nodorum* blotch in tetraploid wheat. *Crop Sci.* 2008;48(3):1107-1116. DOI 10.2135/cropdsci2007. 09.05.16.

Conner R.L. Influence of irrigation timing on black point incidence in soft white spring wheat. *Can. J. Plant Pathol*. 1987;9(4):301-306. DOI 10.1080/07060668709501861.

Damidaux R., Autran J.C., Grinac P., Feillet P. Mise en evidence de relation application en selection entre electrophoregramme des giadines et les proprieties viscoelastiques du gluten du *Triticum durum* Desf. *C. R. Acad. Sci. Series D.* 1978;287:701-704.

Dolgalev M.P., Tikhonov M.P. Adaptive Breeding of Spring Wheat in the Orenburg Cis-Ural Region. Orenburg, 2005. (in Russian)

Dorofeev V.F., Udachin R.A., Semenova L.V., Novikova M.V., Gradchaninova O.D., Shitova I.P., Merezhko A.F., Filatenko A.A. Wheats of the World. Leningrad: Agropromizdat Publ., 1987. (in Russian)

Dragavtsev V.A. (Ed.) Companion of Nikolai Ivanovich Vavilov. Plant Gene Pool Researchers. St. Petersburg: VIR Publ., 1994. (in Russian)

Dragavtsev V.A., Mikhailenko I.M., Proskuryakov M.A. On how we can non-canonically increase hereditary drought resistance in plants (by an example of cereals). *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology*. 2017;52(3):487-500. DOI 10.15389/agrobiology.2017.3.487rus. (in Russian)

Elouafi I., Nachit M.M., Martin L.M. Identification of a microsatellite on chromosome 7B showing a strong linkage with yellow pigment in durum wheat (*Triticum turgidum* L. var. *durum*). *Hereditas.* 2001; 135(2-3):255-261. DOI 10.1111/j.1601-5223.2001.t01-1-00255.x.

Eurostat. Agriculture and Agri-Food Canada. Area, Yield, and Production of Canadian Principal Field Crops Report. 2019. Available online: https://agriculture.canada.ca/en/sector/crops/reports-statistics.

Evdokimov M.G., Yusov V.S., Morgunov A.I., Zelensky Yu.I. Drought tolerance gene pool in developing adaptive varieties of durum wheat identified in study nurseries under the Kazakhstan-Siberian program. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2017;21(5):515-522. DOI 10.18699/VJ17.23-o. (in Russian)

Evdokimov M.G., Belan I.A., Yusov V.S., Kovtunenko A.N., Rosseeva L.P. Adaptive potential of wheat varieties (winter, common and durum spring wheat) selected in the Omsk Agricultural Scientific Centre. *Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology in Agro-Industrial Complex*. 2020;34(10):9-15. DOI 10.24411/0235-2451-2020-11001. (in Russian)

Evdokimov M.G., Yusov V.S., Pakhotina I.V. The main trends in yield and quality of grain of durum spring wheat in the southern forest steppe of Western Siberia. *Vestnik KrasGAU = Bulletin of KSAU*. 2021;169(4):33-41. DOI 10.36718/1819-4036-2021-4-33-41. (in Russian)

Evdokimov M.G., Yusov V.S., Kiryakova M.N., Meshkova L.V., Pakhotina I.V., Glushakov D.A. Promising genetic sources for the creation of varieties of durum spring wheat in Western Siberia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(7):609-621. DOI 10.118699/VJGB-22-75. (in Russian)

Fissyura N.I., Bessemeltsev V.I., Terekhov V.I., Afonin S.P. Dependence of the incubation period of wheat powdery mildew on weather factors. *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology.* 1987;8:50-53. (in Russian)

Gaponov S.N., Shutareva G.I., Popova V.M., Tsetva N.M., Parshikova T.M. Breeding results of spring durum wheat for the arid Volga region. *Agrarnyy Vestnik Yugo-Vostoka = Agrarian Reporter of South East.* 2017;1(16):16-17. (in Russian)

Gaponov S.N., Shutareva G.I., Tsetva N.M., Tsetva I.S., Milovanov I.V. Improvement of the method of rheological assessment of grain quality in the spring wheat breeding. *Zernovoye Khozyaystvo Rossii = Grain Economy of Russia*. 2020;1(67):49-53. DOI 10.31367/2079-8725-2020-67-1-49-53. (in Russian)

Genkel P.A. Physiology of Heat and Drought Tolerance in Plants. Moscow: Nauka Publ., 1982. (in Russian)

Germantsev L.A., Ilyina T.F. Breeding spring durum in Krasnokutskoy Station, 1909–2018. *Agrarnyy Vestnik Yugo-Vostoka = Agrarian Reporter of South-East.* 2019;1:4-9. (in Russian)

Golik V.S., Golik O.V. The Breeding of *Triticum durum* Desf. Kharkov: Magda LTD Publ., 2008. (in Russian)

Goncharov N.P. Comparative Genetics of Wheats and their Related Species. Novosibirsk: Acad. Publ. House "Geo", 2012. (in Russian)

Groshev S.V. Hard wheat is waiting for a firm decision. *Krest'yanin = The Farmer*. 2020;50:10-11. (in Russian)

Gultyaeva E., Yusov V., Rosova M., Malchikov P., Shaydayuk E., Wanyera R., Morgunov A., Yskakova G., Rsaliev A. Evalution of resistance of spring durum wheat germplasm from Russia and Kazakhstan to fungal foliar pathogens. *Cereal Res. Commun.* 2020;48(1): 71-79. DOI 10.1007/s42976-019-00009-9.

Gurkin N.A. The effect of the short stature gene on spring durum wheat. In: Physiological and Genetic Fundamentals of Breeding. Saratov, 1984;137-143. (in Russian)

Hovmøller M.S. GRRC report: Samples of stem rust infected wheat from Russia. 2017. Available online: https://agro.au.dk/fileadmin/Country_report_Russia_-_August2017.pdf.

Joppa L.R., Du C., Hart G.E., Harland G.A. Mapping gene(s) for grain protein in tetraploid wheat (*Trticum turgidum* L.) using a population of recombinant inbred chromosome lines. *Crop Sci.* 1997; 37(5):1586-1589. DOI 10.2135/CROPSCI1997.0011183X0037000 50030X.

Kadushkina V.P., Grabovets A.I., Kovalenko S.A. Results of chemical mutagenesis in breeding of spring durum wheat. *Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology in Agro-industrial Complex*. 2016;2:82-84. (in Russian)

Khlebova L.P., Barysheva N.V. Genetic control of resistance to stem rust in durum wheat introgressive lines derived from *Triticum timopheevii* Zhuk. *Biologicheskiy Vestnik Melitopol'skogo Gosudarstvennogo Pedagogicheskogo Universiteta im. Bogdana Khmel'nitskogo = Biological Bulletin of Bogdan Khmelnitskiy Melitopol State Pedagogical University*. 2016;6(3):121-131. DOI 10.15421/201678. (in Russian)

Koishibaev M. Diseases of Wheat. FAO. Ankara, 2018. (in Russia)

Konstantinov P.N. To Combat Droughts in the Volga Region. Pokrovsk: Unsere Wirtshaft Publ., 1923. (in Russian)

Körnicke F.K. Sysstematische Ubersicht der Cerealien und Monocarpischen Leguminosen. Popelsdorf, 1873.

Kosmolak F.G., Dexter J.E., Matsuo R.R., Lessle D., Marchylo B.A. A relationship between durum wheat quality and gliadin electrophoregrams. *Can. J. Plant Sci*. 1980;60(2):427-432. DOI 10.4141/CJPS80-064.

Kozlovskaya V.F., Grigoryeva L.P., Shatilova N.V. Using interspecific hybridization for the development of new donors of stem rust resistance in wheat. *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology*. 1990;1:65-71. (in Russian)

Krupnov V.A. Strategy and technology of wheat breeding for race-specific leaf rust resistance in the Volga region. *Agrarnyy Vestnik Yugo-Vostoka = Agrarian Reporter of South-East.* 2016;14-15(1-2):22-26. (in Russian)

Kryuchkov A.G. Basic Principles and Methodology of Agroecological Zoning of Grain Crops in the Steppe of the Southern Urals. Moscow: Kolos Publ., 2006. (in Russian)

Kudryavtsev A.M. Genetics of gliadin in spring durum wheat (*Triticum durum* Desf.). *Genetika = Genetics*. 1994;30(1):77-84. (in Russian)

Kudryavtsev A.M., Dedova L.V., Melnik V.A., Shishkina A.A., Upelniek V.P., Novoselskaya-Dragovich A.Y. Genetic diversity of modern Russian durum wheat cultivars at the gliadin-coding loci. *Russ. J. Genet*. 2014;50(5):483-488. DOI 10.1134/S10227954140 50093.

Kulkarni M., Soolanayakanahally R., Ogawa S., Uga Y., Selvaraj M.G., Kagale S. Drought response in wheat: Key genes and regulatory mechanisms controlling root system architecture and transpiration efficiency. *Front. Chem*. 2017;5:106. DOI 10.3389/fchem.2017.00106.

Kumakov V.A. Physiological Grounds of Models of Wheat Varieties. Moscow: Kolos Publ., 1985. (in Russian)

Lamari L., Bernier C.C. Toxin of *Pyrenophora tritici-repentis*: host-specificity, significance in disease, and inheritance of host reaction. *Phytopathology.* 1989;79:740-744. DOI 10.1094/Phyto-79-740.

Malchikov P.N., Myasnikova M.G. Inheritance of straw solidness in *Triticum durum* Desf. *Izvestiya Samarskogo Nauchnogo Centra RAN. Spets. Vypusk: Razvitiye Nauchnogo Naslediya Akademika N.M. Tulaykova = Proceedings of the Samara Scientific Center of the RAS. Special Issue*: *Development of the Scientific Heritage of Academician N.M. Tulaikov*. 2008;133-141. (in Russian)

Malchikov P.N., Myasnikova M.G. Formation of gene associations that code for general homeostasis and performance components of durum wheat (*Triticum durum* Desf.). *Vavilovskii Zhurnal Genetiki i Gelektsii = Vavilov Journal of Genetics and Breeding*. 2015;19(3):323-332. DOI 10.18699/VJ15.042. (in Russian)

Malchikov P.N., Myasnikova M.G. The content of yellow pigments in durum wheat (*Titicum durum* Desf.) grains: biosynthesis, genetic control, marker selection. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2020;24(5):501-511. DOI 10.18699/VJ20.642. (in Russian)

Malchikov P.N., Myasnikova M.G., Leonova I.N., Salina E.A. Introgression of stability to powdery mildew (*Blumeria graminis* DC. f. *tritici*) from *Triticum timofeevii* Zhuk. and *Triticum dicoccum* Shuebl. in genome *Triticum durum* Desf. *Zernovoye Khozyaystvo Rossii = Grain Economy of Russia.* 2015;2:63-67. (in Russian)

Malchikov P.N., Chakheeva T.V., Myasnikova M.G. Source material of spring durum wheat for breeding varieties resistant to pathogens that cause blackening of the grain germ. *Russian Agricultural Sciences*. 2022;48(6):446-453. DOI 10.3103/S106836742206009X.

Marchylo B.A., Dexter J.E., Clarke F.N., Clarke J.M., Preston K.R. Relationships among bread-making quality, gluten strength, physical dough properties, and pasta cooking quality for some Canadian durum wheat genotypes. *Can. J. Plant Sci*. 2001;81(4):611-620. DOI 10.4141/P00-13.

Martínez F., Sillero J.C., Rubiales D. Resistance to leaf rust in cultivars of bread wheat and durum wheat grown in Spain. *Plant Breed*. 2007;126(1):13-18. DOI 10.1111/j.1439-0523.2007.01287.x.

Martynov S.P., Dobrotvorskaya T.V., Pukhalskiy A.V. Analysis of the genetic diversity of spring durum wheat (*Triticum durum* Desf.) cultivars released in Russia in 1929–2004. *Russ. J. Genet.* 2005; 41(10):1113-1122. DOI 10.1007/s11177-005-0208-4.

Marushev A.I. Grain Quality in Volga Wheats. Saratov Publ., 1968. (in Russian)

McIntosh R.A., Yamazaki Y., Dubkovsky J., Rogers W.J., Morris C., Appels R., Xia X.C. Catalogue of gene symbols for wheat. In: Procedings of the 12th Internanional Wheat Genetics Symposium, 8–13 September 2013; Yokahama, Japan. Springer Open, 2013.

Mitrofanova O.P. Wheat genetic resources in Russia: current status and pre-breeding. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2012;16(1):10-20. (in Russian)

Myasnikova M.G., Malchikov P.N., Shabolkina E.N., Sidorenko V.S., Tugareva F.V., Rozova M.A., Chakheeva T.V., Tsygankov V.I. Changes of protein content in spring durum wheat grain during the breeding of high-yield varieties. *Zernobobovyye i Krupyanyye Kul'tury = Legumes and Groat Crops.* 2018;4(32):112-119. DOI 10.24411/2309-348X-2019-11141. (in Russian)

N'Diaye A., Haile J.K., Cory A.T., Clarke F.R., Clarke J.M., Knox R.E., Pozniak C.J. Single marker and haplotype-based association analysis of semolina and pasta colour in elite durum wheat breeding lines using a high-density consensus map. *PloS One.* 2017;12(1): e0170941. DOI 10.1371/journal.pone.0170941.

Ordonez M.E., Kolmer J.A. Virulence phenotypes of a worldwide collection of *Puccinia triticina* from durum wheat. *Phytopatology*. 2007;97:344-351.

Percival J. The Wheat Plant. A Monograph. London, 1921.

Pogna N., Lafiandra D., Feillet P., Autran J.C. Evidence for a direct causal effect of low molecular weight subunits of glutenins on gluten viscoelasticity in durum wheat. *J. Cereal Sci*. 1988;7(3):211-214. DOI 10.1016/S0733-5210(88)80001-8.

Reynolds M.P., Pask A.J.D., Mullan D.M. (Eds.). Physiological Breeding I: Interdisciplinary Approaches to Improve Crop Adaptation. Mexico, D.F.: CIMMYT, 2012.

Ronchallo P.F., Guzman C., Larsen A.O., Achilli A.L., Dreisigsker S., Molfese E., Astiz V., Echenique V. Allelic variation at gluten loci (*Glu-1*, *Glu-2*, and *Glu-3*) in worldwide durum wheat collection and its effect quality atributies. *Foodes.* 2021;10(11):2845. DOI 10.3390/foods10112845.

Rozova M.A., Ziborov A.I., Yegizaryan E.E. Ecologic, genetic and evolutionary aspects of variation of yield and its components in spring durum wheat genotypes under the conditions of the Altai region's the Ob river forest-steppe. *Vestnik Altayskogo Gosudarstvennogo Agrarnogo Universiteta = Bulletin of the Altai State Agrarian University*. 2017;157(11):5-13. (in Russian)

Rsaliyev A.S., Gultyaeva E.I., Malchikov P.N., Shaydayuk E.L., Kovalenko N.M., Yakovleva D.R., Baigutov M.Zh. Resistance of perspective spring durum wheat accessions to foliar diseases. *Vestnik Zashchity Rasteniy = Plant Protection News.* 2020;103(2):105-112. DOI 10.31993/2308-6459-2020-103-2-13334. (in Russian)

Savitskaya V.A., Sinitsyn S.S., Shirokov A.I. Durum Wheat in Siberia. Moscow: Kolos Publ., 1980. (in Russian)

Shamanin V., Salina E., Wanyera R., Zelenskiy Y., Olivera P., Morgounov A.I. Genetic diversity of spring wheat from Kazakhstan and Russia for resistance to stem rust Ug99. *Euphytica.* 2016;212(2): 287-296. DOI 10.1007/s10681-016-1769-0.

Shekhurdin A.P. Breeding and Seed Production of Spring Wheat in the Southeast Region. Moscow: Kolos Publ., 1961. (in Russian)

Shestakova A.P., Vyushkov A.A. Study of spring wheat source material for resistance to loose smut and its breeding use. In: Science and Efficiency of Agricultural Production. Kuybyshev, 1975;89-99. (in Russian)

Shevchenko S.N., Malchikov P.N., Myasnikova M.G., Natoli V., Chakheeva T.V. Accelerated development of high-quality analogues of durum wheat genotypes adapted to the Middle Volga conditions using genomic selection methods. *Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology in Agro-industrial Complex*. 2019;33(12):38-42. DOI 10.24411/0235-2451-2019-11207. (in Russian)

Singh P.K., Mergoum M., Ali S., Adhikari T.B., Elias E.M., Hughes G.R. Identification of new sources of resistance to tan spot, *Stagonospora nodorum* blotch, and *Septoria tritici* blotch of wheat. *Crop Sci.* 2006;46(5):2047-2053. DOI 10.2135/cropsci2005.12.0469.

Singh R.P., Hodson D.P., Jin Y., Lagudah E.S., Ayliffe M.A., Bhavani S., Rouse M.N., Pretorius Z.A., Szabo L.J., Huerta-Espino J., Basnet B.R., Lan C., Hovmøller M.S. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology*. 2015;105(7):872-884. DOI 10.1094/PHYTO-01-15-0030-FI.

Smirnova O.G., Pshenichnikova T.A. The relationship between the genetic status of the *Vrn-1* locus and the size of the root system in bread wheat (*Triticum aestivum* L.). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2021;25(8):805-811. DOI 10.18699/VJ21.093. (in Russian)

State Registers of Selection Achievements Authorized for Use for Production Purposes. Plant Varieties. Ministry of Agriculture of Russia, 1993–2022. (in Russian)

Tajibayev D., Yusov V.S., Chudinov V.A., Malchikov P.N., Rozova M.A., Shamanin V.P., Shtpelev S.S., Sharma R., Tsygankov V.I., Morgunov A.I. Genotype by environment interactions for spring durum wheat in Kazakhstan and Russia. *Ecol. Genet. Genom.* 2021; 21:100099. DOI 10.1016/j.egg. 2021.100099.

Talanov V.V. The Best Varieties of Spring Wheats According to the State Variety Testing Network of the People's Commissariat of Agriculture of the RSFSR in 1924 and 1925. Leningrad, 1926. (in Russian)

Vasil'chuk N.S. Spring Durum Wheat Breeding. Saratov: Novaya Gazeta Publ., 2001. (in Russian)

Vavilov N.I. Scientific Foundations of Wheat Breeding. Moscow; Leningrad: Selkhozgiz Publ., 1935;3-214. (in Russian)

Vavilov N.I. The New Systematics of Cultivated Plants. Oxford: The Clarendon Press, 1940.

Vedder C. Dier Trennung von Kornproteinen mit Hiffe der Elektrophorese als Hilfsmittel bei Elternwahl und Selektion in der Selbstbefruchterzuchtung am Beispil des Winterweizens: Inaugural-Diss. Bonn, 1992.

Vyushkov A.A. Spring Wheat Breeding in the Middle Volga Region. Samara: SamLUX Publ., 2004. (in Russian)

Xu S.S., Friesen T.L., Mujeeb-Kazi A. Seedling resistance to tan spot and *Stagonospora nodorum* blotch in synthetic hexaploid wheats. *Crop Sci*. 2004;44(6):2238-2245. DOI 10.2135/cropsci2004.2238.

ORCID ID

P.N. Malchikov orcid.org/0000-0002-2141-6836
M.G. Myasnikova orcid.org/0000-0002-7224-0308

Acknowledgements. The study was supported by the Russian Science Foundation grant No. 23-16-00041 (https://rscf.ru/project/23-16-00041/).

Conflict of interest. The authors declare no conflict of interest.

Received May 20, 2023. Revised July 9, 2023. Accepted July 11, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):609-622
DOI 10.18699/VJGB-23-72

Estimation of genetic components, heterosis and combining ability of elite Pakistani wheat varieties for yield attributing traits and stripe rust response

M.S. Ahmed[1, 2] ✉, M. Qamar[1], S. Waqar[1], A. Naeem[2], R.A. Javaid[1], S.K. Tanveer[1], I. Hussain[1]

[1] Wheat Program, Crop Sciences Institute, National Agricultural Research Center, Islamabad, Pakistan
[2] Rice Program, Crop Sciences Institute, National Agricultural Research Center, Islamabad, Pakistan
✉ shahzad@parc.gov.pk

Abstract. Wheat (*Triticum aestivum* L.) is a staple food and major source of dietary calories in Pakistan. Improving wheat varieties with higher grain yield and disease resistance is a prime objective. The knowledge of genetic behaviour of germplasm is key. To achieve this objective, elite wheat varieties were crossed in 4 by 3, line × tester design, and tested in 2019 in a triplicate yield trial to estimate genetic variance, general and specific combining ability, mid-parent heterosis and stripe rust (*Puccinia striiformis* L.). High grain 3358 kg·ha^{-1} was recorded in F_1 hybrid (ZRG-79 × PAK-13). Analysis of variance (ANOVA) revealed significant genotypic variance in grain yield. Broad sense heritability (H^2) was recorded in the range of 28 to 100 %. General combining ability (GCA) significant for grain yield in parents except FSD-08 and PS-05 was recorded, while specific combining ability (SCA) was recorded to be highly significant for grain yield only in two crosses (ZRG-79 × NR-09 and ZRG-79 × PAK-13). Mid-parent heterosis was estimated in the range of –28 to 62.6 %. Cross combinations ZRG-79 × PAK-13 depicted highly significant mid-parent heterosis (62.6 %). Highly significant correlation was observed among spike length, spikelets per spike, plant height and 1000-grain weight. Rust resistance index was recorded in the range of 0 to 8.5. These findings suggest exploitation of GCA for higher grain yield is important due to the presence of additive gene action and selection in the filial generations will be effective with improved rust resistance, while cross combinations ZRG-79 × PAK-13 high GCA are best suited for hybrid development.
Key words: wheat; combining ability; mid-parent heterosis; stripe rust; rust resistance index.

For citation: Ahmed M.S., Qamar M., Waqar S., Naeem A., Javaid R.A., Tanveer S.K., Hussain I. Estimation of genetic components, heterosis and combining ability of elite Pakistani wheat varieties for yield attributing traits and stripe rust response. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):609-622. DOI 10.18699/VJGB-23-72

Анализ генетических компонент, гетерозиса и комбинационной способности пакистанских элитных сортов пшеницы по признакам урожайности и устойчивости к желтой ржавчине

М.Ш. Ахмед[1, 2] ✉, М. Камар[1], С. Вакар[1], А. Наим[2], Р.А. Джавайд[1], С.Х. Танвир[1], И. Хусейн[1]

[1] Программа по выращиванию пшеницы, Институт растениеводства, Национальный исследовательский центр сельского хозяйства, Исламабад, Пакистан
[2] Программа по выращиванию риса, Институт растениеводства, Национальный исследовательский центр сельского хозяйства, Исламабад, Пакистан
✉ shahzad@parc.gov.pk

Аннотация. Пшеница (*Triticum aestivum* L.) – основной пищевой и кормовой продукт на территории Пакистана. Одной из приоритетных задач является работа по улучшению сортов пшеницы, отличающихся более высокой урожайностью зерна и устойчивостью к заболеваниям. Ключевым фактором такой работы стало изучение генетического разнообразия сортового материала. С этой целью были оценены показатели урожайности у потомства, полученного от скрещивания четырех элитных сортов пшеницы с тремя тестерными линиями. На основе проведенного анализа получены данные о генетической дисперсии, общей и специфической комбинационной способности, гетерозисе и степени устойчивости к желтой ржавчине (*Puccinia striiformis* L.). Высокая урожайность зерна (3358 кг/га) была отмечена среди растений первого поколения от скрещивания ZRG-79 × PAK-13. Дисперсионный анализ (ANOVA) выявил статистически достоверную генотипическую дисперсию по данному признаку. Значение показателя наследуемости (H^2) фиксировалось в диапазоне от 28 до 100 %. Выявленная общая комбинационная способность (ОКС) по признаку «урожайность зерна» была статистически достоверной для всех родительских сортов, кроме FSD-08 и PS-05, в то время как специфическая комбинационная способность (СКС) по данному признаку оказалась высокодостоверной только для гибридных растений от двух скрещиваний: ZRG-79 × NR-09 и ZRG-79 × PAK-13. Величина гетерозиса составила от –28 до 62.6 %. В комбинациях скрещивания ZRG-79 × PAK-13 была показана высокодостоверная величина гетерозиса (62.6 %). Наблюдалась высокодо-

стоверная корреляция по признакам «длина колоса», «число колосков в колосе», «высота растения» и «масса 1000 зерен». Значение индекса устойчивости к ржавчине изменялось в диапазоне от 0 до 8.5. На основании полученных результатов сделаны следующие выводы: 1) учет ОКС важен при отборе на более высокую урожайность зерна, обусловленную действием аддитивных генов; 2) отбором в дочерних поколениях обеспечится повышенная устойчивость к ржавчине; 3) комбинации скрещивания ZRG-79 × PAK 13 с высокой ОКС лучше всего подойдут для создания гибридных сортов.
Ключевые слова: пшеница; комбинационная способность; гетерозис; желтая ржавчина злаков; индекс устойчивости к ржавчине.

Introduction

Wheat (*Triticum aestivum* L.) is an important cereal crop worldwide playing a crucial role in the daily dietary and nutritional requirement not only for human beings but also for animals. It is the major food for one third of world population and its chief use is the flour for making bread. It is grown around all continents. Increasing human population, climate change and global pandemics have an overwhelming impact on food security, especially wheat on crop with current inadequate genetic improvement of wheat to meet future demand. In Pakistan, wheat is grown in an area of 9.2 million ha with the production of around 25.5 m tonnes (FAOSTAT, 2016) and hardly meets the total requirement of the country. But this figure is continuously under fluctuation because of stagnant yield of cultivars, disease impact, drought, and floods. Apart from these factors, injudicious selection of parental selection for a breeding program without prior knowledge of genetic behaviour in germplasm and lack of indigenous breeding programs for genetic improvement of wheat is another constraint in the yield.

Genetic recombination in germplasm by hybridization is a robust conventional breeding tool for obtaining transgressive segregants and genetic variation, which provides means of selection of ideotypes. Gene action and combining ability analysis are a most reliable biometric procedure for the study of genetic behaviour of yield and yield-related components (Rashid et al., 2007). General combining ability is the average performance of genotypes in a series of cross combinations, while specific combining ability is the performance of a particular genotype in a specific cross combination. Mode of selection depends based on genetic action in traits of interest (Arzu, 2017).

In self-pollinated crops, especially in wheat, plant breeders are usually interested in selection of segregants having additive gene action with high specific combining ability. Additive gene action boosts yield and yield components by cumulative addition of genes. Dominance genetic variance exploits heterosis in cross combinations and specific combining ability provides the presence of dominant or non-additive gene action in a particular trait (Kaushik, 2019) and provides optimal parental identification (Fakthongphan et al., 2016). Equal magnitude of both general and specific combining ability in a breeding population means preponderance of both additive and dominant genes for the traits of interest; selection in this case is most effective for variety development (Ahmad et al., 2012). The term combining ability was first introduced and further refined as general combining ability (GCA) and specific combining ability (SCA) by Sprague and Tatum (1942). GCA distinguishes between the mean performances of parents in cross combinations whiles SCA is the deviation of individual crosses from the average performance of the parents involved. GCA and SCA represent the additive and non-additive portions of genotypic variance respectively (Hallauer et al., 1988). The estimates from GCA and SCA provide an assessment of relative merits of the individual genotypes in cross combinations to guide selection and testing schemes. Thus, line × tester analysis is among the genetic-statistical approaches developed to assist in selection of parents based on their combining ability and the potential to produce promising segregating populations (Okello et al., 2006). According to GCA and SCA impacts, positive values are desirable for most crop plants characteristics, such as growth and yield-related attributes. Negative GCA and SCA impacts, on the other hand, are desirable for characters where minimum values are essential and appealing, such as early flowering.

Heterosis is a phenomenon where F_1 hybrids are superior in traits as compared to their parental genotypes. There are several theories that explain the genetic basis of heterosis, including over-dominance, dominance, and genetic balance. The over-dominance theory of heterosis, first proposed by Shull and East (1908), suggests that heterozygous individuals, since they carry two different alleles, have an advantage over homozygous individuals as they carry two identical alleles for a particular gene. This advantage is thought to mean that the two different alleles can supplement with each other, leading to a vigorous phenotype in F_1 hybrids. The dominance theory, presented by Jones (1917), suggests that hybrid vigour is caused by dominant alleles that are more valuable than the recessive alleles. According to this theory, F_1 hybrids accede two copies of the dominant allele, resulting in a vigorous phenotype. The third heterosis theory is the "Lerner's genetic balance theory", suggested by Lerner (1954), that describes that heterosis is the result of a balance between the expression of genes that promote growth and those that hamper growth. In F_1 hybrids, the expression of growth-promoting genes is increased, whereas the expression of growth-retarding genes is decreased, leading to better growth and development.

Heterotic studies for increasing wheat grain yield have been an interest of early wheat researchers. Mid-parent heterosis is the percent of the increase or decrease in the F_1 value as compared to the average value of both parents for any metric trait. In the early green revolution era Pal and Alam (1938) reported mid-parent heterosis (MPH) in wheat. After the green revolution and introduction of semi-dwarf wheat varieties, various wheat researchers reported MPH heterosis in wheat (i. e., Knott, 1965; Shamsuddin, 1985; Uddin et al., 1992). Barbosa-Neto et al. (1996) reported MPH in soft red winter wheat in the range of –20 to 57 %. Liu et al. (1999), Dreisigacker et al. (2005), Basnet et al. (2019) reported MPH in CIMMYT wheat varieties in the range of 9.5 to 14 %.

Wheat crop faces numerous challenges that cause yield losses, including stripe rust (*Puccinia striiformis* f. sp. *tritici*),

which is a major disease in areas where cool to mild warm temperature prevails during the months of February and March in the wheat-growing season. Under conducive environmental conditions, disease causes yield losses ranging from 10 to 70 % depending upon susceptibility of genotypes (Raza et al., 2018). Development of cultivars containing genetic resistance is the most cost-effective and environmentally friendly strategy to mitigate yield losses by stripe rust (Ali Y. et al., 2020). Stripe rust spores continue to mutate and evolve new virulent races causing damage to previously resistant cultivars (Chen et al., 2010). Wheat crop in Pakistan has faced severe damage caused by stripe rust pathogen in recent years (Ali Y. et al., 2020). Due to climate change and rapid mutation in stripe rust pathogen, new races overwintering on alternative host barberry in hilly areas at high altitudes evolve (Figueroa et al., 2020). Under these circumstances, the already resistant genotypes become susceptible (Javaid et al., 2018).

There are two types of resistance mechanism against rust pathogens in wheat, vertical resistance, and horizontal resistance. Vertical resistance is conferred by a single gene to a specific pathogenic race of rust, while horizontal resistance involves the use of multiple genes that provide broad spectrum disease resistance against multiple pathogenic races of rust. There are several resistance genes present in the Pakistani bread wheat varieties that confer resistance against yellow rust, which include *Yr5*, *Yr10*, *Yr15*, *Yr17*, and *Yr18*. Qamar et al. (2014) reported *Lr34/Yr18* gene complex that confers broad spectrum resistance against yellow rust and leaf rust in most of Pakistani wheat varieties. Intikhab et al. (2021) reported the presence of *Lr46/Yr29* gene complex in Punjab-2011 and Pirsabak-2005 cultivars that confer resistance against stripe rust. Khan S.N. et al. (2022) reported the presence of *Yr17* and *Yr5* gene complex in Pakistani wheat varieties Punjab-2011 and Pirsabak-2005. Utilization of these resistance sources in the breeding program for development of varieties resistant against stripe rust is an ultimate objective to ensure high yield on sustainable basis.

Various biometrical techniques and breeding designs are used for genetic evaluation and genetic behaviour of germplasm to be utilized in crop breeding programs, but line × tester analysis is an efficient mating design providing reliable information about GCA and SCA that ultimately depicts the mode of gene action in a particular trait (Fellahi et al., 2013). GCA and SCA are important to apprehend the genetic architecture of quantitative traits and create the road map for initiation of an efficient breeding program (Fasahat et al., 2016).

Several studies investigating the GCA and SCA effects have been conducted in wheat. Zhao et al. (2013) reported significant effects for both GCA and SCA for yield and its components and inferred that selecting parental genotypes with high GCA and SCA effects could lead to the development of high-yielding wheat hybrids. Similarly, researchers assessed the GCA and SCA effects in spring wheat and durum wheat F_1 hybrids by using line × tester model for combining ability estimate and concluded that GCA effects were more important than SCA effects for grain yield and yield-related traits, and selection of parental genotypes with high GCA effects could increase the prospective yield of wheat hybrids (Iqbal A. et al., 2017; Ishaq et al., 2018; Dragov, 2022). They found that both GCA and SCA effects were significant for grain yield and its components and suggested that selecting parents with high GCA and SCA effects could lead to the development of high-yielding wheat hybrids. Selecting parents with high GCA and SCA effects can improve the yield potential and disease resistance of wheat hybrids, and the use of line × tester designs can provide valuable information about the genetic effects of parents and their hybrids.

The objectives of this study is to elucidate the general and specific combing ability, heterotic potential, and stripe rust (*Puccinia striiformis* f. sp. *tritici*) resistance behaviour of indigenous elite wheat varieties and their breeding population.

Material and methods

Experimental site and plant material. The research was carried out at the experimental site of a wheat research program, National Agricultural Research Center, Islamabad Pakistan (Latitude: 33.71° N, Longitude: 73.06° E, Elevation: 683 m) during 2017–2018 wheat growing season. The soil type of the site is clay loam from 0 to 20 cm, and at the 20–40 cm depth it is moderate clay loam. Five widely adopted approved wheat varieties were used as lines (Faisalabad-2008, Punjab-2011, Pirsabak-2005, Miraj-2008 and Zargoon-79) and three widely adopted, registered and approved varieties for rainfed areas of Pakistan were used as a tester, namely, NARC-2009, Pakistan-2013 and Borlaug-2016 (Table 1). These testers are widely adopted and due to their ability to withstand rainfed and drought-prone areas of Pakistan their leaves have the ability to stay green during high terminal heat and drought stress.

Field experiment and crossing scheme. Eight parents were hybridized to produced 15 F_1 cross combinations according to line × tester crossing fashion as described by Kempthorne (1957) during 2017–2018 wheat growing season and crossing was conducted during March 2018. 15 cross combinations and seven parents were planted in Randomized Complete Block Design (RCBD) with three replications during 2018–2019 wheat growing season. In every replication, parents and F_1 hybrids were sown in 1 m length with row-to-row spacing 25 cm and plant-to-plant spacing 15 cm. The experiment was conducted in an irrigated field and a total of 6 irrigations were applied after sowing to harvesting time. Recommended doses of fertilizers, i. e. 120 kg N · ha^{-1} and 80 kg P · ha^{-1}, were applied. Half of the fertilizers were used at the time of soil preparation, the second half was applied at the time of tillering, and weedicides (Ally Max™ Syngenta and Axial™ Syngenta) were used for eradication of broad leaves and narrow leaves weeds respectively according to the doses mentioned by the manufacturer. Herbicide was applied before the jointing stage of the crop. Leaf area was measured when leaves were fully turgid and green.

Data collection. Grain yield and some yield-related parameters were measured in parents and hybrid combinations. Grain yield per plant was measured in grams and 1000-grain weight was measured after counting 500 grains of each wheat grain sample on a counting tray once and the second sample was repeated for the other 500 grains.

Canopy temperature was measured by using a portable thermal gun (Model: AG-42, Telatemp Crop, CA). Readings for canopy temperature were taken at three Feeks stages (Large, 1954) like booting, kernel water ripening and grain milking stages (Feeks 10, 10.5.4 and 11.2). All readings were

Table 1. Details of wheat parents used in the study and their pedigree/parentage, year of release and parental Institute

Parents		Codes	Pedigree	Institute	Year of release	Area of adoptability
Lines	Faisalabad-2008	FSD-08	PBW-65/2*PASTOR	AARI, Faisalabad, Punjab	2008	Irrigated/rainfed
	Punjab-2011	PB-11	AMSEL/ATTILA//INQUILAB-91/(SIB)PEWEE		2011	Irrigated
	Pirsabak-2005	PS-05	MUNIA/CHORLITO//AMSEL	CCRI, Pirsabak, KPK	2005	Irrigated/rainfed
	Miraj-2008	MRJ-08	SPARROW/INIA//V-7394/WL-711-3/(SIB)BAGULA	AARI, Faisalabad, Punjab	2008	Irrigated
	Zargoon-79	ZRG-79	CORRECAMINOS/INIA-66/3/TOBARI-66/CENTRIFEN//BLUEBIRD/4/SIETE-CERROS-66	WRI, Quetta Baluchistan	1979	Irrigated/rainfed
Testers	NARC-2009	NR-09	INQALAB 91*2/TUKURU	Wheat Program, NARC, Islamabad	2009	Rainfed
	Pakistan-2013	PAK-13	MEX94.27.1.20/3/SOKOLL//ATTILA/3*BCN		2013	Rainfed
	Borlaug-2016	BOR-16	SOKOLL/3/PASTOR//HXL7573/2*BAU		2016	Irrigated/rainfed

taken at the angle of 30° and above 50 cm of the crop canopy, avoiding land temperature by pointing thermal gum only at the canopy. The observations were taken between 11:00 am and 14:00 pm under stagnant air conditions and clear sky as described by (Reynolds et al., 1998). Observations for Normalized difference vegetative index (NDVI) were recorded 50 cm above the canopy by using a hand-held Green Seeker with an optical sensor unit (Model: 505, CA, USA) at three stages of booting and grain filling between 11:00 hours to 14:00 hours with clear sky (Sultana et al., 2014). Values of NDVI range from –1 (NDVI value usually in the water) to +1 (the strongest green vegetative stage) (Kumar, Silva, 1973).

Statistical analysis. Data for other traits (days to 50 % heading, plant height, number of tillers per plant, peduncle length, spike length, days to maturity, number of spikelets per spike, number of grains per spike) were recorded from 6 randomly selected plants. Data recorded were arranged in mean data and subjected to Analysis of Variance (ANOVA) according to Steel and Torrie (1980) and Line × Tester analysis, according to Kempthorne (1957), combining ability and gene action were studied (Singh R.K., Chaudhary, 1977) by using R Package agricolae (De Mendiburu, Simon, 2015; The R Project…, 2017). Genotypic variance and phenotypic variance were estimated as mentioned by Almutairi (2022) in MS Excel 2016, by using the following formula:

$$\text{Genotypic variance } (\sigma^2 g) = \frac{\text{Mean square genotypes (MSG)} - \text{mean square error (MSE)}}{\text{Reps (r)}},$$

$$\text{Phenotypic variance } (\sigma^2 ph) = \sigma^2 g + \sigma^2 e,$$

$$\text{Environmental variance } (\sigma^2 e) = \text{Mean square error (MSE)}.$$

Environmental variance was estimated according to Comstock and Robinson (1952). Broad sense heritability was calculated by using the following formula as described by Burton and Devane (1953):

$$\text{Broad sense heritability } (H^2) = \sigma^2 g / \sigma^2 Ph.$$

Heterosis was estimated in percentage increase or decrease of the F_1 hybrids value over mid-parental value by following the formula as described by Fonseca and Patterson (1968):

$$\text{Mid-Parent Heterosis (\%)} = (F_1 \text{ hybrid} - \text{Mid-parent})/\text{Mid-parent} \times 100.$$

Disease observations and scoring. Observations for stripe rust were recorded at the time of appearance of disease and data were recorded when rust pathogen was fully developed on leaves of a susceptible check cultivar and leaves' surface was fully covered with rust's spores. Disease observation was recorded in three replicates of each parental line and F_1 hybrids according to the Cobb Scale method as described by Peterson et al. (1948). The severity of disease was expressed as the percentage of leaf area covered, and 0 % score was given when there was no infection on the leaf and 100 % score was considered when the leaf area was fully covered with rust spores and infection. Readings of percent severity were recorded with the following descriptions for scoring and response values: (R, resistant = 0.2; S, susceptible = 0.3; MR, moderately resistant = 0.4; MRMS, moderately resistant to moderately susceptible = 0.6; MS, moderately susceptible = 0.8; MSS, moderately susceptible to susceptible = 0.9; S, Susceptible = 1.0), response values, coefficient of infection (CI), average coefficient of infection (ACI), country average relative percentage attack (CARPA) and rust resistance index (RRI) according to Akhtar et al. (2002). The following formula was used for the calculation of RRI:

$$\text{RRI} = \frac{100 - \text{CARPA}}{100} \times 9.$$

RRI was calculated by considering the scale of 0 to 9 from CARPA, where 0 represents a most susceptible genotype and 9 represents a highly resistant response of the genotype to rust pathogen.

Results and discussion

Analysis of variance (ANOVA)

Analysis of variance (ANOVA) results presented in Table 1 show that the lines (female) had statistically significant differences for all the trails. The testers (male) showed statistical differences for days to heading, plant height, peduncle length, spike length, days to maturity, grains per spike, 1000-grain weight and grain yield, while non-significant results for flag leaf area, tillers per plant and spikelets per spike were shown. Interaction of line × tester was significant in case of plant height, flag leaf area, peduncle length, days to maturity, grains per spike and 1000-grain weight. Parents (male and

female) used in this study provided a broad range of expression for various characters as shown in Table 2. There were significant differences ($p \leq 0.05$) among the means of genotypes (Table 3) for days to heading (DH), highly significant ($p \leq 0.01$) for plant height (PH), flag leaf area (FLA), tillers per plant (TPP), peduncle length (PL), spikelets per spike (SPS), days to maturity (DM), grains per spike (GPS), thousand grain weight (TGW), grain yield (GY) and NDVI value.

The values for days to heading (DH) were maximum in the tester (male) PAK-13 (119 days) and minimum in the lines (female) PB-11, PS-05 and MRJ-08 (117 days). DH are the key indicator of earliness in crop production. Plant breeders are keen to create new varieties of wheat genotypes with early maturity. So, early heading is a desirable trait. Delayed heading leads to a reduction in yield (Ullah et al., 2018) and early heading increases the grain filling duration, which ultimately results in high yield (Iqbal A. et al., 2017). Plant height (PH) was the highest in the female parent FSD-08 (103 cm) and the lowest in the male parent NR-09 (83 cm). Minimum PH is preferred due to expected lodging losses. Similarly, the tester NR-09 (83 cm) can be assessed for developing drought tolerant variety with reduced plant height for future breeding programs. Likewise, minimum flag leaf area (FLA) is also desirable for drought tolerance due to reduced transpiration losses from a reduced area exposed to sunlight. The testers (male) PAK-13 and BOR-16 showed the minimum values of flag leaf area: 29.57 and 29.83 cm^2, respectively. Peduncle length was longest in the female parent PS-05 (16.67 cm) and shortest in the male parent NR-09 (7.20 cm). PS-05 produced the maximum grain yield (2531.8 $kg \cdot ha^{-1}$) while minimum grain yield (1696.1 $kg \cdot ha^{-1}$) was recorded in ZRG-79.

Mean performance of parents and their cross combinations

Mean performance for line, testers and cross combinations for days to maturity (DH) ranged from 117 to 119 days. The parental lines FSD-08, PB-11, PS-05, MRJ-08, ZRG-79 and BOR-16 were revealed to have 117 DH, while NR-09 and PAK-13 had 118 and 119 days to heading, respectively (see Table 2). Among F_1 hybrids Zargoon-79 × Pakistan-2013 had 119 days for heading while the rest of the cross combinations showed 117 DH. The grand means for parents, crosses, lines and testers were 117.67, 117.56, 117.27 and 118.33, respectively. The coefficient of variance 0.67 % obtained for DH was also in the acceptable range.

Average minimum plant height was recorded in NARC-2009 (83 cm) followed by cross combination of Punjab-2011 × Pakistan-2013 (86 cm), and maximum plant height of 104 cm was recorded in the cross-combination Faisalabad-2008 × Borlaug-2016 followed by one of parent viz. Faisalabad-2008 (103 cm). Grand mean, coefficient of variance (CV) and least significant variance (LSD) for plant height of lines, testers and their parental combinations was revealed to be 95.92 cm, 3.64 and 5.73, respectively.

The cross-combination Punjab-2011 × Pakistan-2013 showed minimum value (26.8 cm^2) for flag leaf area followed by the lines Pakistan-2013 (29.5 cm^2) and Borlaug-2016 (29.8 cm^2). Maximum leaf area was recorded in the line Punjab-2011 (39.2 cm^2) followed by the F_1 combination, Faisalabad-2008 × Pakistan-2013 (36.9 cm^2). Grand mean, CV, LSD and standard error for leaf area of lines, testers and cross combinations was recorded as 10.46, 5.7 and 2.0 cm^2 respectively.

Maximum 13 tillers per plant (TPP) was recorded in the tester Pakistan-2013 followed by 12.3 tillers in Borlaug-2016 while minimum 7.6 tillers were observed in the female parent Punjab-2011 followed by Pirsabak-2005 (8.33). In the cross combinations a maximum of 10 tillers was recorded in Punjab-2011 × NARC-2009 and Miraj-2008 × NARC-2009 and grand mean for TTP was recorded as 9.49 with CV, LSD and SE 11.66, 3.61 and 1.27 respectively.

Maximum peduncle length was observed in the line Pirsabak-2005 while minimum peduncle length was recorded in the tester parent NARC-2009. Grand mean for peduncle length was observed to be 11.96 cm with CV 9.7 % and LSD (α 0.05) value 1.9.

Mean performance for spike length (SL) was observed 12.36 cm in parents and their cross combinations. Maximum SL was observed in cross combinations Miraj-2008 × Pakistan-2013 followed by Miraj-2008 × Borlaug-2016, but minimum SL was observed in the parental line Faisalabad-2008.

Grand mean for spikelets per spike (SPS) for parents and cross combination was recorded as 20.33 with a maximum of 21.9 spikelets observed in Pakistan-2013 in addition to the F_1 hybrid combination of Miraj-2008 × Pakistan-2013 and in the Pirsabak-2005 × Pakistan-2013. Grains per spike (GPS) was recorded maximum in parental line Punjab-2011. Average thousand grain weight was calculated to be 34.52 with higher TGW in Faisalabad-2008 (43.73 g) and the cross combination of Faisalabad-2008 × Pakistan-2013 (43.53 g), and lower value for TGW was depicted by the parental line NARC-2009 (22.47 g). Average grain yield was obtained in all the parents and cross combinations (2344.5 $kg \cdot ha^{-1}$), while average GY was higher in crosses as compared to the parents' grain yield, the maximum was recorded in the cross combination ZRG-79 × PAK-13 (3358 $kg \cdot ha^{-1}$), followed by PB-11 × NR-09 (2820 $kg \cdot ha^{-1}$), while minimum grain yield was recorded in the cross combination of ZRG-79 × NR-09 (1372 $kg \cdot ha^{-1}$).

Maximum normalized differences in vegetative index (NDVI) value was observed in the parental line PS-05 (0.73) followed by PB-11 × PAK-13 and PS-05 × NR-09 with the same value. Average NDVI value for parents and crosses was revealed to be 0.67; lines, testers and crosses also contained similar values for NDVI.

There were significant differences among the means of crosses combinations for almost all the traits studies except DH, TPP an SL. The lines (female parents) also depicted highly significant differences in all the parameters under consideration except SL. The testers (male parents) also revealed highly significant differences for all the traits except for FLA, SL, SPS, and NDVI. Interaction of lines × testers depicted highly significant differences in their mean performance for the traits of PH, FLA, PL, DM, GPS, TGW and NDVI value.

Estimates of genetic variance components

Estimation of genotypic variance, phenotypic variance, environmental variance, variance due to general combining ability, variance due to specific combining ability and variance due to GCA over SCA is mentioned in Table 4.

Table 2. Mean performance of lines, testers, and their cross combinations under rainfed conditions

Entries	DH	PH	FLA	TPP	PL	SL	SPS	DM	GPS	TGW	GY	NDVI	CT
FSD-08	117.67^{bc}	103^{ab}	32.33^{bcdef}	9.00^{bc}	16.13^{ab}	10.77^{f}	19.20^{de}	156^{abc}	47.90^{abcde}	43.73^{a}	2123.4^{bcde}	0.63^{abc}	25.45^{a}
PB-11	117.00^{c}	99^{abcd}	39.20^{a}	7.67^{c}	13.63^{cde}	12.57^{abcdef}	19.03^{de}	154^{bcdef}	57.47^{ab}	36.87^{cd}	2139.7^{bcde}	0.65^{abc}	24.40^{a}
PS-05	117.00^{c}	98^{bcdef}	33.70^{abcde}	8.33^{c}	16.67^{a}	13.33^{abcd}	21.87^{a}	153^{efgh}	53.33^{abcd}	36.33^{cd}	2531.8^{abcd}	0.73^{a}	24.70^{a}
MRJ-08	117.00^{c}	97^{cdef}	32.27^{bcdef}	10.00^{ab}	14.23^{bcd}	11.57^{def}	18.77^{de}	153^{defgh}	48.00^{abcde}	36.67^{cd}	1916.9^{cde}	0.64^{abc}	25.20^{a}
ZRG-79	117.67^{ab}	93^{fgh}	32.73^{bcde}	9.33^{bc}	12.33^{defg}	13.57^{ab}	21.53^{abc}	155^{abcde}	42.67^{de}	28.60^{hij}	1696.1^{cde}	0.63^{abc}	24.75^{a}
NR-09	118.33^{ab}	83^{j}	32.67^{bcde}	11.00^{abc}	7.20^{j}	11.90^{bcdef}	20.80^{abcd}	155^{abcde}	39.57^{e}	22.47^{k}	2165.8^{bcde}	0.63^{abc}	24.60^{a}
PAK-13	119.00^{c}	91^{ghi}	29.57^{ef}	13.00^{a}	10.67^{ghi}	12.67^{abcde}	21.90^{a}	156^{ab}	53.23^{abcd}	33.60^{def}	2433.4^{abcde}	0.71^{abc}	25.20^{a}
BOR-16	117.67^{ab}	91^{ghi}	29.83^{ef}	12.33^{ab}	9.90^{hi}	13.80^{a}	21.80^{ab}	157^{a}	53.67^{abcd}	27.20^{ij}	2248.6^{abcde}	0.64^{abc}	25.10^{a}
FSD-08 × NR-09	117.00^{c}	94^{defgh}	35.27^{abcde}	8.67^{c}	11.20^{fgh}	12.13^{abcdef}	19.90^{abcde}	153^{defgh}	50.57^{abcde}	32.60^{efg}	2542.9^{abcd}	0.71^{abc}	24.90^{a}
FSD-08 × PAK-13	117.67^{ab}	93^{efgh}	36.90^{ab}	8.33^{c}	12.47^{defg}	12.70^{abcde}	19.67^{bcde}	154^{bcdef}	49.00^{abcde}	36.80^{cd}	3054.9^{ab}	0.64^{abc}	25.50^{a}
FSD-08 × BOR-16	117.00^{c}	104^{a}	30.63^{cdef}	8.67^{c}	15.33^{abc}	11.67^{cdef}	17.90^{e}	154^{bcdef}	46.20^{abcde}	43.53^{a}	2512.6^{abcd}	0.68^{abc}	25.55^{a}
PB-11 × NR-09	117.33^{ab}	89^{hij}	33.37^{bcde}	10.00^{ab}	13.10^{def}	11.67^{cdef}	18.77^{de}	152^{gh}	44.43^{cde}	30.00^{ghi}	2820.3^{abc}	0.64^{abc}	25.25^{a}
PB-11 × PAK-13	118.33^{c}	86^{ij}	26.77^{f}	9.00^{bc}	8.90^{ij}	11.57^{def}	19.67^{bcde}	155^{abcde}	50.57^{abcde}	25.27^{jk}	2572.0^{abcd}	0.73^{ab}	25.10^{a}
PB-11 × BOR-16	117.00^{c}	98^{bcdef}	33.53^{abcde}	8.33^{c}	11.47^{fgh}	12.80^{abcde}	20.80^{abcd}	151^{h}	50.33^{abcde}	37.87^{bc}	1595.6^{de}	0.66^{abc}	25.60^{a}
PS-05 × NR-09	117.00^{c}	99^{abcd}	30.50^{cdef}	8.00^{c}	12.77^{def}	11.67^{cdef}	19.90^{abcde}	154^{cdefg}	55.43^{abc}	38.60^{bc}	2430.4^{abcde}	0.73^{a}	25.85^{a}
PS-05 × PAK-13	117.00^{c}	99^{abcde}	29.93^{def}	9.00^{bc}	11.47^{fgh}	13.43^{abc}	21.90^{a}	155^{abcd}	47.10^{abcde}	35.07^{cdef}	2735.3^{abc}	0.64^{abc}	25.55^{a}
PS-05 × BOR-16	117.00^{c}	100^{abcd}	33.17^{bcde}	8.67^{c}	12.20^{efg}	12.23^{abcdef}	20.33^{abcd}	152^{fgh}	52.33^{abcd}	37.67^{bc}	2355.7^{abcde}	0.67^{abc}	25.55^{a}
MRJ-08 × NR-09	117.67^{c}	99^{abcdef}	33.00^{bcde}	10.33^{abc}	12.97^{def}	11.47^{ef}	19.47^{cde}	152^{gh}	53.43^{abcd}	37.53^{bc}	2424.0^{abcde}	0.66^{abc}	25.25^{a}
MRJ-08 × PAK-13	118.33^{c}	96^{defg}	33.87^{abcde}	9.67^{abc}	9.67^{hi}	13.90^{a}	21.90^{a}	154^{bcdef}	56.43^{abc}	31.93^{fgh}	2559.8^{abcd}	0.67^{abc}	24.80^{a}
MRJ-08 × BOR-16	117.67^{c}	96^{defg}	36.77^{ab}	10.33^{abc}	10.10^{hi}	12.57^{abcdef}	21.47^{abc}	156^{abc}	54.13^{abcd}	35.47^{cdef}	1700.3^{cde}	0.60^{c}	25.70^{a}
ZRG-79 × NR-09	117.67^{c}	100^{abcd}	35.63^{abcd}	8.67^{c}	10.57^{ghi}	11.90^{bcdef}	20.57^{abcd}	155^{abcde}	58.13^{a}	35.53^{cde}	1372.9^{e}	0.62^{bc}	25.90^{a}
ZRG-79 × PAK-13	119.00^{c}	96^{defg}	35.87^{abc}	8.00^{c}	9.77^{hi}	11.57^{def}	20.77^{abcd}	156^{abc}	47.57^{abcde}	29.93^{ghi}	3358.4^{a}	0.63^{abc}	24.50^{a}
ZRG-79 × BOR-16	117.67^{c}	102^{abc}	34.50^{abcde}	9.00^{bc}	12.23^{efg}	12.90^{abcde}	19.67^{bcde}	154^{bcdef}	45.90^{bcde}	40.80^{ab}	2631.4^{abcd}	0.68^{abc}	23.70^{a}
Grand Mean	117.59	95.92	33.17	9.49	11.96	12.36	20.33	154.25	51.01	34.52	2344.5	0.67	25.13
Mean of parents	117.67	94.54	32.81	10.17	12.60	12.52	20.61	154.92	49.60	33.18	2156.96	0.66	25.12
Mean of crosses	117.56	96.82	33.36	9.13	11.61	12.28	20.18	153.89	51.77	35.24	2431.07	0.67	25.14
Mean of lines	117.27	98.13	34.11	9.00	14.60	12.36	20.08	154.27	49.87	36.44	2081.58	0.6553	25.37
Mean of testers	118.33	88.55	30.64	12.11	9.26	12.79	21.50	156.00	48.82	27.75	2282.60	0.6800	24.70
CV (%)	0.67	3.64	10.46	11.66	9.71	8.97	6.56	0.64	3.84	6.31	23	2.13	1.00
LSD (α 0.05)	1.28	5.73	5.70	3.61	1.90	1.82	2.19	2.17	12.04	3.58	564.35	0.11	
SE	0.45	2.01	2.00	1.27	0.67	0.64	0.77	0.57	1.13	1.26	399.06	0.0082	–1.0
SED	0.64	2.19	1.64	0.90	0.95	0.91	1.09	0.80	1.60	1.78	0.85	0.0116	–1.0

Note. Hereinafter: DH, days to heading; PH, plant height, cm; FLA, flag leaf area, cm^2; TPP, tillers per plant; PL, peduncle length, cm; SL, spike length, cm; SPS, spikelets per spike; DM, days to maturity; GPS, grains per spike; TGW, 1000-grain weight; GY, grain yield per plant, $kg \cdot ha^{-1}$; NDVI, normalized difference in vegetative index, and CT, canopy temperature, °C.

Table 3. ANOVA for line × tester (including parents) mean squares for morpho-physiological and agronomic traits under rainfed conditions

SOV	DF	DH	PH	FLA	TPP	PL	SL	SPS	DM	GPS	TGW	GY	NDVI	CT
Reps	2	1.49^{ns}	15.50**	2.04^{ns}	0.71**	1.19^{ns}	6.33**	1.56	1.62^{ns}	2.95^{ns}	54.27**	453264.39^{ns}	0.0003^{ns}	–0.000^{ns}
Genotypes	22	1.21*	84.36**	24.66**	5.00**	15.94**	2.17^{ns}	4.30**	6.04**	88.94**	88.01**	664369.03*	0.0037**	0.8471^{ns}
Crosses	14	1.13^{ns}	68.04**	24.12**	1.42^{ns}	8.32**	1.68^{ns}	3.67*	5.70**	79.98**	65.28**	468535.90*	0.0034**	0.5088**
Lines (c)	4	1.89*	109.59**	35.73**	3.41*	7.83**	0.61^{ns}	4.79*	5.61**	50.64**	60.72**	907796.47**	0.0024**	0.9032^{ns}
Testers (c)	2	2.96*	131.49**	5.55^{ns}	0.80^{ns}	15.22**	3.09^{ns}	4.45^{ns}	9.76*	121.92**	199.70**	915502.50*	0.0005^{ns}	0.2435**
L × T (c)	8	0.29^{ns}	31.41**	22.96**	0.58^{ns}	6.84**	1.87^{ns}	2.91^{ns}	4.73**	84.16**	33.95**	135237.50^{ns}	0.0046**	0.3779^{ns}
Parents	7	1.52*	117.42**	28.60**	10.48**	31.29**	3.34*	5.75**	5.21**	109.12**	136.61**	1138947.60*	0.0050**	1.6434**
Lines (p)	4	0.40^{ns}	41.77**	28.52**	3.17*	9.62**	4.21*	6.67*	3.73**	92.32**	86.18**	907796.47**	0.0054**	0.3848**
Testers (p)	2	1.33^{ns}	69.44**	9.21^{ns}	3.11^{ns}	9.95**	2.74^{ns}	1.11^{ns}	2.33^{ns}	192.89**	93.66**	915502.50**	0.0049**	3.7200^{ns}
L *vs.* T	1	6.40**	516.00**	67.68**	54.44**	160.67**	1.03^{ns}	11.34*	16.90**	8.80^{ns}	424.24**	135237.47^{ns}	0.0034**	2.52^{ns}
C *vs.* P	1	0.19^{ns}	81.41**	4.69^{ns}	16.71**	15.11**	0.92^{ns}	2.96^{ns}	16.53**	73.20**	66.21**	83982.72^{ns}	0.0000^{ns}	0.0095^{ns}
Error	44	0.61	7.20	4.02	1.23	1.35	1.23	1.78	0.97	3.83	4.74	477623.54	0.0002	0.00

Note. SOV, source of variation; DF, degree of freedom. * Significant at $p < 0.05$, ** significant at $p < 0.01$, ^{ns} significant at $p > 0.05$.

Table 4. Estimates of genetic components of variance and degree of dominance of yield attributing traits under rainfed conditions

SOV	DH	PH	FLA	TPP	PL	SL	SPS	DM	GPS	TGW	GY	NDVI	CT
σ^2g	0.20	25.72	6.88	1.26	4.86	0.31	0.84	1.69	28.37	27.76	62,248	0.00	0.28
σ^2p	0.40	28.12	8.22	1.67	5.31	0.72	1.43	2.01	29.65	29.34	221,456	0.00	0.28
σ^2e	0.20	2.40	1.34	0.41	0.45	0.41	0.59	0.32	1.28	1.58	159,207	0.00	0.00
H^2, %	49.59	91.47	83.70	75.40	91.53	43.32	58.60	83.94	95.69	94.61	28.11	94.59	100
σ^2gca	0.03	1.29	0.04	0.03	0.05	–0.01	0.03	0.03	–0.15	1.11	11,783.28	0.00	0.00
σ^2sca	–0.12	9.14	6.13	–0.07	1.99	0.22	0.70	1.32	26.81	10.60	58,407.31	0.00	0.13
σ^2sca/σ^2gca	–0.25	0.14	0.01	–0.43	0.03	–0.05	0.04	0.02	–0.01	0.10	0.20	0.00	0.00

Note. σ^2g, genotypic variance; σ^2p, phenotypic variance; σ^2e, environmental variance; H^2, broad sense heritability; σ^2gca, general combining ability variance; σ^2sca, specific combining ability variance.

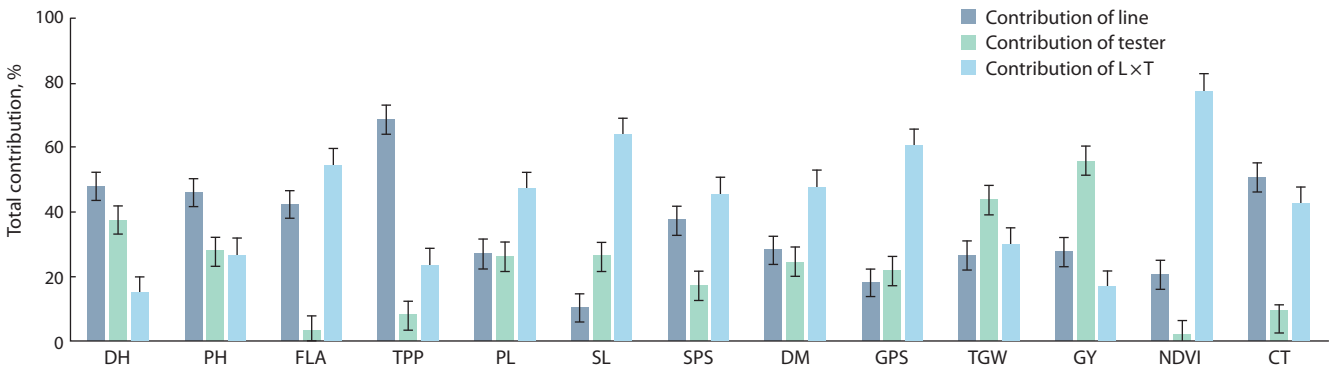

Fig. 1. Proportional contribution of lines, testers, and their interactions to total variance under rainfed conditions.

Here and in Figure 3: DH, days to heading; PH, plant height, cm; FLA, flag leaf area, cm^2; TPP, tillers per plant; PL, peduncle length, cm; SL, spike lenth, cm; SPS, spikelets per spike; DM, days to maturity; GPS, grains per spike; TGW, 1000-grain weight; GY, grain yield per plant, $kg \cdot ha^{-1}$; NDVI, normalized difference in vegetative index, and CT, canopy temperature, °C.

Phenotypic variance was depicted more as compared to genotypic variance in some traits, i. e. PH, FLA, SL, DM, and GY, while only GY showed high environmental variance. Broad sense heritability (H^2) was estimated in the range of 28.11 % (GY) to 100 % (CT). PH, PL, GPS, TGW, NDVI value and CT showed broad sense heritability of more than 91 %, while DH and GY depicted less heritability. The traits with high genetic variance, low environmental variance and high broad sense heritability have preponderance of additive genes and these are stable characters and selection in the filial generations can be made by keeping eye on these traits. Grain yield (GY) and DH attained low broad sense heritability and showed that environmental influence is more important for the expression of these traits. Selection in the filial generation should be made for these traits by considering disease incidence and drought proxy parameters, i. e. NDVI and CT values.

Proportional contribution of lines, testers, and their interactions to total variance

Proportional contribution of total variance for yield and yield-related metric traits for lines, testers and their cross combinations was estimated (Fig. 1). For DH, PH, TPP and CT it was recorded to be higher as compared to the testers and combinations of both lines and testers. Contribution of L×T to total variance was recorded as high in FLA, PL, SL, SPS, DM, GPS and NDVI value, while variance contribution of testers to TGW and GY was estimated higher as compared to the lines and L×T combinations.

General combining ability

General combining ability (GCA) estimates for all the traits are given in Table 5. Both positive and negative GCA effects were observed for lines and testers. For DH, the value of GCA effects ranged between 0.00 and 0.56. As a good general combiner, significant positive (0.56) and negative (–0.56) GCA effects were observed for lines Zargoon-79 and Pirsabak-2005, respectively. Similarly, in the testers, positive and significant (0.51) GCA effect was observed for Pakistan-2013 only (see Table 5). Patel et al. (2020) demonstrated ($p \leq 0.01$) significant negative and desirable GCA effects in lines and non-additive gene action was primarily involved in days to heading.

For plant height, negative general combining ability effects are more important since more emphasis is placed upon selection for short stature in segregating the population because it ultimately turns out that a short stature line is more responsive to fertilizer and tolerant to lodging. In this study, GCA effects ranged between –5.82 and 3.24 for PH. Significant positive (3.07) and negative (–5.82) GCA effects were observed for the lines Pirsabak-2005 and Punjab-2011, respectively. Similarly, highly significant positive (3.24) was estimated for the tester BOR-16 and highly significant but negative (–2.56) GCA effects were observed for the testers PAK-13, respectively. These results are in accordance with the results of (Singh S. et al., 2003; Gorjanović, Kraljević-Balalić, 2007).

For flag leaf area (FLA), negative general combining ability effects are more important because FLA is much influenced by transpiration losses due to disclosure to sunlight, which eventually affects the grain yield. Hence, more emphasis is retained on the selection of genotypes with smaller FLA. From that, among the female parents, Pirsabak-2005 and Punjab-2011 showed a highly significant negative GCA effect: –2.16 and –2.10, respectively. On the other side, no significant GCA effects were observed among the testers for FLA. These results confirm the findings of (Saeed A. et al., 2001; Arshad, Chowdhry, 2002; Chowdhary et al., 2007).

In case of tillers per plant (TPP), GCA effects ranged between –0.58 and 0.98. As a good general combiner, highly significant positive (0.98) GCA effects were observed only for the line MRJ-08 while there were no significant GCA effects among the testers for TPP. To begin with, TPP is a significant yield-boosting characteristic that contributes to increased grain yield. A higher number of tillers per plant confirms optimal plant populations and as a result higher grain yield (Tilley et al., 2019). For this point of view, the female line MRJ-08 showed better performance. These findings are in accordance with the results of (Iqbal M.M., 2007; Khan A. et al., 2020; Rashmi et al., 2020).

GCA effects ranging between –1.16 and 1.39 were observed for peduncle length (PL). Highly significant positive (1.39) and negative (–0.76) GCA effects were observed for the lines (female) FSD-08 and ZRG-09, respectively. In the same way, highly significant positive (0.65) and negative (–1.16) GCA effects were observed for the testers BOR-16 and PAK-13,

respectively. Likewise, in PH, shorter PL is preferred because an increase in PL ultimately increases the PH and we prefer a plant with short stature. In current study, two female parents, ZRG-79 (–0.76) and MRJ-08 (–0.70), showed negative general combining ability. Also, one male parent, PAK-13, showed superior general combining ability for this trait. So, it can be concluded that the above-mentioned parents are desirable for use in the breeding program. The findings of (Sharma, Garg, 2005) supported the results.

Greater spike length (SL) and larger number of spikelets per spike (SPS) are essential for enhanced yield. Among parents, one line (female), MRJ-08, showed significant positive values (0.77) for SPS. One tester, Pakistan-2013, exhibited high GCA for SPS. These results were quite close to the findings of (Awan et al., 2005; Sharma, Garg, 2005; Hassan et al., 2007). Number of grains per spike (GPS) is also an important factor for enhanced grain yield. Therefore, positive GCA effects are more important due to positive contribution of grain yield. Among male parents, only NR-09 showed positive and higher values (3.26) of GCA effects for GPS. Among female parents, MRJ-08 and PS-05 showed positive and higher values, i. e. 2.90 and 2.02 respectively. It should be noted that values of male parents were higher than those of female parents. These findings match with the results of (Saeed A. et al., 2001; Ahmadi et al., 2003; Saeed M.S. et al., 2005; Hassan et al., 2007). These results are different from the findings of Nazir et al. (2005).

For grain yield per plant (GY), only one female parent MRJ-08, and among the male parents, BOR-16 and NR-09, exhibited positive general combining ability effects. Similar results were also found by (Malik et al., 2005).

Specific combining ability

Specific combining ability (SCA) estimates for all the traits are given in Table 6. Both positive and negative SCA effects were observed among the crosses.

As for SCA effects for DH, all the fifteen crosses were of non-significant nature with positive and negative magnitude (see Table 6). The result indicates the involvement of both additive and non-additive genetic effects in the inheritance of DH, with greater proportion of additive genetic effect. Lines with maximum SCA effects can be used in development of hybrid cultivars. Only six among fifteen crosses depicted negative SCA effects for plant height. If parents with tallness are the ideal ones, then the crosses FSD-08 × NR-09, FSD-08 × PAK-13, PB-11 × NR-09, PB-11 × PAK-13, PS-05 × BOR-16 and MRJ-08 × BOR-16 would be considered good. However, the remaining crosses exhibited higher SCA effects. These findings confirmed the results of (Arshad, Chowdhry, 2002; Hasnain et al., 2006; Chowdhary et al., 2007). Furthermore, non-additive type of gene action is detected for PH and supported by (Babar et al., 2022). Also, our results concur with Ali F.K.H. and Abdulkhaleq (2019) for plant height.

GCA effects for flag leaf area range from negative –3.80 to positive 3.33. Roughly 50 % of the crosses showed smaller values of SCA effects for flag leaf area, which is desirable. As less flag leaf area is required for drought tolerance, the crosses with significant SCA effects, i. e. FSD-08 × BOR-16 and PB-11 × PAK-13 may be used in a future breeding program because they have high negative SCA values contributing towards minimum FLA. However, the remaining crosses exhibited higher positive SCA effects for FLA. Comparable results have also been stated by (Saeed A. et al., 2001; Arshad, Chowdhry, 2002; Chowdhary et al., 2007).

Negative SCA effects are needed to reduce the peduncle length (PL). In this study, two crosses showed significantly negative SCA effects. FSD-08 × NR-09 and MRJ-08 × BOR-16 are the best hybrids for reduced PL. Similar results were reported by (Chowdhary et al., 2007).

In case of spike length (SL), all the fifteen crosses were of non-significant nature with positive and negative magnitude (see Table 6). For a number of SPS, positive specific combining ability effects were shown in 6 out of 15 crosses but only two crosses, FSD-08 × NR-09 and PB-11 × BOR-16, have significant GCA effects. These hybrids performed best and can be suggested for future breeding programs. These results are in the conformity with those of (Mahantashivayogayya et al., 2004).

For grain yield per plant, SCA effects found varied much among crosses. The poorest cross with respect to SCA for grain yield per plant was ZRG-79 × PAK-13 whereas the cross that appeared to be the best and the most promising specific combination was ZRG-79 × NR-09. Positive specific combining ability effects were displayed in 8 out of 15 crosses. But only ZRG-79 × NR-09 showed such significant positive effects among crosses. Similar results were also reported by (Saeed A. et al., 2001).

Mid-parent heterosis estimation for grain yield

Mid-parental heterosis (MPH) for GY was estimated for 15 F_1 hybrids (Fig. 2). F_1 hybrids ZRG-79 × PAK-13 showed higher mid-parental value (62 %) followed by FSD-08 × PAK-13, ZRG-79 × BOR-16, and PB-11 × NR-09, which revealed mid-parent heterosis value above 30 % (34, 33, 31 % respectively). Cross combinations PS-05 × PAK-13, MRJ-08 × NR-09, MRJ-08 × PAK-13, FSD-08 × NR-09 depicted mid-parental heterosis value more than 15 %. Three cross combinations, ZRG-79 × NR-09, MRJ-08 × BOR-16 and PB-11 × BOR-16, depicted negative heterosis.

Cross combinations with more than 30 % mid-parental heterosis can be used in hybrid breeding in wheat. Heterotic studies for increasing wheat grain yield has been an interest of early wheat researchers. Pal and Alam (1938) reported mid-parent heterosis in the pre-green revolution era. After the introduction of semi-dwarf wheat in the post-green revolution era, various wheat researchers reported mid-parent heterosis in wheat, i. e. (Knott, 1965; Shamsuddin, 1985; Uddin et al., 1992). Barbosa-Neto et al. (1996) reported MPH in red soft winter wheat in the range of –20 to 57 %. Liu et al. (1999), Dreisigacker et al. (2005), Basnet et al. (2019) studied MPH in CIMMYT wheat varieties and reported MPH in the range of 9.5 to 14 %. Parental lines and tester used in present studies have CIMMYT background and the majority of the genotypes exhibited similar results for MPH. However, crosses combination ZRG-79 × PAK-13 has one indigenous parent ZRG-79 and exhibited a high percentage of MPH. These finding can demonstrate that crosses among parents with CIMMYT background have low heterotic potential and additive gene action governed the GY potential in these cross combinations and selection in the filial generation will be key for transgressive

Table 5. General combining ability effects of wheat genotypes, lines and testers for yield and its components under rainfed conditions

Entries	DH	PH	FLA	TPP	PL	SL	SPS	DM	GPS	TGW	GY	NDVI	CT
Lines													
FSD-08	–0.33^{ns}	0.40^{ns}	0.91^{ns}	–0.58^{ns}	1.39**	–0.11^{ns}	–1.02**	–0.11^{ns}	–0.96^{ns}	2.40**	–0.37^{ns}	0.01*	–0.29**
PB-11	–0.00^{ns}	–5.82**	–2.10**	0.20^{ns}	–0.46^{ns}	–0.27^{ns}	–0.43^{ns}	–1.11**	–2.77**	–4.20**	–2.70**	0.01*	–0.29**
PS-05	–0.56*	3.07**	–2.16**	–0.36^{ns}	0.53^{ns}	0.17^{ns}	0.53^{ns}	0.00^{ns}	2.02**	1.87**	0.65^{ns}	0.01*	–0.08**
MRJ-08	0.33^{ns}	0.07^{ns}	1.27^{ns}	0.98**	–0.70*	0.37^{ns}	0.77*	0.11^{ns}	2.90**	–0.26^{ns}	3.24**	–0.01**	0.29**
ZRG-79	0.56*	2.29**	2.08**	–0.24^{ns}	–0.76*	–0.16^{ns}	0.16^{ns}	1.11**	–1.20^{ns}	0.18^{ns}	–0.82*	–0.02**	0.37**
SE	0.2689	0.6645	0.7139	0.2981	0.3122	0.3671	0.3036	0.2934	0.6438	0.4877	0.3223	0.0050	0.0000
Testers													
NR-09	–0.22^{ns}	–0.69^{ns}	0.25^{ns}	0.13^{ns}	0.51*	–0.51^{ns}	–0.46^{ns}	–0.62*	3.26**	–0.39^{ns}	0.65*	0.01^{ns}	0.15**
PAK-13	0.51*	–2.56**	–0.69^{ns}	–0.27^{ns}	–1.16**	0.36^{ns}	0.60*	0.91**	–1.23*	–3.44**	–1.89**	–0.00^{ns}	–0.06**
BOR-16	–0.29^{ns}	3.24**	0.45^{ns}	0.13^{ns}	0.65*	0.16^{ns}	–0.14^{ns}	–0.29^{ns}	–2.03**	3.83**	1.25**	–0.00^{ns}	–0.08**
SE	0.21	0.51	0.55	0.23	0.24	0.28	0.24	0.23	0.50	0.38	0.25	0.00	0.00

Table 6. Specific combining ability effects of 15 wheat crosses for yield and related traits under rainfed conditions

Crosses	DH	PH	FLA	TPP	PL	SL	SPS	DM	GPS	TGW	GY	NDVI	CT
FSD-08 × NR-09	0.00^{ns}	–2.20*	0.75^{ns}	–0.02^{ns}	–2.31**	0.48^{ns}	1.20*	–0.16^{ns}	3.16**	–4.66**	–1.67**	0.03**	0.45**
FSD-08 × PAK-13	–0.07^{ns}	–1.33*	3.33*	0.04^{ns}	0.63^{ns}	0.18^{ns}	–0.09^{ns}	–0.69^{ns}	–0.58^{ns}	2.60**	1.07^{ns}	–0.03**	–0.39**
FSD-08 × BOR-16	0.07^{ns}	3.53**	–4.08**	–0.02^{ns}	1.68**	–0.66^{ns}	–1.11*	0.84^{ns}	–2.58*	2.06*	0.60^{ns}	0.01^{ns}	–0.07**
PB-11 × NR-09	0.00^{ns}	–1.64*	1.86^{ns}	0.20^{ns}	1.44*	0.17^{ns}	–0.52^{ns}	0.18^{ns}	–7.83**	–0.66^{ns}	–0.77^{ns}	–0.04**	0.20**
PB-11 × PAK-13	0.27^{ns}	–2.11*	–3.80**	–0.40^{ns}	–1.10^{ns}	–0.80^{ns}	–0.68^{ns}	0.98^{ns}	4.80**	–2.34**	0.37^{ns}	0.05**	–0.04**
PB-11 × BOR-16	–0.27^{ns}	3.76**	1.93^{ns}	0.20^{ns}	–0.34^{ns}	0.63^{ns}	1.20*	–1.16*	3.03*	3.00**	0.40^{ns}	–0.01^{ns}	–0.17**
PS-05 × NR-09	0.22^{ns}	0.13^{ns}	–0.95^{ns}	–0.58^{ns}	0.12^{ns}	–0.27^{ns}	–0.35^{ns}	0.73^{ns}	4.72**	1.88*	–0.42^{ns}	0.04**	–0.01**
PS-05 × PAK-13	–0.51^{ns}	2.67*	–0.57^{ns}	0.82^{ns}	0.48^{ns}	0.63^{ns}	0.59^{ns}	0.53^{ns}	–5.46**	1.40^{ns}	0.78^{ns}	–0.03**	0.10**
PS-05 × BOR-16	0.29^{ns}	–2.80*	1.52^{ns}	–0.24^{ns}	–0.60^{ns}	–0.37^{ns}	–0.23^{ns}	–1.27*	0.74^{ns}	–3.27**	–0.36^{ns}	–0.00^{ns}	–0.08**
MRJ-08 × NR-09	–0.00^{ns}	2.47*	–1.61^{ns}	0.09^{ns}	1.55**	–0.67^{ns}	–1.02^{ns}	–1.38*	–4.49**	2.94**	0.92^{ns}	0.01^{ns}	–0.08**
MRJ-08 × PAK-13	–0.07^{ns}	1.67*	–0.07^{ns}	–0.18^{ns}	–0.08^{ns}	0.90^{ns}	0.35^{ns}	–0.58^{ns}	3.00*	0.40^{ns}	0.09^{ns}	0.03**	0.18**
MRJ-08 × BOR-16	0.07^{ns}	–4.13**	1.69^{ns}	0.09^{ns}	–1.46*	–0.23^{ns}	0.67^{ns}	1.96**	1.49^{ns}	–3.34**	–1.01^{ns}	–0.03**	–0.10**
ZRG-79 × NR-09	–0.22^{ns}	1.24^{ns}	–0.06^{ns}	0.31^{ns}	–0.80^{ns}	0.29^{ns}	0.69^{ns}	0.62^{ns}	4.44**	0.50^{ns}	1.94**	–0.03**	–0.56**
ZRG-79 × PAK-13	0.38^{ns}	–0.89^{ns}	1.12^{ns}	–0.29^{ns}	0.07^{ns}	–0.91^{ns}	–0.17^{ns}	–0.24^{ns}	–1.77^{ns}	–2.05*	–2.32**	–0.01^{ns}	0.15**
ZRG-79 × BOR-16	–0.16^{ns}	–0.36^{ns}	–1.06^{ns}	–0.02^{ns}	0.72^{ns}	0.62^{ns}	–0.52^{ns}	–0.38^{ns}	–2.67*	1.55^{ns}	0.38^{ns}	0.04**	0.42**
SE	0.47	1.15	1.24	0.52	0.54	0.64	0.53	0.51	1.12	0.84	0.56	0.01	0.00

* Significant at $p < 0.05$; ** significant at $p < 0.01$; ^{ns} significant at $p > 0.05$.

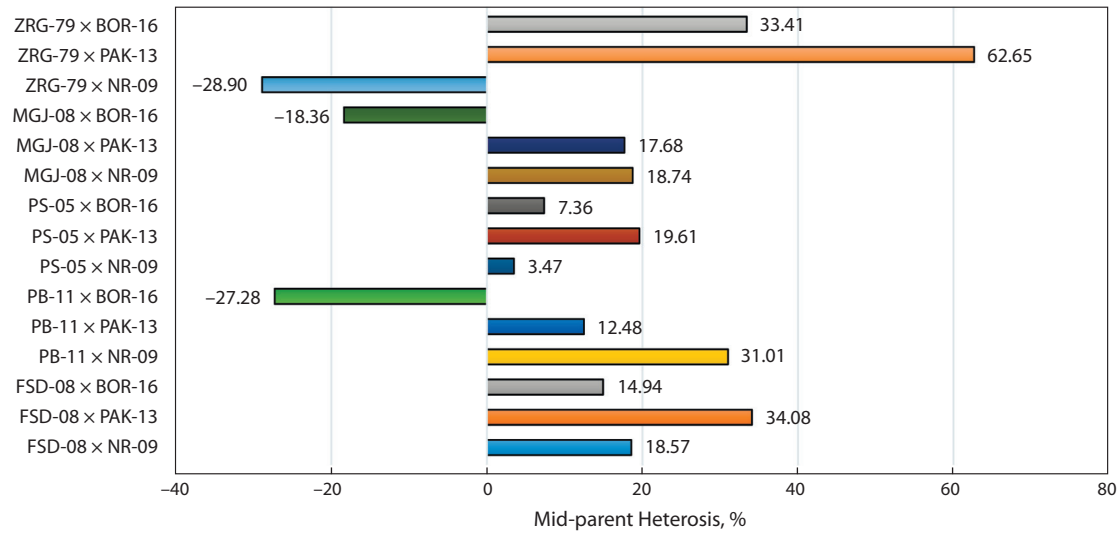

Fig. 2. Estimation of mid-parent heterosis for grain yield (GY) as a percentage increase or decrease in the F_1 hybrids compared to mid-parental value.

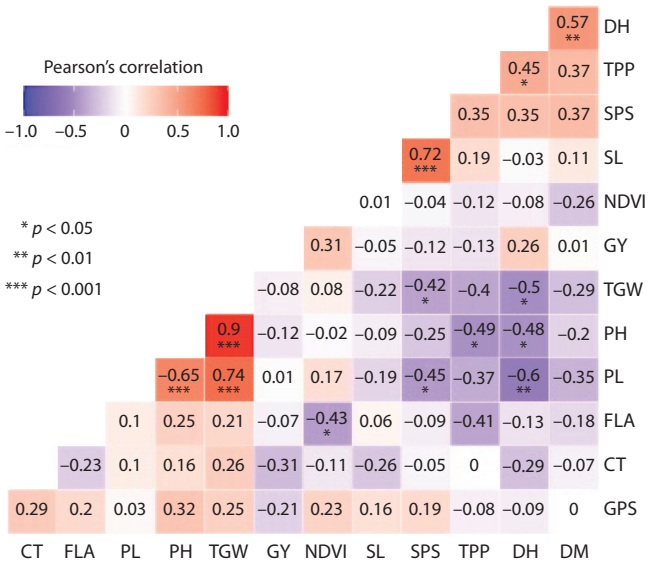

Fig. 3. Correlation study among yield and related traits of eight parents and 15 wheat crosses under rainfed conditions.

segregants, but in case of crosses among indigenous parents and genotypes with a CIMMYT parent it will be good source of hybrid breeding.

Correlation study of agronomic traits

Correlation study among yield and related traits under rainfed conditions of eight parents and 15 wheat crosses is mentioned in Figure 3. High significance was observed between PH and TGW with the value of 0.9 ($p < 0.001$), SL and SPS had a correlation coefficient value of 0.72 ($p < 0.001$) followed by DH and DM with 0.57 ($p < 0.01$). PL also showed highly significant and positive correlation with TGW and PH (0.74 and 0.65 respectively, $p < 0.001$). PL and TGW also revealed significant but negative correlation with DH –0.6 and –0.5, respectively ($p < 0.01$).

Positive and significant correlation among PH and PL with TGW showed that the higher the plant height the higher the thousand grain weight and peduncle length. Careful consideration should be made while selecting the genotypes with stiff and strong stem girth to avoid lodging. Correlation between SL and SPS revealed that an increase in spike length leads to an increase in spikelets per spike, genotypes with long spikes will be a good selection criterion for increasing yield due to the increase in number of spikelets per spike. Positive and significant correlation among DH and DM depicted that genotypes with early DH would mature earlier, so selection of genotypes with early flowering is good for early maturity and short duration variety development. Significant but negative correlation between TGW and DH indicated that a delay in days to flowering leads to a reduced TGW and *vice versa*. TGW showed negative correlation with TPP and these findings are in line with the results of Almutairi (2022). Low correlation of GY with other parameters in wheat was also reported by Gowda et al. (2010).

Stripe rust responses of parental lines and their cross combinations

The response to stripe rust (*Puccinia striiformis* f. sp. *tritici*) on parental lines used in the study and their offspring (crosses) is recorded for disease scoring, coefficient of infection (CI), average coefficient of infection (ACI), country average relative percentage attack (CARPA) and rust resistance index (RRI) (Table 7). All the parental lines showed moderate resistant (MR) to highly resistant (R) reaction against stripe rust (*Pst*). The female parents (lines) FSD-08, PB-11, PS-05, MRJ-08 and ZRG-79 showed 20M, 20M, 5M, 30M and 40M scores respectively, while the pollen parents (testers) viz. PAK-13, BOR-16 and NR-09 depicted 10MR, 5R and 40M response against stripe rust. All these parents showed a slow rusting response against rust pathogen that is under the control of multiple genes.

Cross combinations of these parental lines showed a varied response, moderately resistant to moderately susceptible reac-

Table 7. Response of parental genotypes and their cross combinations against stripe rust infection under rainfed conditions

Parents and their crosses	Stripe rust observations			CI total	ACI	CARPA	RRI
	Rep-1	Rep-2	Rep-3				
FSD-08	20M	20M	30M	42	14.0	35.90	**5.8**
PB-11	20M	30M	40S	70	23.3	59.83	3.6
PS-05	5M	5R	5MR	7	2.0	5.13	**8.5**
MRJ-08	30M	50MS	50MSS	103	34.3	88.03	1.1
ZRG-79	40M	5MS	60S	88	29.3	75.21	2.2
PAK-13	10MR	20MR	20MR	20	6.7	17.09	**7.5**
BOR-16	5R	10MR	20MR	13	4.3	11.11	**8.0**
NR-09	40M	40M	40M	72	24.0	61.54	3.5
FSD-08 × PAK-13	20M	30MS	40M	60	20.0	51.28	4.4
FSD-08 × BOR-16	20M	30MS	40M	60	20.0	51.28	4.4
FSD-08 × NR-09	30MSS	30MR	60S	99	33.0	84.62	1.4
PB-11 × PAK-13	20M	4MS	30MS	39	13.1	33.50	**6.0**
PB-11 × BOR-16	20M	30R	30M	36	12.0	30.77	**6.2**
PB-11 × NR-09	30MSS	60MSS	40MSS	117	39.0	100	0.0
PS-05 × PAK-13	10MR	50MSS	30MR	61	20.3	52.14	4.3
PS-05 × BOR-16	10MR	10MR	10MR	12	4.0	10.26	**8.1**
PS-05 × NR-09	5MR	40M	30M	44	14.7	37.61	**5.6**
MRJ-08 × PAK-13	30M	5R	10M	25	8.3	21.37	**7.1**
MRJ-08 × BOR-16	20M	10M	10M	72	24.0	61.54	3.5
MRJ-08 × NR-09	40MS	40S	30MS	96	32.0	82.05	1.6
ZRG-79 × PAK-13	20MS	40MS	30MSS	75	25.0	64.10	3.2
ZRG-79 × BOR-16	20MS	30MSS	30MSS	70	23.3	59.83	3.6
ZRG-79 × NR-09	30M	30MSS	30MSS	72	24.0	61.54	3.5

Note. R, resistant; S, susceptible; MR, moderately resistant; MS, moderately susceptible; MSS, moderately susceptible to susceptible; CI, coefficient of infection; ACI, average coefficient of infection; CARPA, country average relative percentage attack; RRI, relative rust index.

tion against stripe rust. The F_1 hybrids combinations PS-05 × PAK-13, PS-05 × BOR-16 and PS-05 × NR-09 showed 10MR, 10MR and 5MR reaction, the crosses FSD-08 × PAK-13, FSD-08 × BOR-16, PB-11 × PAK-13, PB-11 × BOR-16, and MRJ-08 × BOR-16 showed 20M reaction, the cross combination MRJ-08 × PAK-13 showed 30M reaction, while the rest of the crosses showed moderately susceptible to susceptible reaction against stripe rust.

Average coefficient of infection (ACI) for the parents PS-05, PAK-13, BOR-16 and FSD-08 was recorded as 2.0, 6.7, 4.3 and 14.0 respectively and these varieties revealed a very good level of resistance against stripe rust. Rust resistance index (RRI) of these parents was also high (ranged 5.8 to 8.5), which indicated a good resistance response of these varieties. Among the cross combinations, PS-05 × BOR-16, MRJ-08 × PAK-13, PB-11 × BOR-16, PB-11 × PAK-13 and PS-05 × NR-09 depicted ACI values of 4.0, 8.3, 12.0, 13.1, and 14.7, respectively. These F_1 hybrids had a resistant response to stripe rust. RRI value of the F_1 hybrids (PS-05 × BOR-16, MRJ-08 × PAK-13, PB-11 × BOR-16, PB-11 × PAK-13 and PS-05 × NR-09) was higher (ranging from 5.6 to 8.1).

The higher the RRI value and the lower the ACI value means of genotypes with a resistant response to the disease pathogen and under the influence of slow rusting genes, the slower the disease progress and the lesser the yield losses. Genotypes with higher RRI values (>5.0) represent moderately resistant to highly resistant response against rust pathogen. The parental genotypes viz. PS-05, PAK-13 and BOR-16 had higher values for RRI (8.5, 7.5 and 8.0 respectively) showing a highly resistant response against stripe rust pathogen. Cross combinations revealed an intermediate response against stripe rust as compared to parents, especially testes, and resistant genes are under the control of additive gene action. These results indicate that repeated backcross can be a better strategy for accumulation of resistant genes in these cross combinations. Selection in these cross combinations by following backcrosses with recurrent parents is efficient for disease resistance in the filial generations. These results are very much in line with the findings of Afzal et al. (2009) and Mahmoud et al. (2015).

Conclusion

According to these findings, it can be concluded that higher general combining ability and low broad sense heritability for grain yield suggest the presence of additive genes, and exploitation of general combining ability for high grain yield is important due to presence of additive gene action, and selection in the filial generations and family rows will be effective.

For development of heterotic population, it is important to exploit specific combining ability for dominant gene action by crossing indigenous genotypes with exotic germplasm with improved rust resistance, which will be a useful future breeding strategy.

References

Afzal S.N., Haque M.I., Ahmedani M.S., Munir M., Firdous S.S., Rauf A., Ahmad I., Rattu A.R., Fayyaz M. Resistance potential of wheat germplasm (*Triticum aestivum* L.) against stripe rust disease under rainfed climate of Pakistan. *Pak. J. Bot.* 2009;41(3):1463-1475.

Ahmad M.W., Ahmed M.S., Tahir H.N. Combining ability analysis for achene yield and related traits in sunflower (*Helianthus annuus* L.). *Chil. J. Agric. Res*. 2012;72(1):21-26. DOI 10.4067/S0718-58392012000100004.

Ahmadi J.A., Zali A., Yazdi-Samadi B., Talaie A., Ghannadha M.R., Saeidi A. A study of combining ability and gene effect in bread wheat under drought stress condition by diallel method. *Iranian J. Agric. Sci*. 2003;34(1):1-8.

Akhtar M.A., Ahmad I., Mirza J.I., Rattu A.R., Hakro A.A., Jaffery A.H. Evaluation of candidate lines against stripe and leaf rusts under national uniform wheat and barley yield trial 2000–2001. *Asian J. Plant Sci*. 2002;1:450-453. DOI 10.3923/ajps.2002.450.453.

Ali F.K.H., Abdulkhaleq D.A. Inheritance of some growth and yield traits in bread wheat using line × tester analysis. *J. Zankoy Sulaimani*. 2019;21(2):131-149. DOI 10.17656/jzs.10763.

Ali Y., Khan M.A., Hussain M., Sabir W., Atiq M., Aatif H.M., Ahmad S., Ijaz M., Ahmad J.N. Virulence analysis of leaf and stripe rust populations in Pakistan through avirulence to virulence formula. *Arch. Phytopathol. Plant Prot*. 2020;53(17-18):844-855. DOI 10.1080/03235408.2020.1802565.

Almutairi M.M. Genetic parameters estimation for some wild wheat species and their F_1 hybrids grown in different regions of Saudi Arabia. *Saudi J. Biol. Sci.* 2022;29(1):521-525. DOI 10.1016/j.sjbs.2021.09.015.

Arshad M., Chowdhry M.A. Impact of environment on the combining ability of bread wheat genotypes. *Pak. J. Biol. Sci*. 2002;5(12):1316-1320. DOI 10.3923/pjbs.2002.1316.1320.

Arzu K. Gene action and combining ability in line × tester population of safflower (*Carthamus tinctorius* L.). *Turk. J. Field Crops*. 2017;22(2):197-203. DOI 10.17557/tjfc.356216.

Awan S.I., Malik F.A., Siddique M. Combining ability analysis in intervarietal crosses for component traits in hexaploid wheat. *J. Agric. Soc. Sci*. 2005;1(4):316-317.

Babar M., Ullah H., Afridi K., Raza H., Ali S., Ali M., Subhan G., Khan J., Khan H. Line × tester analysis for combining ability and identification of gene action in F_2 populations of bread wheat. *Int. J. Agric. Biol.* 2022;27(6):381-392. DOI 10.17957/IJAB/15.1938.

Barbosa-Neto J., Sorrells M., Cisar G. Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship. *Genome*. 1996;39(6):1142-1149. DOI 10.1139/g96-144.

Basnet B.R., Crossa J., Dreisigacker S., Pérez-Rodríguez P., Manes Y., Singh R.P., Rosyara U.R., Camarillo-Castillo F., Murua M. Hybrid wheat prediction using genomic, pedigree, and environmental covariables interaction models. *Plant Genome*. 2019;12(1):180051. DOI 10.3835/plantgenome2018.07.0051.

Burton G., Devane H. Estimating heritability in tall fescue (*Festuca arundinacea*), from replicated clonal material. *Agron. J*. 1953;45(10):478-481. DOI 10.2134/agronj1953.00021962004500100005x.

Chen X., Penman L., Wan A., Cheng P. Virulence races of *Puccinia striiformis* f. sp. *tritici* in 2006 and 2007 and development of wheat stripe rust and distributions, dynamics, and evolutionary relationships of races from 2000 to 2007 in the United States. *Can. J. Plant. Pathol*. 2010;32(3):315-333. DOI 10.1080/07060661.2010.499271.

Chowdhary M.A., Sajad M., Ashraf M.I. Analysis on combining ability of metric traits in bread wheat, *Triticum aestivum*. *J. Agric. Res.* 2007:45(1):11-17.

Comstock R., Robinson H. Genetic parameters, their estimation and significance. In: Proceedings of the Sixth International Grassland Congress (Pennsylvania, August 17–23, 1952). Pennsylvania: State College, 1952;248-291.

De Mendiburu F., Simon R. Agricole – ten years of an open source statistical tool for experiments in breeding, agriculture and biology. *PeerJ PrePrints*. 2015:3:e1404v1. DOI 10.7287/peerj.preprints.1404v1.

Dragov R.G. Combining ability for quantitative traits related to productivity in durum wheat. *Vavilov J. Genet. Breed*. 2022;26(6):515-523. DOI 10.18699/VJGB-22-63.

Dreisigacker S., Melchinger A., Zhang P., Ammar K., Flachenecker C., Hoisington D., Warburton M.L. Hybrid performance and heterosis in spring bread wheat, and their relations to SSR-based genetic distances and coefficients of parentage. *Euphytica*. 2005;144:51-59. DOI 10.1007/s10681-005-4053-2.

Fakthongphan J., Graybosch R.A., Baenziger P.S. Combining ability for tolerance to pre-harvest sprouting in common wheat (*Triticum aestivum* L.). *Crop Sci*. 2016;56(3):1025-1035. DOI 10.2135/cropsci2015.08.0490.

FAOSTAT. Food and Agriculture Organization of the United Nations. Rome (Italy): FAOSTAT, 2016.

Fasahat P., Rajabi A., Rad J.M., Derera J. Principles and utilization of combining ability in plant breeding. *Biom. Biostat. Int. J*. 2016;4(1):1-22. DOI 10.15406/bbij.2016.04.00085.

Fellahi Z.E.A., Hannachi A., Bouzerzour H., Boutekrabt A. Line × tester mating design analysis for grain yield and yield related traits in bread wheat (*Triticum aestivum* L.). *Int. J. Agron*. 2013;2013:201851. DOI 10.1155/2013/201851.

Figueroa M., Dodds P.N., Henningsen E.C. Evolution of virulence in rust fungi – multiple solutions to one problem. *Curr. Opin. Plant Biol*. 2020;56:20-27. DOI 10.1016/j.pbi.2020.02.007.

Fonseca S., Patterson F.L. Hybrid vigour in seven parent diallel crosses in common wheat (*Triticum aestivum* L.). *Crop Sci*. 1968;8(1):85-89. DOI 10.2135/cropsci1968.0011183X000800010025x.

Gorjanović B.M., Kraljević-Balalić M.M. Inheritance of plant height, spike length and number of spikelets per spike in durum wheat. *Proc. Nat. Sci. Matica Srpska Novi Sad*. 2007;112:27-33. DOI 10.2298/ZMSPN0712027G.

Gowda M., Kling C., Würschum T., Liu W., Maurer H.P., Hahn V., Reif J.C. Hybrid breeding in durum wheat: heterosis and combining ability. *Crop Sci*. 2010;50(6):2224-2230. DOI 10.2135/cropsci2009.10.0637.

Hallauer A.R., Carena M.J., Miranda Filho J.D. Quantitative Genetics in Maize Breeding: Handbook of plant breeding. Springer, 1988.

Hasnain Z., Abbas G., Saeed A., Shakeel A., Muhammad A., Rahim M.A. Combining ability for plant height and yield related traits in wheat (*Triticum aestivum* L.). *J. Agric. Res*. 2006;44(3):167-173.

Hassan G., Muhammad F., Afridi S.S., Khalil I.H. Combining ability in F_1 generations of diallel cross for yield components in wheat. *Sarhad J. Agric*. 2007;23(4):937-942.

Intikhab A., Awan S.I., Mur L.A.J., Saeed M.S., Ahmed M.S. Molecular and phenotypic analysis of bread wheat varieties in relation to durable rust resistance. *Int. J. Agric. Biol*. 2021;26(2):329-336. DOI 10.17957/IJAB/15.1841.

Iqbal A., Khalil I.H., Shah S.M., Kakar A.M.S. Estimation of heritability, genetic advance and correlation for morphological traits in spring wheat. *Sarhad J. Agric*. 2017;33(4):674-679. DOI 10.17582/journal.sja/2017/33.4.674.679.

Iqbal M.M. Combining ability analysis in wheat. *Pak. J. Agric. Sci*. 2007;44(1):1-5.

Ishaq M., Ahmad G., Afridi K., Ali M., Khan T.U., Shah I.A., Ahmad B., Ahmad N., Ahmad I., Saleem A., Miraj M. Combining abi-

lity and inheritance studies for morphological and yield contributing attributes through line × tester mating design in wheat (*Triticum aestivum* L.). *Pure Appl. Biol*. 2018;7(1):160-168.

Javaid M.M., Zulkiffal M., Ali Y., Mehmood A., Ahmed J., Hussain M., Muhammad F., Sabir W., Tanveer M.H., Yasin O. Impact of environmental and pathogenic variability on breaking of host rust resistance in wheat cultivars under changing climatic conditions. *Adv. Zool. Bot*. 2018;6(1):31-40. DOI 10.13189/azb.2018.060104.

Jones D.F. Dominance of linked factors as a means of accounting for heterosis. *Proc. Natl. Acad. Sci. USA*. 1917;3(4):310-312. DOI 10.1073/pnas.3.4.310.

Kaushik P. Line × tester analysis for morphological and fruit biochemical traits in eggplant (*Solanum melongena* L.) using wild relatives as testers. *Agronomy*. 2019;9(4):185. DOI 10.3390/agronomy 9040185.

Kempthorne O. An Introduction to Genetic Statistics. New York: John Wiley and Sons, 1957.

Khan A., Ahmad M., Shah M.K.N., Ahmed M. Genetic manifestation of physio-morphic and yield related traits conferring thermotolerance in wheat. *Pak. J. Bot*. 2020;52(5):1545-1552. DOI 10.30848/PJB2020-5(27).

Khan S.N., Hassan G., Khan M.R., Facho Z.H., Singh D., Sandhu K.S., Sanaullah M., Imtiaz M., Ali S. Field assessment and molecular markers-based characterization of yellow rust resistance in wheat hybrid progenies. *J. Anim. Plant Sci*. 2022;32(1):127-137. DOI 10.36899/JAPS.2022.1.0409.

Knott D.R. Heterosis in seven wheat hybrids. *Can. J. Plant Sci*. 1965; 45:499-501.

Kumar R., Silva L. Light ray tracing through a leaf cross section. *Appl. Opt*. 1973;12(12):2950-2954. DOI 10.1364/AO.12.002950.

Large E.C. Growth stages in cereals illustration of the Feekes scale. *Plant Pathol*. 1954;3(4):128-129. DOI 10.1111/j.1365-3059.1954.tb00716.x.

Lerner I.M. Genetic Homeostasis. New York: John Wiley and Sons, 1954.

Liu Z., Pei Y., Pu Z. Relationship between hybrid performance and genetic diversity based on RAPD markers in wheat (*Triticum aestivum* L.). *Plant Breed*. 1999;118(2):119-123. DOI 10.1046/j.1439-0523.1999.118002119.x.

Mahantashivayogayya K., Hanchinal R., Salimath P. Combining ability in dicoccum wheat. *Karnataka J. Agric. Sci*. 2004;17(3):451-454.

Mahmoud A.F., Hassan M.I., Amein K.A. Resistance potential of bread wheat genotypes against stripe rust disease under Egyptian climate. *Plant Pathol. J*. 2015;31(4):402-413. DOI 10.5423/PPJ.OA.12.2014.0127.

Malik M.F.A., Awan S.I., Ali S. Genetic behaviour and analysis of quantitative traits in five wheat genotypes. *J. Agric. Soc. Sci*. 2005; 1(4):313-315.

Nazir S., Khan A.S., Ali Z. Combining ability analysis for yield contributing traits in bread wheat. *J. Agric. Soc. Sci*. 2005;1:129-132.

Okello D.K., Manna R., Imanyowoha J., Pixley K., Edema R. Agronomic performance and breeding potential of selected inbred lines for improvement of protein quality of adapted Ugandan maize germplasm. *African Crop Sci. J*. 2006;14(1):37-47.

Pal B., Alam N. The effect of certain external factors upon the manifestation of hybrid vigour in wheat. *Proc. Indian Acad. Sci*. 1938;7: 109-124. DOI 10.1007/BF03051095.

Patel P.U., Patel B.C., Sidapara M.P., Sharma D.D. Combining ability and gene action studies for yield and its component traits in bread wheat (*Triticum aestivum* L.). *Int. J. Curr. Microbiol. Appl. Sci*. 2020;9:2463-2469.

Peterson R.F., Campbell A.B., Hannah A.E. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Can. J. Res*. 1948;26(5):496-500. DOI 10.1139/cjr48c-033.

Qamar M., Ahmad S.D., Rabbani M.A., Shinwari Z.K., Iqbal M. Determination of rust resistance genes in Pakistani bread wheats. *Pak. J. Bot*. 2014;46(2):613-617.

Rashid M., Cheema A.A., Ashraf M. Line × tester analysis in Basmati rice. *Pak. J. Bot*. 2007;39(6):2035-2042.

Rashmi K., Prasad L.C., Prasad R., Thakur P., Chandra K. Inheritance study of yield traits and spot blotch disease resistance in wheat under epiphytotic environment. *Plant Arch*. 2020;20(2):6590-6594.

Raza W., Ghazanfar M.U., Rehman A.U., Fayyaz M. Screening of wheat germplasm against stripe rust disease under field conditions in Pakistan. *Plant Prot*. 2018;2(3):87-92.

Reynolds M.P., Singh R.P., Ibrahim A., Ageeb O.A.A., Larque-Saavedra A., Quick J.S. Evaluating the physiological traits to complement empirical selection for wheat in warm environments. *Euphytica*. 1998;100:85-94. DOI 10.1023/A:1018355906553.

Saeed A., Chowdhry M.A., Saeed N., Khaliq I., Johar M.Z. Line × tester analysis for some morpho-physiological traits in bread wheat. *Int. J. Agric. Biol*. 2001;3(4):444-447.

Saeed M.S., Chowdhry M.A., Ahsan M. Genetic analysis for some metric traits in *Aestivum* species. *Asian J. Plant Sci*. 2005;4(4):413-416. DOI 10.3923/ajps.2005.413.416.

Shamsuddin A K. Genetic diversity in relation to heterosis and combining ability in spring wheat. *Theor. Appl. Genet*. 1985;70:306-308.

Sharma A., Garg D. Combining ability over environments in bread wheat (*Triticum aestivum* L.). *J. Maharasahtra Agric. Univ*. 2005; 30(2):153-156.

Shull G.H., East E.M. The composition of a field of maize. *J. Hered*. 1908;4(1):296-301. DOI 10.1093/jhered/os-4.1.296.

Singh R.K., Chaudhary B.D. Biometrical Methods in Quantitative Genetic Analysis. New Delhi: Kalyani Publ., 1977.

Singh S., Singh L., Singh D., Kumar R. Combining ability in common wheat (*Triticum aestivum* L.) grown in sodic soil. *Progress. Agric*. 2003;3:78-80.

Sprague G.F., Tatum L.A. General *vs.* specific combining ability in single crosses of corn. *Agron. J*. 1942;34(10):923-932.

Steel R.G.D., Torrie J.H. Principles and Procedures of Statistics. New York: McGraw Hill Book Company Inc., 1980.

Sultana S.R., Ali A., Ahmad A., Mubeen M., Zia-Ul-Haq M., Ahmad S., Ercisli S., Jaafar H.Z. Normalized Difference Vegetation Index as a tool for wheat yield estimation: a case study from Faisalabad, Pakistan. *Sci. World J*. 2014;2014:725326. DOI 10.1155/2014/725326.

The R Project for Statistical Computing. A language and environment for statistical computing R foundation for Statistical Computing Department of Agronomy, Faculty of Agriculture of the University of the Free State. Vienna, Austria, 2017. www.r-project.org.

Tilley M.S., Heiniger R.W., Crozier C.R. Tiller initiation and its effects on yield and yield components in winter wheat. *Agron. J*. 2019;111(3):1323-1332. DOI 10.2134/agronj2018.07.0469.

Uddin M.N., Ellison F.W., O'Brien L., Latter B.D.H. Heterosis in F_1 hybrids derived from crosses of adapted Australian wheats. *Aust. J. Agric. Res*. 1992;43(5):907-919. DOI 10.1071/AR9920907.

Ullah N., Ullah H., Afridi K., Alam M., Jadoon S.A., Khan W.U., Ahmad M., Uddin H. Genetic variability, heritability and correlation analysis among morphological and yield traits in wheat advanced lines. *Biol. Divers. Conserv.* 2018;11(1):166-180.

Zhao Y., Zeng J., Fernando R., Reif J.C. Genomic prediction of hybrid wheat performance. *Crop Sci*. 2013;53(3):802-810. DOI 10.2135/cropsci2012.08.0463.

Conflict of interest. The authors declare no conflict of interest.

Received January 2, 2023. Revised March 9, 2023. Accepted March 9, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):623-632
DOI 10.18699/VJGB-23-73

Original Russian text https://vavilovj-icg.ru/

Prebreeding studies of leaf rust resistant *Triticum aestivum/T. timopheevii* line L624

S.N. Sibikeev[1] ✉, I.G. Adonina[2, 3], A.E. Druzhin[1], O.A. Baranova[4]

[1] Federal Center of Agriculture Research of the South-East Region, Saratov, Russia
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
[4] All-Russian Research Institute of Plant Protection, Pushkin, St. Petersburg, Russia
✉ sibikeev_sergey@mail.ru

Abstract. *Triticum timopheevii* Zhuk. attracts the attention of bread wheat breeders with its high immunity to the leaf rust pathogen. However, introgressions from this species in *Triticum aestivum* L. are little used in practical breeding. In the presented study, the agronomic value of *T. aestivum/T. timopheevii* line L624 was studied in comparison with the parent cultivars Saratovskaya 68, Dobrynya and the standard cultivar Favorit during 2017–2022. Introgressions from *T. timopheevii* in L624 were detected by the FISH method with probes pSc119.2, pAs1 and Spelt1, as well as microsatellite markers *Xgwm312*, *Xgpw4480* and *Xksum73*. Translocations of 2AS.2AL-2A^{t}L and on 2DL were detected as well. Line L624 is highly resistant to *Puccinia triticina* both under the background of natural epiphytotics and under laboratory conditions. PCR analysis with the DNA marker of the *LrTt1* gene (*Xgwm312*) revealed that it is not identical to the *Lr* gene(s) in L624. According to a five-year study, the grain yield of L624 was, on average, higher than that of Favorit and Dobrynya, but lower than that of Saratovskaya 68. Line L624 had a lower weight of 1000 grains than the recipients, and was at the same level with the standard cultivar Favorit. Introgressions from *T. timopheevii* in L624 increased the grain protein content by comparison with Saratovskaya 68 and Favorit, but it was at the same level as in Dobrynya. As for parameters of flour and bread, L624 was not inferior to the recipient cultivars, but by volume and porosity of bread, it surpassed Saratovskaya 68. Moreover, L624 surpassed Favorit by the elasticity of the dough, the ratio of the elasticity of the dough to the extensibility and the strength of the flour. Thus, the results obtained suggest that introgressions in chromosomes 2A and 2D in L624 do not impair baking properties.
Key words: *Triticum aestivum/T. timopheevii* line; 2AS.2AL-2A^{t}L translocation; leaf rust resistance; impact on productivity and grain quality.

For citation: Sibikeev S.N., Adonina I.G., Druzhin A.E., Baranova O.A. Prebreeding studies of leaf rust resistant *Triticum aestivum/T. timopheevii* line L624. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):623-632. DOI 10.18699/VJGB-23-73

Пребридинговые исследования устойчивой к листовой ржавчине *Triticum aestivum/T. timopheevii* линии Л624

С.Н. Сибикеев[1] ✉, И.Г. Адонина[2, 3], А.Е. Дружин[1], О.А. Баранова[4]

[1] Федеральный аграрный научный центр Юго-Востока, Саратов, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия
[4] Всероссийский научно-исследовательский институт защиты растений, Пушкин, Санкт-Петербург, Россия
✉ sibikeev_sergey@mail.ru

Аннотация. Вид *Triticum timopheevii* Zhuk. привлекает внимание селекционеров мягкой пшеницы высоким иммунитетом к возбудителю листовой ржавчины. Однако интрогрессии от этого вида в *T. aestivum* L. мало используются в практической селекции. В представленном исследовании изучена агрономическая ценность *T. aestivum/T. timopheevii* линии Л624 по сравнению с родительскими сортами Саратовская 68, Добрыня и сортом-стандартом Фаворит в течение 2017–2022 гг. Интрогрессии от *T. timopheevii* у Л624 выявлены с помощью метода FISH с зондами pSc119.2, pAs1 и Spelt1, а также микросателлитных маркеров *Xgwm312*, *Xgpw4480* и *Xksum73*. Обнаружены транслокации 2AS.2AL-2A^{t}L и в длинном плече хромосомы 2D. Линия Л624 высокоустойчива к *Puccinia triticina* как на фоне естественной эпифитотии, так и в лабораторных условиях. С использованием ПЦР-анализа с ДНК-маркером гена *LrTt1* (*Xgwm312*) установлена его неидентичность *Lr*-гену(ам) у линии Л624. По данным пятилетнего изучения, урожайность зерна у Л624 была в среднем выше, чем у сортов Фаворит и Добрыня, но ниже, чем у сорта Саратовская 68. По массе 1000 зерен Л624 уступала реципиентам и была одного уровня с сортом-стандартом Фаворит. Интрогрессии от *T. timopheevii* у Л624 увеличили содержание белка в зерне по сравнению с сортами Саратовская 68 и Фаворит, но с сортом Добрыня оно было на одном уровне. В целом по показателям качества муки и хлеба линия Л624 не уступила сортам-реципиентам, а по объ-

ему и пористости хлеба превзошла Саратовскую 68. В то же время Л624 превышала сорт-стандарт Фаворит по упругости теста, отношению упругости теста к растяжимости и силе муки. Таким образом, полученные результаты позволяют сделать предположение, что интрогрессии в хромосомах 2А и 2D у линии Л624 не ухудшают хлебопекарные свойства.
Ключевые слова: *Triticum aestivum/T. timopheevii* линия; 2AS.2AL-2A^{t}L транслокация; устойчивость к листовой ржавчине; влияние на продуктивность и качество зерна.

Introduction

Leaf rust (caused by the fungus *Puccinia triticina* Eriks.) is a harmful disease both in Russia and abroad. Despite the fact that the importance of this disease has decreased for a number of grain-growing Russian regions, the damage remains quite extensive (Gultyaeva et al., 2021). The production of resistant bread wheat cultivars makes it possible to avoid economically significant damage to plants by this pathogen. Resistance to *P. triticina* in bread wheat is controlled by *Lr*-genes. To date, 82 *Lr*-genes have been identified (McIntosh et al., 2022), but most of them have been overcome by the pathogen. In general, the following genes are used in Russian bread wheat cultivars: *Lr1*, *Lr3*, *Lr9*, *Lr10*, *Lr19*, *Lr20*, *Lr24*, *Lr26*, *Lr34*, *Lr37* and *Lr6Agi1*, *Lr6Agi2*, *LrSp*. These genes are used in various combinations, but only the *LrSp*, *Lr6Agi1*, and *Lr6Agi2* genes are not overcome (Gultyaeva et al., 2021).

The low genetic diversity of effective resistance genes to the leaf rust pathogen can be solved by involving species related to bread wheat in hybridization. Thus, out of 82 identified *Lr*-genes, 39 were transferred from alien species (McIntosh et al., 2013, 2018, 2022). *Triticum timopheevii* Zhuk. (A^{t}A^t GG, 2*n* = 28) is one of the sources of effective resistance genes. This species is very popular among breeders because of its unique high resistance to a complex of diseases. Both in Russia and abroad, numerous attempts have been made to transfer pathogens resistance genes from *T. timopheevii* into bread and durum wheat by direct hybridization followed by backcrossing to cultivated species (Allard, Shands, 1954; Jørgensen, Jensen, 1972; Skurygina, 1984; Tomar et al., 1988; Kozlovskaya et al., 1990; Budashkina, Kalinina, 2001; Brown-Guedria et al., 2003; Singh et al., 2017).

In addition, synthetic alloploid forms from crossing *T. timopheevii* with *Aegilops tauschii* Coss. were produced as an intermediate form – a "bridge". As a result, forms with 2*n* = 42 and genome composition A^{t}A^{t}GGDD – *T. kiharae* Dorof. et Migusch. were obtained (Dorofeev et al., 1979), as well as the synthetic of Dr. Savov (Leonova et al., 2007). With similar hybridization with the natural mutant *T. timopheevii* – *T. militinae* Zhuk. et Migusch, a synthetic form *Triticum miguschovae* was obtained (Zhirov, Ternovskaya, 1984).

Subsequently, when bread wheat was crossed with these synthetics, a number of lines that were resistant to leaf and stem rust pathogens were obtained. Thus, a set of pathogen-resistant lines in the gene background of the Saratovskaya 29 cultivar, the so-called C29 immune (C29im), as well as the spring bread wheat cultivar Pamyati Maistrenko, was obtained at the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences (Laikova et al., 2013).

However, despite numerous studies, only two *Lr*-genes from *T. timopheevii* have been identified – *Lr18* in the 5BS.5BL-5G#1L translocation and *Lr50* localized on 2BL (McIntosh et al., 2013). In addition, the "Catalogue of Genes Symbols for Wheat" (McIntosh et al., 2013) lists three *Lr*-genes with a temporary (laboratory) designation: *LrTt1* located on chromosome 2A, *LrTt2* located on chromosome 5BL (Leonova et al., 2004, 2010), *LrSelG12* located on chromosome 3BL (Singh et al., 2017). Thus, the list of *Lr*-genes from *T. timopheevii* is small.

The effectiveness of these genes varies. Thus, the pathogen of leaf rust has virulence to the *Lr18* gene in Germany, Switzerland and Russia (McIntosh et al., 1995; Sibikeev et al., 2020), but this gene is more effective in Australia (McIntosh et al., 1995). The MBRL and PNMQ races are virulent to the *Lr50* gene at the seedling stage, but this gene is effective at the adult plant stage in Kansas and Texas (Brown-Guedria et al., 2003). The *LrTt1* gene is effective at the seedling phase (Leonova et al., 2004), while *LrTt2* and *LrSelG12* are effective both at seedlings and adult plants (Leonova et al., 2010; Singh et al., 2017).

However, practical wheat breeding has shown little use of these genes in commercial cultivars. Only *Lr18* is used in Australia in the Timvera cultivar and its derivatives, and in the Sabikei cultivar and its derivatives (McIntosh et al., 2013). One of the reasons for the insignificant use of introgressions with leaf rust resistance genes from *T. timopheevii* is their insufficient prebreeding study, which leads to caution in their use by breeders due to fear of linkages with genes that negatively affect agronomic valuable traits.

The purpose of our research was: based on the results of studying the spring *T. aestivum/T. timopheevii* line L624, to identify its prospects for practical breeding both in terms of effectiveness against *P. triticina* and in terms of its effect on grain productivity and the quality of flour and bread.

Materials and methods

The material used included the following genotypes: 1) cultivars of spring bread wheat Saratovskaya 68 (S68), Dobrynya and standard cultivar Favorit; 2) *T. aestivum/T. timopheevii* line L624 = Saratovskaya 68/*T. timopheevii**4//Dobrynya. The cultivars Saratovskaya 68 and Dobrynya differ from each other by morphotype, presence of different *Lr*-genes, and flour and dough quality. The first cultivar is awned, red-grained, white-haired, tall, mid-ripening, susceptible to the leaf rust pathogen, contains the ineffective *Lr10* gene (Gultyaeva et al., 2020), belongs to the category of valuable wheat in terms of the quality of flour and bread.

The second cultivar is awnless, red-grained, white-haired, tall, mid-ripening, according to the quality of flour and bread, it belongs to the category of strong wheat. The cultivar Dobrynya contains the 7DS-7DL-7Ae#1L translocation from *Agropyron elongatum* (Host) Beauv. with *Lr19/Sr25* genes resistant to leaf and stem rust. These resistance genes have been overcome by pathogens in the Middle and Lower Volga, Central and Central Black Earth regions of Russia (Sibikeev et al., 1996; Baranova et al., 2021).

The standard cultivar Favorit is awnless, red-grained, white-haired, tall, and mid-ripening, according to the quality of flour and bread, it belongs to the category of valuable wheat. The cultivar is resistant to leaf rust and powdery mildew pathogens and is characterized by the substitution of the bread wheat chromosome 6D by the chromosome $6Ag^{i}$ *A. intermedium* (Host) Beauv. (Sibikeev et al., 2017).

For production of L624, a sample of *T. timopheevii* of unknown origin was provided by Dr. S.A. Stepanov (Saratov State University after name N.G. Chernyshevsky, Saratov). *T. aestivum/T. timopheevii* line L624 was obtained from direct crossing of S68 (female form) with *T. timopheevii* followed by fourfold backcrossing to the Dobrynya cultivar. Since S68 is susceptible to the leaf rust pathogen, and Dobrynya carries the overcome *Lr19* gene, resistance to *P. triticina* was the main selection criterion during backcrossing. To do this, in artificial infection with *P. triticina*, we used a population of a pathogen with a high presence of the *pp19* pathotype, virulent to *Lr19*. The stable pathogen-resistant line was isolated from the sixth generation after the last crossing.

The studies included three stages. The first stage was an evaluation of the L624 line resistance to the leaf rust pathogen in the field conditions – the phase of milky-wax ripeness (experimental field of the Agricultural Research Center for Southeast Regions, conditions of strong pathogen epiphytoty in 2017 and of medium pathogen epiphytoty in 2022) and in laboratory conditions – at the seedling phase (third leaf) in 2018–2020. The main differences in *P. triticina* populations of 2017 and 2022 were that the latter had a decreased percentage of the presence of the *pp19* pathotype virulent to the *Lr19* gene. In the field conditions, the resistance degree was evaluated according to the scale of A.P. Roelfs et al. (1992), where R is resistance, MR is moderate resistance, MS is moderate susceptibility and S is susceptibility. The degree of rust damage (%) was evaluated according to the scale of R.F. Peterson et al. (1948).

For laboratory evaluation, we used combined Saratov populations of *P. triticina* with an artificially increased presence of the *pp19* pathotype, virulent to the *Lr19* gene. Seedlings grown in pots with soil were sprayed with an aqueous suspension of population spores with the addition of Tween 80 detergent. The suspension concentration was 1 mg of inoculum per 1 ml of water. After plant infection, a dark stage was created for 20 hours with 100 % relative humidity, then seedlings were grown at a temperature of 20–22 °C, photoperiod consisted of two stages: day (16 hours) and night (8 hours).

The type of wheat reaction to the pathogen was determined according to the scale of E.B. Mains, H.S. Jackson (1926), where 0 is the absence of symptoms; 0; – necrosis without pustules; 1 – very small pustules surrounded by necrosis; 2 – pustules of medium size, surrounded by necrosis or chlorosis; 3 – pustules of medium size without necrosis; 4 – large pustules without necrosis; X – pustules on the same leaf of different types, chlorosis and necrosis are present. Plants with reaction types 0, 0; 1, 2 were considered resistant (R), and 3, 4 and X (S) were considered susceptible.

The second stage was the cytogenetic evaluation of *T. aestivum/T. timopheevii* line L624. The purpose of the cytogenetic analysis of the introgressive line was to identify alien genetic material and determine its state in the reconstructed genome of bread wheat, in the form of additional or substituted chromosomes, translocations. This made it possible to evaluate the stability of inheritance of the target trait – resistance to *P. triticina.*

Preparations of mitotic chromosomes were prepared from the meristem of seedling roots in accordance with the method of E.D. Badaeva et al. (2017). To analyze the L624 line karyotype, we used the FISH method (fluorescent *in situ* hybridization) using probes based on various repeating sequences. The probe pSc119.2 (Bedbrook et al., 1980) is localized mainly on the chromosomes of the B genome of bread wheat, while pAs1 (Rayburn, Gill, 1986) is localized mainly on the chromosomes of the D genome. Simultaneous use of these probes makes it possible to identify all chromosomes of the B and D genomes, and some chromosomes of the A genome (Schneider et al., 2003). In addition, the chromosomes of the G genome of *T. timopheevii* can be identified by the localization of hybridization signals with the pSc119.2 probe (Jiang, Gill, 1994).

The repeating sequence Spelt1 was isolated from the genomic DNA of *Ae. speltoides* Tausch. (Salina et al., 1997); the repeat blocks are localized in the subtelomeric regions of the chromosomes of this species. Individual Spelt1 sites are found in some accessions of tetra- and hexaploid wheats, in particular *T. timopheevii*, and can serve as markers of these chromosomes (Salina et al., 2006). FISH was performed according to the method of E.A. Salina et al. (2006) with minor modifications. Microscopic analysis was carried out at the Multiple-Access Centre for Microscopy of Biological Subjects (Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia).

In addition to cytogenetic analysis, molecular genetic analysis was performed. Total DNA was isolated from 5–7 day old seedlings by the method of J. Plaschke et al. (1995). Microsatellite markers were used for analysis: *Xgwm312* (Röder et al., 1998), most closely located to the *LrTt1* resistance gene on the long arm of chromosome 2A (Leonova et al., 2004); markers of the long arm of chromosome 2D – *Xgpw4480* (http://www.graingenes.org.) and *Xksum73* (Yu et al., 2004). The polymerase chain reaction (PCR) procedure was carried out according to the protocols of the authors. Separation of the amplification products was performed in a 2 % agarose gel.

The third stage was the evaluation of grain productivity traits, physical and baking properties of dough and bread in the recipient cultivars S68, Dobrynya, the standard cultivar Favorit and the *T. aestivum/T. timopheevii* line L624. The studies were carried out in 2017–2022.

The hydrothermal coefficient (HTC) for the growing season of bread wheat in 2017 was 1.0 (satisfactorily moistened

conditions), in 2018 and 2019 the HTC was 0.6 (very dry conditions), in 2020 – 0.8 (dry conditions), in 2021 – 0.9 (dry conditions), and in 2022 – 0.8 (dry conditions). The main differences in weather conditions in 2018, 2019 and 2021 were high temperatures during the flowering period (above the long-term average by 5.0, 4.2 and 8.0 °C, respectively) with a reduced precipitation amount (below the long-term average by 13–15 mm), which sharply reduced grain yield. At moderate temperatures, the most humid was the growing season of 2017, but in 2020, as in 2017, during the flowering phase, the air temperature was lower than the long-term average, namely by 0.7 and 1.0 °C, while the precipitation amount was 23 and 48 mm higher, respectively, in 2017 and 2020, which increased grain yield. The weather conditions in 2022 were characterized by an increased air temperature during the flowering phase by 1.0 °C with a precipitation amount reduced by 12 mm, but in the next ten days (beginning of July) an increase in precipitation by 16 mm was noted at a low temperature by 1.0 °C, which in general allowed to neutralize the harmful effect.

The experimental material was randomly sown in 7 m^2 plots in three replicates. The seeding rate was 400 grains per 1 m^2. The bread making and flour quality was evaluated by the content of raw gluten, gluten strength and the indicators of the IDK-3 device (deformation index of gluten) and the Chopin alveograph with the baking of experimental bread samples. The protein content of grain, harvested in 2020–2022, was determined on the grain analyzer Infratec TM 1241. The obtained data on *T. aestivum/T. timopheevii* line L624 and recipient cultivars S68, Dobrynya and the standard cultivar Favorit were subjected to one-way analysis of variance with multiple comparisons according to Duncan using the Agros 2.10 breeding and genetic software package (Martynov et al., 2000).

Results

Phytopathological analysis of resistance to the leaf rust causative agent

Under strong leaf rust epiphytotic condition in the growing season of 2017, and medium epiphytoty in 2022, line L624 showed a resistant reaction type (R) (severity 0 %, type of reaction – (IT) = 0;), while recipient cultivars: S68 – susceptibility to the pathogen (S, IT = 3+ in 2017 and S, IT = 3 in 2022) (severity 20–40 %) and Dobrynya – susceptibility to the pathogen (S, IT = 3 in 2017 and RS, IT = 0;/3 in 2022). Similar results were obtained during artificial inoculation of cultivars and line L624 at the seedling phase under greenhouse conditions (Table 1).

Thus, the phytopathological analysis of resistance to the leaf rust pathogen *T. aestivum/T. timopheevii* line L624 under field and laboratory conditions showed its high level in comparison with the recipient parental cultivars and the donor species. It should be clarified that the artificial inoculation of cultivars and L624 was carried out three times during 2018–2020 by Saratov populations of the pathogen with the addition of the *pp19* pathotype collected from the infected cultivar Dobrynya. All three evaluations (for 2018, 2019 and 2020) had the same results in terms of ITs for all studied cultivars and L624. The high efficiency of L624 resistance to *P. triticina* both at the seedlings and at milky ripeness stages indicates the juvenile nature of resistance.

Cytogenetic analysis of *T. aestivum/T. timopheevii* line L624

In general, according to the results of cytogenetic analysis, line L624 is cytologically stable and is characterized by a number of chromosomes standard for hexaploid wheat – 42 chromosomes. Chromosomal substitutions and translocations with chromosomes of the *T. timopheevii* G genome have not been identified. FISH results with pSc119.2 and pAs1 probes on L624 metaphase chromosomes are shown in Fig. 1.

Hybridization with pSc119.2 and pAs1 probes did not reveal any changes in the B and D genomes chromosomes of the L624 line, except for chromosome 2D. Chromosome 2D lacks the pAs1 signal at the end of the long arm (Schneider et al., 2003) (see Fig. 1), which may indicate a translocation of unknown origin. The performed molecular genetic analysis showed that when using microsatellite markers of the 2DL chromosome (*Xgpw4480*, *Xksum73*) (Fig. 2, *b*), the L624 line amplification fragments correspond in length to the fragments of control wheat samples. Thus, it can be concluded that the detected translocation from *T. timopheevii* in the long arm of chromosome 2D does not include the localization area of the *Xgpw4480*, *Xksum73* markers, i. e. it is subtelomeric.

One of the chromosomes of the A genome had a weak pSc119.2 signal at the end of the long arm (see Fig. 1, Fig. 3, *a*). Hybridization with pSc119.2 and Spelt1 probes (see Fig. 3) showed that this chromosome also carries the Spelt1

Table 1. Resistance of line L624, parental cultivars and standard cultivar to *P. triticina* under field conditions (natural epiphytotics) and under greenhouse conditions (artificial inoculation)

Cultivar, line	*Lr*-genes	Infection type (IT) and resistance degree		
		2017	2022	2018–2020
		Field conditions, milky ripeness phase		Greenhouse conditions, seedling phase
Saratovskaya 68	10	3+, 40S	3, 20S	3
Dobrynya	19	3, 15S	0;/3, RS*	3
L624	?	0, R	0, R	0;
Favorit	6Ag^{i}	0;, R	0;, R	0;
T. timopheevii	?	0;, R	0;, R	0;

* This infection type and resistance degree are caused by the low presence of the *pp19* pathotype in the *P. triticina* population.

repeat block at the end of the long arm (see Fig. 3, *b*). Since similar localization of these probes was previously shown on 2A^t chromosomes in some specimens of *T. araraticum* Jakubz. [syn. *T. timopheevii* (Zhuk.) Zhuk. subsp. armeniacum (Jakubz.) van Slageren)] (Salina et al., 2006), it could be assumed that the L624 line had the whole 2A^t chromosome of *T. timopheevii* or a translocation in the long arm of 2A chromosome (T2AS.2AL-2A^{t}L).

However, molecular genetic analysis using the *Xgwm312* marker (see Fig. 2, *a*) showed that the L624 amplification fragment differs in length from the *T. timopheevii* fragment and corresponds to a fragment of one of the parent wheat cultivars (Saratovskaya 68). This indicates the absence of chromosomal substitution 2A^t(2A) in line L624 and indicates a terminal translocation in the long arm of chromosome 2A (T2AS.2AL-2A^{t}L), which does not include the area of localization of the *Xgwm312* marker. In addition, there is reason to assume that the *Lr* gene(s) in L624 are not identical to the *LrTt1* gene, which is located close to the *Xgwm312* microsatellite marker.

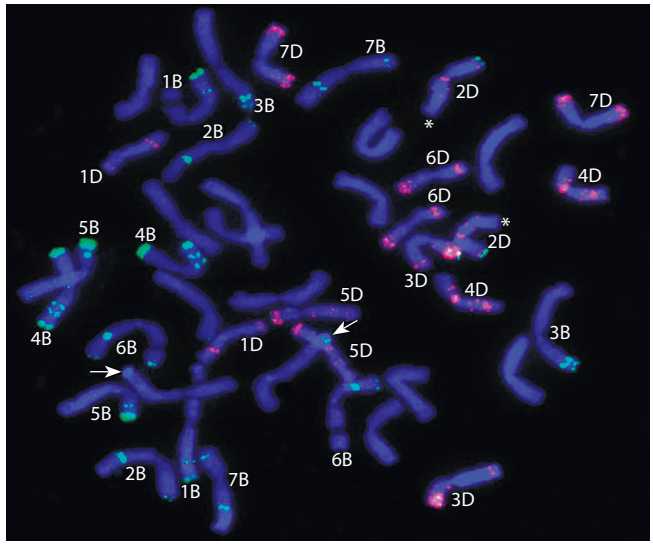

Fig. 1. FISH on metaphase chromosomes of the line of spring bread wheat L624 = Saratovskaya 68/*T. timopheevii* *4//Dobrynya with pSc119.2 probe (green signal) and pAs1 probe (red signal).

The arrows indicate the sites of hybridization with the pSc119.2 probe on the long arms of one of the A genome chromosomes. The asterisks indicate the long arms of the 2D chromosomes.

Prebreeding study of *T. aestivum/T. timopheevii* line L624

The results of the study of grain productivity in *T. aestivum/T. timopheevii* line L624, resistant to the leaf rust causative

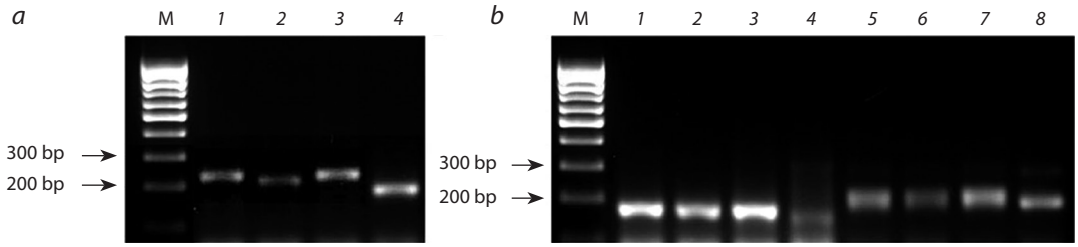

Fig. 2. Electrophoregram of the results of PCR amplification of microsatellite markers with DNA of the L624 line and parental forms.

a – Xgwm312; b – Xgpw4480 (*1–4*), *Xksum73* (*5–8*): *1, 5 – T. aestivum* (Saratovskaya 68), *2, 6 – T. aestivum* (Dobrynya), *3, 7* – line L624, *4, 8 – T. timopheevii*. M – marker of fragment length.

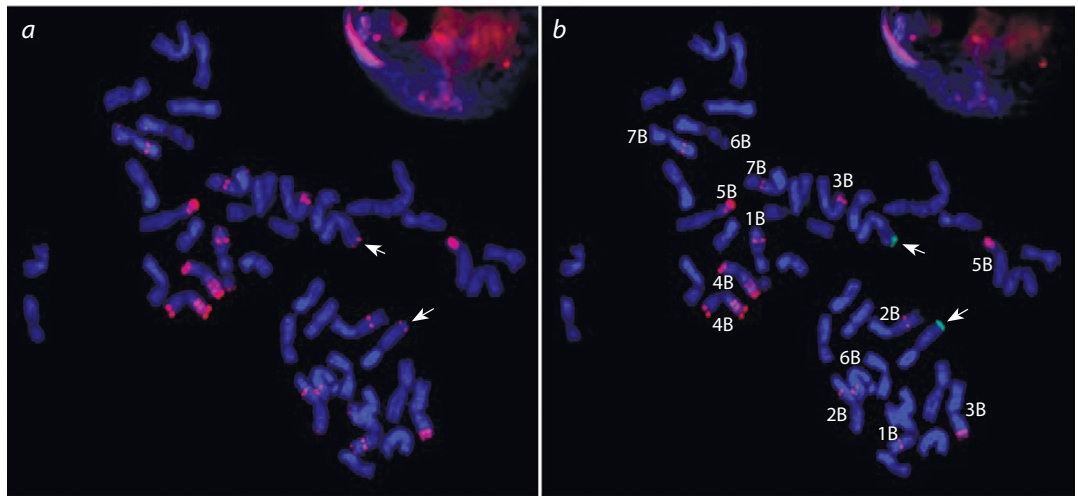

Fig. 3. FISH on metaphase chromosomes of spring bread wheat line L624 = Saratovskaya 68/*T. timopheevii* *4//Dobrynya with pSc119.2 probe (red signal) and Spelt1 probe (green signal).

Localization of pSc119.2 probe only is shown on the same metaphase plate (*a*); arrows indicate pSc119.2 signals on long arms of presumably 2A chromosomes; (*b*) arrows indicate Spelt1 signals on 2AL.

Table 2. Indicators of grain productivity in *T. aestivum/T. timopheevii* line L624

Cultivar, line	Germination-earing, days average, 2018–2022	Grain yield, kg/ha						1000 grain weight, g in 2018–2022	Protein, % in 2020–2022
		2018	2019	2020	2021	2022	Average		
S68	42.0	1427	668	3485	797	2699	1815	31.12	16.2
Dobrynya	41.6	852	670	2805	651	2829	1561	30.32	17.1
L624	44.2	1095	590	2873	759	2797	1623	27.96	16.9
Favorit (St)	43.8	861	602	2631	519	2401	1403	28.48	16.2
LSD_{05}	1.3	188	NS	240	193	343	240	2.05	0.5

agent, showed that, on average, from 2018 to 2022 there were no significant differences for grain yield in the line compared to the recipient cultivars Saratovskaya 68 and Dobrynya, as well as the standard cultivar Favorit. It is expected, since the productivity traits in 2020 and 2022 are three to five times higher than the yield in 2019 and 2021 (Table 2).

However, the analysis of grain productivity by years revealed that out of five years of study, two years of grain yield (2018 and 2020) in L624 were significantly lower than that of the recipient cultivar Saratovskaya 68 and three years (2019, 2021 and 2022) were on the same level. Compared to the cultivar Dobrynya, the grain productivity of L624 was at the same level for four years (2019–2022) and surpassed the recipient cultivar Dobrynya in 2018. Compared to the standard cultivar Favorit, the grain yield of *T. aestivum/ T. timopheevii* line L624 was higher than of Favorit for four years (2018, 2020–2022) and was at the same level only in 2019 (see Table 2).

As mentioned above, the growing seasons of 2018, 2019 and 2021 were characterized by a hard drought, while the 2020 season was characterized by excess moisture and moderate air temperature from germination to the beginning of flowering, and then a drought with high temperatures was noted until full maturity. At the same time, the 2022 season was characterized by the opposite distribution of precipitation, that is, from germination to the beginning of flowering – a lack of moisture, and during grain filling (July), an excess of moisture. However, the entire period of 2018–2022 relates to arid conditions to varying degrees. Thus, there are grounds to state that L624 has a high drought resistance and that the introgressive material from *T. timopheevii* has a neutral effect on drought resistance in this line.

On average for 2018–2022, an analysis of 1000 grain weight, as an important element of grain productivity, showed a significant decrease in the line L624 (27.96 g) compared to the cultivars Saratovskaya 68 (31.12 g) and Dobrynya (30.32 g), but a slight decrease compared to the standard cultivar Favorit (28.48 g) (see Table 2).

In terms of the "germination–earing period" trait for the growing seasons of 2018–2022, in L624, compared with recipient cultivars, there were significant differences, namely 44.2 days versus 42.0 days in S68 and 41.6 days in Dobrynya, LSD_{05} = 1.3 days. At the same time, there were no significant differences with the standard cultivar Favorit (43.8 days).

Plant height of L624 did not differ from S68, Dobrynya, and Favorit. However, lodging resistance differed between L624 (4.24 points) and Dobrynya (4.60 points), but the line did not differ from S68 (4.22 points) and Favorit (4.26 points), LSD_{05} = 0.15.

The study of the effect of chromosomes or translocations from bread wheat related species on bread making qualities in introgressive lines of bread wheat is an important step in the evaluation of their breeding value. In general, a number of studies of *T. aestivum/T. timopheevii* lines (Timonova et al., 2012), as well as lines obtained on the spring bread wheat cultivar Saratovskaya 29 using Dr. Savov's synthetic (A^tA^tGGDD), revealed a positive or neutral effect of the *T. timopheevii* genes material on the quality of flour and bread (Laikova et al., 2007, 2013).

In 2020–2022, grain protein content in *T. aestivum/ T. timopheevii* line L624 was significantly higher than in S68 and Favorit (16.9 versus 16.2 % for cultivars), but was on the same level as Dobrynya (17.1 %) (see Table 2). According to the indicators of the IDK-3 device, gluten content in line L624 was significantly lower than in S68 and Dobrynya, namely 38.7 and 39.7 % versus 36.4 % in L624, with LSD_{05} = 2.0. At the same time, the line was on the same level with the standard cultivar Favorit (37.2 %). Gluten strength in L624 was significantly higher than in the recipient cultivars and the standard cultivar Favorit, namely 78.1 u.d. against 85.8 (S68), 83.5 (Dobrynya) and 91.3 u.d. (Favorit), LSD_{05} = 4.0.

When studying the alveograph indicators (Table 3), it was found that dough tenacity, tenacity to extensibility ratio and flour strength in line L624 did not differ significantly from the recipient cultivars. However, all these traits were significantly increased in *T. aestivum/T. timopheevii* line L624 compared to the standard cultivar Favorit. Crumb porosity and bread volume in L624 were on the same level as in Dobrynya and Favorit, but significantly higher than in S68. Differences in the color of the breadcrumb were observed. Thus, in S68

Table 3. Bread making quality traits in *T. aestivum/T. timopheevii* line L624 in 2018–2022

Cultivar, line	Alveograph*			Bread**		
	P, mm	P/L	W, units	V, см^3	Porosity, score	Crumb color
Saratovskaya 68	88.8	1.75	195	761	4.50	White
Dobrynya	84.4	1.68	204	814	4.88	Yellow
L624	84.7	1.51	203	799	4.85	White
Favorit	69.1	1.20	170	780	4.75	Cream
LSD_{05}	8.0	0.3	25	30	0.26	

* Indicators of the alveograph: P – dough tenacity; P/L – tenacity to extensibility ratio; W – flour strength.
** Indicators of bread evaluation: V – bread volume, porosity.

and L624, the crumb is white, in Favorit, it is cream, and in Dobrynya, it is yellow, which is associated with the presence of 7DS-7DL-7Ae#1L, a translocation that carries genes for high carotenoid content (Prins et al., 1996) (see Table 3).

Discussion

As noted above, the *T. timopheevii* species attracts breeders with its high immunity to a complex of fungal diseases. A set of *T. aestivum/T. timopheevii* lines with resistance genes to various diseases was obtained: to the leaf rust pathogen – *Lr18*, *Lr50*, *LrTt1*, *LrTt2*, *LrSelG12*; to the causative agent of stem rust (*P. graminis* Pers.) – *Sr36*, *Sr37*, *Sr40*; to powdery mildew pathogen (*Blumeria graminis* DC.) – *Pm6*, *Pm27* (Adonina et al., 2021). The *T. aestivum/T. timopheevii* lines resistant to pathogens mainly carry introgressive material from chromosomes 6G, 2G, 5G, in addition, from chromosomes $1A^t$, $2A^t$, $3A^t$, $5A^t$, $7A^t$, 3G, 4G, 7G (Badaeva et al., 2010).

In our studies, introgressions from *T. timopheevii* affected chromosome 2A (T2AS.2AL-$2A^t$L – translocation), and chromosome 2D lacks the pAs1 signal at the end of the long arm, which may indicate a translocation of unknown origin. Previously, in the long arm of chromosome 2A (translocation 2AS-$2A^t$S.$2A^t$L-2AL) of line 842 = Saratovskaya 29 *2/*T. timopheevii* spp. *viticulosum* between *Xgwm817* and *Xgwm312* markers, the *LrTt1* gene was mapped (Leonova et al., 2004, 2008, 2011). This gene was found to be inherited in a recessively monogenic manner (Leonova et al., 2004). In L624, the leaf rust resistance gene (according to segregation in the F_2 families Saratovskaya 68/*T. timopheevii* *4//Dobrynya) is also inherited in a recessively monogenic manner (Sibikeev, unpublished data).

However, in our studies, using PCR analysis with the DNA marker of the *LrTt1* gene (*Xgwm312*), it was established that it is not identical to the *Lr* gene(s) in L624. In addition, *LrTt1* is effective against the European population of *P. triticina* at the seedling stage, but only has a restraining effect (infection type IT = 3) at the stage of adult plants against the West Siberian population of the pathogen (Leonova et al., 2004; Timonova et al., 2012). In addition, translocation with *LrTt1* is not linked to resistance to the stem rust causative agent in Western Siberia, the share of damage is 80 % (Timonova et al., 2012).

Introgressions from *T. timopheevii* in L624 protected against the Saratov population of *P. triticina*, both at the seedlings stage (three-leaf phase) and at the stage of adult plants (phase of milky ripeness), infection type IT = 0;. Line L624 was affected by the stem rust causative agent at the seedlings stage during artificial inoculation of both the Saratov and Omsk populations, IT = 3.

When trying to identify in L624 genes of resistance to *P. graminis* (*Sr*-genes): *Sr2*, *22*, *24*, *25*, *26*, *31*, *32*, *35*, *36*, *38*, *39*, *57* using DNA markers, none of the indicated resistance genes was identified (Baranova, unpublished data). At the same time, earlier in our studies, the effectiveness of L624 (in these studies, L624 is designated by serial number 49) was shown at the seedling stage against the Saratov, Krasnodar, Dagestan and Chelyabinsk populations of *P. triticina* collected in 2018, as well as against test clones of the pathogen with virulence to the leaf rust resistance genes – *Lr9*, *Lr19*, *Lr26*, the infection type to the pathogen was IT = 0.

In addition, the use of DNA markers for genes *Lr1*, *3*, *9*, *10*, *19*, *20*, *21*, *22a*, *24*, *25*, *28*, *29*, *34*, *35*, *37*, *39 = 41*, *47*, *50*, *53*, *66*, *6Agi* made it possible to identify the *Lr10* and *Lr28* genes in L624. The *Lr19* gene from the Dobrynya cultivar was not identified. The detection of the SCS421 gene marker for *Lr28* indicates the presence of *T. timopheevii* (*LrTtim*) genetic material in the L624 since this marker is not strictly specific for determining the *Lr28* gene transferred from *Ae. speltoides*, and is also present in specimens of *T. timopheevii* (Gultyaeva et al., 2014). Among the *Lr* genes from *T. timopheevii*, the *Lr18* gene belongs to the group of ineffective in the conditions of the Volga region and has a different infection type to the pathogen (IT = 3) from IT = 0 in L624.

The infection type of the line with *Lr50* (the second *Lr* gene from *T. timopheevii*) at inoculation with the Saratov population of the leaf rust pathogen varied from 0–1 to 2+ points and differed from that in *L624*. The *Xgwm382* marker of the *Lr50* gene indicated the absence of *Lr50* in this line (Gultyaeva et al., 2020). Thus, there are grounds to suggest that L624 in the T2AS.2AL-$2A^t$L translocation contains a leaf rust resistance

gene that is different from *Lr18*, *Lr50* and, possibly, non-allelic to the *LrTt1* gene.

Analyzing the effect of introgressive material with pathogen resistance genes from *T. timopheevii* on agronomically important traits, its multidirectional influence should be noted (Leonova, 2018). However, there are very few data on the specific effect of introgressions involving the $2A^t$ chromosome on productivity indicators.

There is a study of *T. aestivum/T. timopheevii* lines 5352-104 and 5360-191/5, obtained by three backcrosses of hybrid parental lines 744 and 832 (*T. aestivum*−*T. timopheevii*, $2n = 42$) with Saratovskaya 29 and subsequent self-pollination (Timonova et al., 2012). Line 5352-104 contains introgressive fragments of chromosomes $1A^t$ and $2A^t$, while line 5360-191/5 contains $2A^t$ and 5GL. In terms of plant height, these lines did not differ from Saratovskaya 29 (S29), and the ear length of the line 5352-104 was higher than that of the S29 cultivar.

According to the number of spikelets per spike, no differences from S29 were found, but in terms of the number and weight of grains per spike, the line's indicators were higher than those of the recipient cultivar. In terms of grain weight per plant and of 1000 grain weight, line 5352-104 did not differ from S29. Line L 5360-191/5, as a possible carrier of the *LrTt1*($2A^t$) and *LrTt2* (5GL) genes, did not differ from the recipient cultivar S29 in all parameters of ear productivity, as well as grain weight per plant and 1000 grain weight and plant height (Timonova et al., 2012).

In our studies, introgressions in L624 affected chromosomes 2A and 2D, resulting in highly effective leaf rust resistance. Of the five years of studying L624, in 2018 and 2020, the grain yield was significantly lower than that of the recipient cultivar Saratovskaya 68, in 2019, 2021 and 2022, it was on the same level. Compared to Dobrynya, the grain productivity of L624 was at the same level for four years (2019–2022) and surpassed the recipient cultivar Dobrynya in 2018. Compared to the standard cultivar Favorit, the grain yield of *T. aestivum/T. timopheevii* line L624 surpassed Favorit for four years (2018, 2020–2022) and was on the same level only in 2019. In general, it can be assumed that L624 does not reduce grain productivity. However, by the weight of 1000 grains, L624 was inferior to both recipients and was on the same level with the standard cultivar Favorit. L624 differs from lines L5352-104 and 5360-191/5 in terms of its effect on this trait.

Unfortunately, the results of other researchers on the study of the effect of introgressions involving the $2A^t$ chromosome from *T. timopheevii* on the bread making quality traits is not known to us. It can be stated that introgressions on the 2A and 2D chromosomes in L624 increased the grain protein content compared to the recipient cultivar S68 and the standard cultivar Favorit, but this trait remained at the same level as Dobrynya. In terms of flour and bread quality, L624 was not inferior to cultivar recipients; moreover, it surpassed S68 in terms of bread volume and porosity. At the same time, L624 surpassed the standard cultivar Favorit in all parameters of the alveograph: dough tenacity, tenacity to extensibility ratio and flour strength. Thus, introgressions in chromosomes 2A and 2D in L624 do not impair baking properties. However, the decrease in gluten content in comparison with the recipient cultivar should be noted. For comparison, introgressions from *T. timopheevii* in the cultivar Pamyati Maystrenko – 2B(2G), 6B(6G) and $1D(1D^t)$ increased protein and gluten content, as well as traits of flour strength and bread volume (Laikova et al., 2013).

The 7DS-7DL-7Ae#1L translocation from *Ag. elongatum* with *Lr19/Sr25* genes linked to yellow flour was lost in L624 = Saratovskaya 68/*T. timopheevii* *4//Dobrynya, despite four backcrosses for the Dobrynya cultivar. As a result, L624 has a white color of flour and bread. The absence of the 7DS-7DL-7Ae#1L translocation in L624 was confirmed using a DNA marker for *Lr19*/*Sr25*, SCS265 (Gultyaeva et al., 2020). A possible reason for this is that during the L624 breeding, a constant selection of resistant plants was carried out under the background of inoculation with a population of *P. triticina* with a high presence of the *pp19* pathotype virulent to *Lr19*.

Conclusion

In general, regarding the entire complex of agronomic traits, *T. aestivum/T. timopheevii* line L624 is promising, both in comparison with the recipient cultivars and in comparison with the standard cultivar Favorit. L624 did not reduce grain yield during five years of study, which is apparently due to a fairly high level of drought resistance. The leaf rust resistance gene in L624 is highly effective both at the seedling stage and at the stage of adult plants. For further use of L624 in breeding programs, it is necessary to carry out additional studies on the combination of resistance to *P. triticina* with resistance genes to the stem rust causative agent, which has become necessary for the Lower Volga region.

References

Adonina I.G., Timonova E.M., Salina E.A. Introgressive hybridization of common wheat: results and prospects. *Russ. J. Genet*. 2021; 57(4):390-407. DOI 10.1134/S1022795421030029.

Allard R.W., Shands R.G. Inheritance of resistance to stem rust and powdery mildew in cytologically stable spring wheats derived from *Triticum timopheevii*. *Phytopathology*. 1954;44:266-274.

Badaeva E.D., Budashkina E.B., Bilinskaya E.N., Pukhalskiy V.A. Intergenomic chromosome substitutions in wheat interspecific hybrids and their use in the development of a genetic nomenclature of *Triticum timopheevii* chromosomes. *Russ. J. Genet*. 2010;46(7): 769-785. DOI 10.1134/S102279541007001X.

Badaeva E.D., Ruban A.S., Aliyeva-Schnorr L., Municio C., Hesse S., Houben A. *In situ* hybridization to plant chromosomes. In: Liehr T. (Ed.). Fluorescence in Situ Hybridization (FISH). Berlin; Heidelberg, Germany: Springer, 2017;477-494. DOI 10.1007/978-3-662-52959-1_49.

Baranova O.A., Sibikeev S.N., Druzhin A.E., Sozina I.D. Loss of effectiveness of stem rust resistance genes *Sr25* and *Sr6Agi* in the Lower Volga region. *Vestnik Zashchity Rasteniy = Plant Protection News.* 2021;104(2):105-112. DOI 10.31993/2308-6459-2021-104-2-14994. (in Russian)

Bedbrook R.J., Jones J., O'Dell M., Thompson R.J., Flavell R.B. A molecular description of telomeric heterochromatin in *Secale* species. *Cell*. 1980;19(2):545-560. DOI 10.1016/0092-8674(80)90529-2.

Brown-Guedira G.L., Singh S., Fritz A.K. Performance and mapping of leaf rust resistance transferred to wheat from *Triticum timopheevii* subsp. *armeniacum*. *Phytopathology*. 2003;93(7):784-789. DOI 10.1094/PHYTO.2003.93.7.784.

Budashkina E.B., Kalinina N.P. Development and genetic analysis of common wheat introgressive lines resistant to leaf rust. *Acta Phytopathol. Entomol. Hung*. 2001;36(1-2):61-65. DOI 10.1556/aphyt.36.2001.1-2.8.

Gultyaeva E.I., Orina A.S., Gannibal P.B., Mitrofanova O.P., Odintsova I.G., Laikova L.I. The effectiveness of molecular markers for the identication of *Lr28*, *Lr35* and *Lr47* genes in common wheat. *Russ. J. Genet*. 2014;50(2):131-139. DOI 10.7868/S00166758140 20064.

Gultyaeva E.I., Sibikeev S.N., Druzhin A.E., Shaydayuk E.L. Enlargement of genetic diversity of spring bread wheat resistance to leaf rust (*Puccinia triticina* Erics.) in Lower Volga region. *Selskokhozyaystvennaya Biologiya = Agricultural Biology*. 2020;55(1): 27-44. DOI 10.15389/agrobiology.2020.1.27rus. (in Russian)

Gultyaeva E.I., Shaydayuk E.L., Gannibal P.B. Leaf rust resistance genes in wheat cultivars registered in Russia and their influence on adaptation processes in pathogen populations. *Agriculture*. 2021; 11(4):319. DOI 10.3390/agriculture11040319.

Dorofeev V.F., Filatenko A.A., Migushova E.F., Udachin R.A., Yakubtsiner M.M. Cultural Flora of the USSR. Vol. 1. Wheat. Leningrad, 1979. (in Russian)

Jiang J., Gill B.S. Different species-specific chromosome translocation in *Triticum timopheevii* and *T. turgidum* support diphyletic origin of polyploid wheats. *Chromosome Res*. 1994;2(1):59-64. DOI 10.1007/ BF01539455.

Jørgensen J.H., Jensen C.J. Genes for resistance to wheat powdery mildew in derivatives of *Triticum timopheevii* and *T. carthlicum. Euphytica*. 1972;21:121-128. DOI 10.1007/BF00040557.

Kozlovskaya V.F., Grigoryeva L.P., Shatilova N.V. Using interspecific hybridization for the development of new sources for stem rust resistance of wheat. *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology.* 1990;1:65-71. (in Russian)

Laikova L.I., Arbuzova V.S., Efremova T.T., Popova O.M., Ermakova M.F. Estimation productivity and quality of the grain in immune introgressive lines of soft wheat of Saratovskaya 29 variety. *Selskokhozyaystvennaya Biologiya = Agricultural Biology*. 2007;42(5):75-85. (in Russian)

Laikova L.I., Belan I.A., Badaeva E.D., Rosseeva L.P., Shepelev S.S., Shumny V.K., Pershina L.A. Development and study of spring bread wheat variety Pamyati Maystrenko with introgression of genetic material fron syntetic hexaploid *Triticum timopheevii* Zhuk. × *Aegilops tauschii* Coss. *Russ. J. Genet.* 2013;49(1):89-97. DOI 10.1134/ S1022795413010067.

Leonova I.N. Influence of alien genetic material on the manifestation of agronomically important traits of common wheat (*T. aestivum* L.). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2018;22(3):321-328. DOI 10.18699/VJ18.367. (in Russian)

Leonova I.N., Borner A., Budashkina E.B., Kalinina N.P., Unger O., Roder M.S., Salina E.A. Identification of microsatellite markers for a leaf rust resistance gene introgressed into common wheat from *Triticum timopheevii. Plant Breed*. 2004;123(1):93-95. DOI 10.1046/j.0179-9541.2003.00906.x.

Leonova I.N., Laikova L.I., Popova O.M., Unger O., Borner A., Roder M.S. Detection of quantitative trait loci for leaf rust resistance in wheat – *T. timopheevii/T. tauschii* introgression lines. *Euphytica*. 2007;155:79-87. DOI 10.1007/s10681-006-9303-4.

Leonova I.N., Kalinina N.P., Budashkina E.B., Röder M.S. Genetic analysis and localization of loci controlling leaf rust resistance of *Triticum aestivum* × *Triticum timopheevii* introgression lines. *Russ. J. Genet*. 2008;44(12):1431-1437. DOI 10.1134/S1022 795408120077.

Leonova I.N., Budashkina E.B, Flath K., Weidner A., Börner A., Röder M.S. Microsatellite mapping of a leaf rust resistance gene transferred to common wheat from *Triticum timopheevii. Cereal Res. Commun.* 2010;38(2):212-220. DOI 10.1556/CRC.38.2010.2.7.

Leonova I.N., Budashkina E.B., Kalinina N.P., Röder M.S., Börner A., Salina E.A. *Triticum aestivum–Triticum timopheevii* introgression lines as a source of pathogen resistance genes. *Czech J. Genet. Plant Breed*. 2011;47:S49-S55.

Mains E.B., Jackson H.S. Physiological specialization in the leaf rust of wheat *Puccinia triticina* Erikss. *Phytopathology*. 1926;16:89-120.

Martynov S.P., Musin N.N., Kulagina T.V. The Agros software package for statistical biometric and genetic analysis, version 2.10. Russian Academy of Agricultural Sciences. Department of Statistical Analysis, 2000. (in Russian)

McIntosh R.A., Wellings C.R., Park R.F. Wheat Rusts. An Atlas of Resistance Genes. Springer Dordrecht, 1995.

McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers W.J., Morris C., Appels R., Xia X.C. Catalogue of gene symbols for wheat. In: Proceedings of the 12th International Wheat Genetics Symposium; 8-13 September 2013; Yokohama, Japan. Springer Open, 2013. https://host170.sedici.unlp.edu.ar/server/api/core/bitstreams/bc9e 00ea-92ac-4d4a-8df7-a9f51111a8e0/content.

McIntosh R.A., Dubcovsky J., Rogers W.J., Xia X.C., Raupp W.J. Catalogue of Gene Symbols for Wheat: 2018 Supplement. *Annu. Wheat Newsl.* 2018;64:73-93.

McIntosh R.A., Dubcovsky J., Rogers W.J., Xia X.C., Raupp W.J. Catalogue of Gene Symbols for Wheat: 2022 Supplement. *Annu. Wheat Newsl.* 2022;68:68-82.

Peterson R.F., Campbell A.B., Hannah A.E. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Can. J. Res*. 1948;26(5):496-500. DOI 10.1139/cjr48c-033.

Plaschke J., Ganal M.W., Röder M.S. Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theor. Appl. Genet*. 1995;91(6-7):1001-1007. DOI 10.1007/BF00 223912.

Prins R., Marais G.F., Janse B.J.H., Pretorius Z.A., Marais A.S. A physical map of the *Thinopyrum*-derived *Lr19* translocation. *Genome*. 1996;39(5):1013-1019. DOI 10.1139/g96-126.

Rayburn A.L., Gill B.S. Isolation of a *D*-genome specific repeated DNA sequence from *Aegilops squarrosa*. *Plant Mol. Biol. Report*. 1986;4: 104-109. DOI 10.1007/BF02732107.

Roelfs A.P., Singh R.P., Saaru E.E. Rust Diseases of Wheat: Concepts and Methods of Disease Management. Mexico, D.F.: CIMMYT, 1992.

Röder M.S., Korzun V., Wendehake K., Plaschke J., Tixier M.H., Leroy P., Ganal M.W. A microsatellite map of wheat. *Genetics.* 1998; 149(4):2007-2023. DOI 10.1093/genetics/149.4.2007.

Salina E.A., Lim Y.K., Badaeva E.D., Scherban A.B., Adonina I.G., Amosova A.V., Samatadze T.E., Vatolina T.Y., Zoshchuk S.A., Leitch A. Phylogenetic reconstruction of *Aegilops* section *Sitopsis* and the evolution of tandem repeats in the diploids and derived wheat polyploids. *Genome*. 2006;49(8):1023-1035. DOI 10.1139/ g06-050.

Salina E.A., Pestsova E.G., Vershinin A.V. "Spelt1" a novel family of tandemly repeated DNA in cereals. *Russ. J. Genet*. 1997;33(4): 352-357.

Schneider A., Linc G., Molnar-Lang M. Fluorescence *in situ* hybridization polymorphism using two repetitive DNA clones in different cultivars of wheat. *Plant Breed.* 2003;122(5):396-400. DOI 10.1046/j.1439-0523.2003.00891.x.

Sibikeev S.N., Krupnov V.A., Voronina S.A., Elesin V.A. First report of leaf rust pathotypes virulent to highly effective *Lr*-genes transferred from *Agropyron* species to bread wheat. *Plant Breed*. 1996;115(4): 276-278. DOI 10.1111/j.1439-0523.1996.tb00917.x.

Sibikeev S.N., Druzhin A.E., Badaeva E.D., Shishkina A.A., Dragovich A.Y., Gultyaeva E.I., Kroupin P.Y., Karlov G.I., Khuat T.M., Divashuk M.G. Comparative analysis of *Agropyron intermedium* (Host) Beauv $6Ag^{i}$ and $6Ag^{i}2$ chromosomes in bread wheat cultivars and lines with wheat–wheatgrass substitutions. *Russ. J. Genet*. 2017;53(3):314-324. DOI 10.1134/S1022795417030115.

Sibikeev S.N., Konkova E.A., Salmova M.F. Characteristic of the bread wheat leaf rust pathogen virulence in the Saratov region conditions. *Agrarnyy Nautchnyy Zhurnal = Agrarian Scientific Journal*. 2020;9:40-44. DOI 10.28983/asj.y2020i9pp40-44. (in Russian)

Singh A.K., Sharma J.B., Vinod, Singh P.K., Singh A., Mallick N. Genetics and mapping of a new leaf rust resistance gene in *Triticum aestivum* L. × *Triticum timopheevii* Zhuk. derivative 'Selection G12'. *J. Genet*. 2017;96:291-297. DOI 10.1007/s12041-017-0760-4.

Skurygina N.A. Highly effective genes for resistance to the population of leaf rust and powdery mildew in bread wheat lines derived from *T. timopheevii* Zhuk., and their identification. *Trudy po Prikladnoy Botanike, Genetike i Selektsii = Proceedings on Applied Botany, Genetics, and Breeding*. 1984;85:5-13. (in Russian)

Timonova E.M., Leonova I.N., Belan I.A., Rosseeva L.P., Salina E.A. Effect of certain chromosome regions of *Triticum timopheevii* on the formation on pest resistance and quantative traits in common wheat. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2012;16(1):142-159. (in Russian)

Tomar S.M.S., Joshi B.C., Kochumadhavan M., Shrivastava K.D. Transfer of leaf rust resistance into bred wheat from *Triticum timopheevii* Zhuk. *Current Sci.* 1988;57(1):17-19.

Yu J.-K., Dake T.M., Singh S., Benscher D., Li W., Gill B., Sorrells M.E. Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. *Genome*. 2004;47(5):805-818. DOI 10.1139/G04-057.

Zhirov E.G., Ternovskaya T.K. Genome engineering in wheat. *Vestnik Selskokhozyaystvennoy Nauki = Herald of Agricultural Sciences*.1984;10:58-66. (in Russian)

ORCID ID

S.N. Sibikeev orcid.org/0000-0001-8324-9765
I.G. Adonina orcid.org/0000-0002-8460-6119
A.E. Druzhin orcid.org/0000-0002-3968-2470
O.A. Baranova orcid.org/0000-0001-9439-2102

Acknowledgements. Cytogenetic analysis was supported by the project of the Ministry of Science and Higher Education FWNR-2022-0017.

Conflict of interest. The authors declare no conflict of interest.

Received March 28, 2023. Revised May 2, 2023. Accepted May 12, 2023.

ANIMAL GENETICS
Original article

Вавиловский журнал генетики и селекции. 2023;27(6):633-640
DOI 10.18699/VJGB-23-74

In vivo MRS study of long-term effects of traumatic intracranial injection of a culture medium in mice

O.B. Shevelev[1, 2] ✉, O.P. Cherkasova[3, 4], I.A. Razumov[1, 5], E.L. Zavjalov[1]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Institute "International Tomografic Center" of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Institute of Laser Physics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[4] Novosibirsk State Technical University, Novosibirsk, Russia
[5] Novosibirsk State University, Novosibirsk, Russia
✉ shevelev.oleg.nsk@gmail.com

Abstract. Orthotopic transplantation of glioblastoma cells in the brain of laboratory mice is a common animal model for studying brain tumors. It was shown that ^{1}H magnetic resonance spectroscopy (MRS) enables monitoring of the tumor's occurrence and its development during therapy based on the ratio of several metabolites. However, in studying new approaches to the therapy of glioblastoma in the model of orthotopic xenotransplantation of glioma cells into the brain of mice, it is necessary to understand which metabolites are produced by a growing tumor and which are the result of tumor cells injection along the modeling of the pathology. Currently, there are no data on the dynamic metabolic processes in the brain that occur after the introduction of glioblastoma cells into the brain of mice. In addition, there is a lack of data on the delayed effects of invasive brain damage. Therefore, this study investigates the long-term dynamics of the neurometabolic profile, assessed using ^{1}H MRS, after intracranial injection of a culture medium used in orthotopic modeling of glioma in mice. Levels of N-acetylaspartate, N-acetylaspartylglutamic acid, myoinositol, taurine, glutathione, the sum of glycerophosphocholine and phosphocholine, glutamic acid (Glu), glutamine (Gln), and gamma aminobutyric acid (GABA) indicate patterns of neurometabolites in the early stage after intracranial injection similar to brain trauma ones. Most of the metabolites, with the exception of Gln, Glu and GABA, returned to their original values on day 28 after injection. A progressive increase in the Glu/Gln and Glu/GABA ratio up to 28 days after surgery potentially indicates an impaired turnover of these metabolites or increased neurotransmission. Thus, the data indicate that the recovery processes are largely completed on day 28 after the traumatic event in the brain tissue, leaving open the question of the neurotransmitter system impairment. Consequently, when using animal models of human glioma, researchers should clearly distinguish between which changes in neurometabolites are a response to the injection of cancer cells into the brain, and which processes may indicate the early development of a brain tumor. It is important to keep this in mind when modeling human glioblastoma in mice and monitoring new treatments. In addition, these results may be important in the development of approaches for non-invasive diagnostics of traumatic brain injury as well as recovery and rehabilitation processes of patients after certain brain surgeries.
Key words: magnetic resonance spectroscopy; animal model of human glioma; neurometabolites; traumatic brain injury.

For citation: Shevelev O.B., Cherkasova O.P., Razumov I.A., Zavjalov E.L. *In vivo* MRS study of long-term effects of traumatic intracranial injection of a culture medium in mice. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):633-640. DOI 10.18699/VJGB-23-74

Прижизненное МРС исследование долгосрочных последствий травматической внутричерепной инъекции культуральной среды у мышей

О.Б. Шевелев[1, 2] ✉, О.П. Черкасова[3, 4], И.А. Разумов[1, 5], Е.Л. Завьялов[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт «Международный томографический центр» Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Институт лазерной физики Сибирского отделения Российской академии наук, Новосибирск, Россия
[4] Новосибирский государственный технический университет, Новосибирск, Россия
[5] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ shevelev.oleg.nsk@gmail.com

Аннотация. Ортотопическая ксенотрансплантация клеток глиобластомы в головной мозг лабораторных мышей – распространенная животная модель для изучения опухолей головного мозга. Показано, что ^{1}H магнитно-резонансная спектроскопия (МРС) позволяет отслеживать возникновение опухоли и ее развитие в процессе терапии по соотношению нескольких метаболитов. Однако при изучении новых подходов в терапии глиобластомы на модели ортотопической ксенотрансплантации клеток глиомы в головной мозг мышей необходимо понимать, какие изменения уровней метаболитов являются следствием роста опухоли, а какие – результатом инъекции опухоле-

вых клеток в головной мозг в процессе моделирования патологии. В настоящее время отсутствуют данные о динамике метаболических процессов в головном мозге, возникающих после введения клеток глиобластомы в мозг мышей. Мало также данных об отсроченных последствиях инвазивного повреждения головного мозга. Поэтому в нашей работе исследуется долговременная динамика нейрометаболического профиля, оцененного с применением ^{1}H МРС, после внутричерепной инъекции культуральной среды, используемой при ортотопическом моделировании глиомы у мышей. Уровни N-ацетиласпартата, N-ацетиласпартилглутаминовой кислоты, мио-инозитола, таурина, глутатиона, суммы глицерофосфохолина и фосфохолина, глутаминовой кислоты (Glu), глутамина (Gln) и гамма-аминомасляной кислоты (ГАМК) указывают на паттерны нейрометаболитов на ранней стадии после внутричерепной инъекции, схожие с таковыми при травме головного мозга. Большинство метаболитов, за исключением Gln, Glu и ГАМК, возвращались к исходным значениям на 28-й день после инъекции. Прогрессирующее увеличение соотношения Glu/Gln и Glu/GABA до 28 дней после операции потенциально указывает на нарушение обмена этих метаболитов или усиление нейропередачи. Таким образом, данные свидетельствуют о том, что восстановительные процессы в основном завершаются на 28-й день после травматического события в ткани головного мозга, оставляя открытым вопрос о нарушении нейромедиаторной системы. Соответственно, при использовании животных моделей глиомы человека исследователи должны четко различать, какие изменения нейрометаболитов являются реакцией на саму инъекцию раковых клеток в головной мозг, а какие процессы могут свидетельствовать о раннем развитии опухоли головного мозга. Это важно иметь в виду при моделировании глиобластомы человека у мышей и мониторинге новых методов лечения. Кроме того, полученные данные могут быть важны при разработке подходов к неинвазивной диагностике черепно-мозговой травмы, а также при мониторинге процессов восстановления и реабилитации пациентов после некоторых операций на головном мозге.
Ключевые слова: магнитно-резонансная спектроскопия; животная модель глиобластомы человека; нейрометаболиты; черепно-мозговая травма.

Introduction

Glioblastoma (GBM) is the most common and aggressive tumor of the central nervous system (Goodenberger, Jenkins, 2012). Even in the case of aggressive therapy of the brain glioma, such as surgical resection, radiotherapy, and chemotherapy, many types of gliomas almost always have a pessimistic prognosis for the patients' survival (Tykocki et al., 2018; Ostrom et al., 2019). Despite the efforts of scientists and clinicians to increase the life expectancy of GBM patients, survivors do not easily exceed the 15th month (3–5) and the 5-year survival rate is as low as 5.8 % (Ostrom et al., 2018, 2019; Tan et al., 2020).

Xenograft mice models are widely used for modeling GBM and testing developmental therapeutics (Haddad et al., 2021). Orthotopic transplantation of glioblastoma cells in the brain of laboratory mice is a common animal model for studying brain tumors (Zavjalov et al., 2016; Miyai et al., 2017; Haddad et al., 2021). This model is characterized by rapid glioma growth following injection of tumor cells into the subcortical structures of the brain, resulting in destruction of blood vessels and neurons, compression of individual brain structures, impairment of cognitive function and deep irreversible lesions (Hall et al., 2005; Maas et al., 2008). Currently, the most accessible way to detect and visualize glioma is magnetic resonance imaging (MRI), which is widely used in the clinic. The main advantage of MRI is its non-invasiveness with the ability to image unlabeled cells, while optical methods require stable expression of fluorophore-labeled protein in tumor cells, which can lead to a change in the behavior of tumor cells (Dass, Choong, 2007). Finally, MRI is a reliable and sufficiently accurate method that can be used to repeatedly study the dynamic processes for the same organism.

To visualize tumors, reliable MRI contrast agents are used, which, however, do not always allow a correct diagnosis. Up to 45 % of abnormalities in a patient's brain detectable with MRI require biopsy and subsequent histological studies. There is another significant drawback of this approach for intracerebral tumors. Contrast does not always lead to an increase in the MRI signal, and although this happens quite rarely – about 4 % of gliomas, the life of the patient is on the line in every such case (Barker et al., 1997; Ginsberg et al., 1998; Sugahara et al., 1998; Castillo et al., 2001; Nelson, Cha, 2003; Atkinson et al., 2008; Stockhammer et al., 2008).

Recently, *in vivo* studies of the functional state of the brain are increasing due to attempts to assess the metabolic profile using ^{1}H magnetic resonance spectroscopy – MRS (Ratai, Gilberto González, 2016). It was shown that MRS enables monitoring of the tumor's occurrence and its development during therapy based on the ratio of several metabolites (Hishii et al., 2019; Tiwari et al., 2020). However, at study of new approaches in the therapy of glioblastoma in the model of orthotopic xenotransplantation of glioma cells into the brain of mice, it is necessary to understand which metabolites are produced by a growing tumor and which are the result of tumor cells injection along the modeling of the pathology. Currently, there are no data on the dynamic metabolic processes in the brain that occur after the introduction of glioblastoma cells into the brain of mice. Furthermore, there is a lack of data on the delayed effects of invasive brain damage (Singh et al., 2016; Li Y. et al., 2021). These studies require multiple animal MRI sessions under anesthesia, and so this study investigates the long-term dynamics of the neurometabolic profile by ^{1}H MRS repeated sessions after intracranial injection of a culture medium used at orthotopic modeling of glioma in mice.

Materials and methods

The study was carried out in the Center for Collective Use "SPF-vivarium" of the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences (RFMEFI62119X0023). This study was approved by the Inter-Institutional Commission on Biological Ethics at the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences (Permission #78, April 16, 2021). Ten SPF male SCID mice aged 6–7 weeks were used. Their health status was investigated in accordance with the recommendations of the Federation of European Laboratory

Animal Science Associations (FELASA working group…, 2014). The animals were kept in single-sex family groups of 2–5 mice in individually ventilated cages using the Optimice system (Animal Care Systems, Centennial, CO, USA). The mice were maintained under controlled conditions: temperature, 22–26 °C; relative humidity, 30–50 %; and 12/12 light/dark periods with dawn at 02:00. Standart V1534-300 ssniff® diet (ssniff Spezialdiäten GmbH, Soest, Germany) and reverse osmosis water enriched with minerals were provided to the animals *ad libitum*.

Before the operation, the animal was placed in a chamber with an air flow of 300–350 ml/min and an isoflurane concentration of 1.5 % (Baxter International Inc., Deerfield, IL, USA). After 3–5 min, the animal was transferred to a heated operating table with a surface temperature of 37 °C and placed under an anesthetic mask with 1.5 % isoflurane. The culture medium (DMEM\F-12, Thermo Fisher Scientific, Waltham, MA, USA) used for orthotopic glioma modeling was introduced into the subcortical structures of the brain through a hole in the cranium (Zavjalov et al., 2016). A 3–4 mm skin incision was made on the animal's head in the caudal-cranial direction in the region of the bregma, and 5 μl of medium was injected through the hole in the skull. Intravital MRS was performed on the anesthetized animals before surgery and on days 7, 14, 21 and 28 after the injection.

Magnetic resonance imaging and MRS protocols. Neurometabolite analyses were performed on a horizontal tomograph with a magnetic field strength of 11.7 Tesla (Biospec 117/16; Bruker, Billerica, MA, USA). Five minutes prior to the analysis, animals were immobilized with gas anesthesia (Isofluran, Baxter International Inc.) using an anesthesia machine (The Univentor 400 Anaesthesia Unit; Univentor Ltd., Zejtun, Malta). The body temperature of the animals was maintained using a water circuit in a tomographic bed-table with a surface temperature of 30 °C. A pneumatic breathing sensor (SA Instruments Inc., Stony Brook, NY, USA) was placed under the lower torso, which made it possible to control the depth of anesthesia.

All proton spectra of the mouse brain were obtained using a ^{1}H volumetric radiofrequency coil (T11440V3). Correct positioning of spectroscopic voxels ($2.5 \times 2.5 \times 2.5$ mm^3) was performed using the RARE method (rapid acquisition with relaxation enhancement), with the parameters of the pulse sequence set at TE = 11 ms and TR = 2.5 s. T2-weighted high-resolution mouse brain images (slice thickness, 0.5 mm; field of view, 2.0×2.0 cm; and matrix size, 256×256 points) were obtained. The voxel location is shown in Fig. 1, *a*. All proton spectra were obtained using spatially localized single-voxel spectroscopy using the STEAM method (stimulated echo acquisition mode spectroscopy) with pulse sequence parameters of TE = 3 ms, TR = 5 s, and the number of accumulations = 180. Before each spectroscopic measurement, the uniformity of the magnetic field was adjusted within the selected voxel using the FastMap technique (Gruetter, 1993). The water signal in the spectra was suppressed using a variable power pulse and an optimized relaxation sequence delay (VAPOR) (Tkac et al., 1999).

MRS processing. The LCModel software package was used to process the experimental ^{1}H MRS spectra and determine the number of individual metabolites (Provencher, 2001). This software package has a high degree of automation with minimal user intervention, which minimizes biased input data, thus allowing data exchange, acceptance, and reliable comparison with values obtained using different equipment or under different conditions. Metabolites with Cramer-Rao lower bounds, which represent the estimates of the percentage standard deviation of the fit for each metabolite, < 20 % were considered reliable and are presented in this study. MRS processing is described in more detail in (Singh et al., 2016). According to the manual for the LCModel, the following levels were estimated: the sum of N-acetylaspartate and N-acetylaspartylglutamic acid (tNAA), myoinositol (mIno), taurine (Tau), glutathione (GSH), the sum of glycerophosphocholine and phosphocholine (tCho), glutamic acid (Glu), glutamine (Gln), gamma aminobutyric acid (GABA) as a relation to the sum of creatine and creatine phosphate (tCr). Moreover, the ratios of neurotransmitters and their precursors Glu/GABA and Glu/Gln were estimated (see Fig. 1).

Statistical analysis. We used Student's *t*-test for dependent samples to process the results. The values of the studied parameters are presented as means ± standard error of the mean.

Results

In studies using *in vivo* MRS, data are presented both as concentrations or relative units of individual metabolites, and as the ratio of individual metabolites to tCr, using it as an internal control. Since in our work the level of tCr in animals did not differ from the base level at all the studied time points ($p > 0.19$), it would be appropriate in the future to present the data as the ratio of individual metabolites to tCr.

Evaluation of indicators of brain cell viability, including tNAA/tCr (indicator of neuronal viability), mIno/tCr (indicator of glial cell viability), tCho/tCr (integrity of cell membranes), Tau/tCr (osmolyte, neuromodulator, and trophic factor) and GSH/tCr (antioxidant) showed that all of the metabolites returned to their initial levels on day 28 after injection ($p > 0.12$), except mIno (Fig. 2).

The levels of metabolites reflected the viability and integrity of cells decreased in the first week of the study and then recovered two weeks after the injection significantly for tNAA/tCr ($p = 0.0152$) and slightly for tCho/tCr ($p = 0.062$) whereas the mIno/tCr level was significantly lower than the original on the 21st and 28th day after injection ($p = 0.0059$ and $p = 0.0484$, respectively). A different picture was observed for Tau/tCr and GSH/tCr, which reflect the energy processes that take place in cells (see Fig. 2). It was found that the relative values of these metabolites increased significantly for Tau/tCr and slightly for GSH/tCr with maximum at 14 days after administration: $p = 0.001$ and $p = 0.054$, respectively. The levels of Tau/tCr and GSH/tCr returned to the original values 28 days after the injection (see Fig. 2).

The analysis of the relative level of metabolites performing the functions of neurotransmitters (Glu/tCr, Gln/tCr and GABA/tCr) also showed the dynamic processes in brain tissue after injection (Fig. 3). The Gln/tCr level progressively decreased from 7 to 28 days after the injection and was significantly lower than the original level on the 14th and 28th day after the surgery ($p = 0.028$ and $p = 0.026$, respectively). The GABA/tCr level had the same dynamic profile, but on the 7th day, there was a dip before minimal values ($p = 0.015$) with

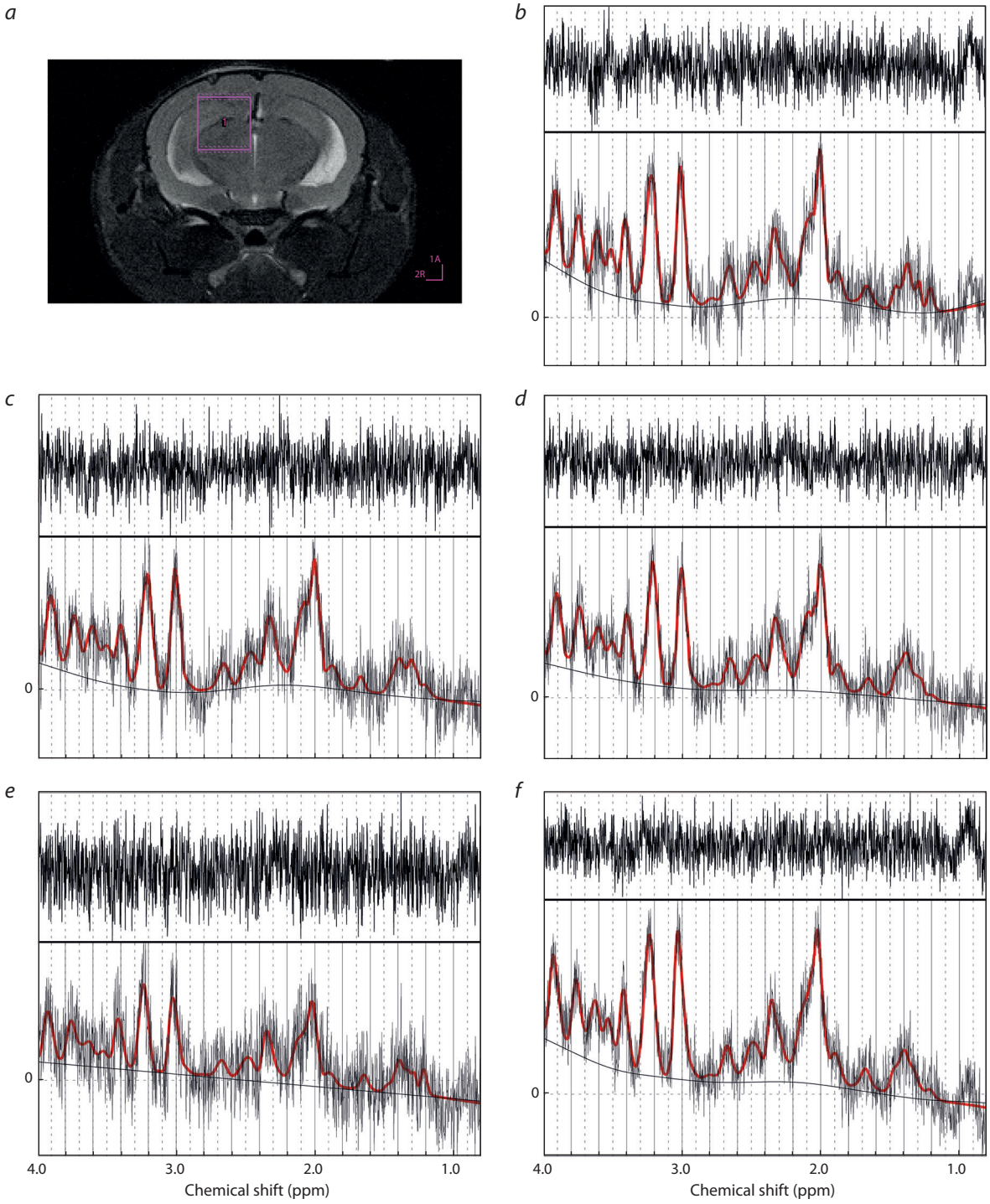

Fig. 1. Voxel location (*a*), characteristic spectra and FIDs (LCModel) before (*b*) and 7 (*c*), 14 (*d*), 21 (*e*) or 28 (*f*) days after injection.

a following recovery and the next extinction on the 28th day ($p = 0.034$). On the contrary, Glu/tCr increased, but only on 14 days after injection, maintaining high values up to 28 days ($p = 0.017$, $p = 0.019$ and $p = 0.045$, respectively).

A particularly interesting results were found at analyzing the Glu/GABA and Glu/Gln ratios that reflect the effectiveness of neurotransmitters metabolism, because both the GABA and Gln have converted from Glu (Fig. 4). It was found that on the 7th day after the injection of the culture medium, the Glu/GABA ratio significantly increased indicating a shift in the balance of neurotransmitters towards excitatory. These differences had disappeared by the 14th day post-injection, but with some progressive growth to 28th day. These differences were significant only for 21 and 28 days ($p = 0.024$ and $p = 0.002$, respectively). Contrarily, the Glu/Gln ratio steadily increased during the experiment before a significant difference appeared on 14, 21 and 28 days after injection ($p = 0.003$, $p = 0.027$ and $p = 0.015$, respectively).

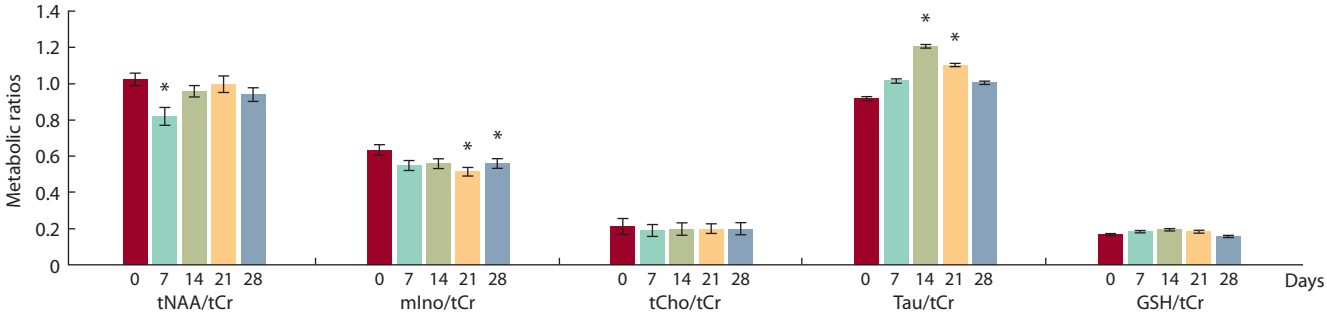

Fig. 2. Dynamics of the relative levels of 5 metabolites (tNAA/tCr, mIno/tCr, Tau/tCr, GSH/tCr, tCho/tCr) with levels before (0 days) and 7, 14, 21, and 28 days after the intracranial injection.

Here and in Fig. 3 and 4: * Significant difference from 0 days ($p < 0.05$; Student's *t*-test for dependent samples).

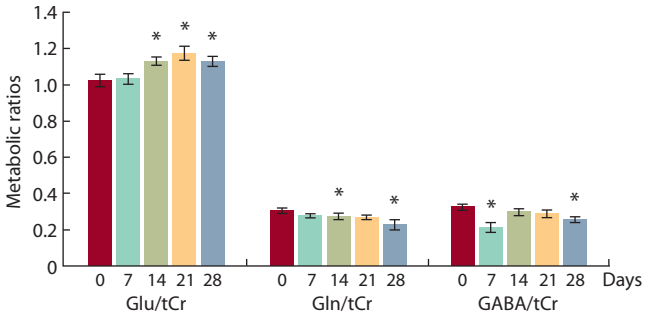

Fig. 3. Dynamics of the relative levels of neurotransmitters (Glu/tCr, Gln/tCr and GABA/tCr) before (0 days) and 7, 14, 21, 28 days after the intracranial injection.

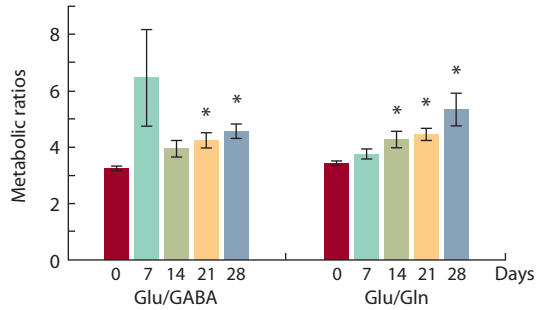

Fig. 4. The Glu/GABA and Glu/Gln ratios before (0 days) and 7, 14, 21, 28 days after the intracranial injection.

Discussion

The change in the tNAA/tCre ratio serves as an indicator of the dynamics of neuronal integrity and is an important biomarker of brain injury (Moffett et al., 2007). tNAA is involved in important processes such as lipid and myelin synthesis (Van Horn et al., 2017). In our study, the decrease in the tNAA level on day seven after surgery may indicate both potential neuronal loss and diffuse axonal injury characteristic of the early stages of brain trauma (Moffett et al., 2007), and its eventually recovery may indicate the stabilization process and the restoration of neurons and their viability.

Another metabolite, mIno, is an osmolyte found mainly in astrocytes and glial cells of the microglia. Increased mIno levels are observed during both the subacute and chronic stages after experimental brain trauma, as well as after moderate or severe head trauma in humans (Brooks et al., 2001; Ashwal et al., 2004; Kierans et al., 2014). These changes are associated with reactive astrocytosis and microgliosis, accompanied by increased glial content and proliferation in response to brain injury (Ashwal et al., 2004), accelerated myelin breakdown, or hypertensive stress (Fisher et al., 2002). However, we found that mIno did not increase but rather decreased during the 28 days of the experiment, reaching the minimum values on day 21. Based on the data obtained, it can be assumed that the destruction of the integrity of the brain tissue caused by the needle injection of the culture medium did not lead to the changes in glial cell function.

tCho, including choline compounds such as acetylcholine, phosphatidylcholine, and phosphocholine, is considered to be a product of nerve myelin disintegration and measures the membrane turnover (Xu et al., 2011; Li J. et al., 2017). In a study on rats (Lescot et al., 2010), two days after brain damage, a slight increase in the tCho level was detected, followed by recovery to the control value by seven days. In our study, we found a slight decrease in the tCho level seven days after the surgery, followed by its recovery to the baseline level, which is consistent with the behavior of the NAA level and may reflect the violation of membrane integrity and myelin degradation processes in the first week after surgery.

Another important indicator of brain cell activity is the balance of Tau levels. Tau is an endogenous amino acid synthesized in large quantities by neurons and astrocytes in the central nervous system. Tau acts as an osmolyte, neuromodulator, trophic factor, stabilizer of membrane integrity, and regulator of intracellular calcium homeostasis (Niu et al., 2018; Gupte et al., 2019). The significant increase in Tau at the injection region at 14 and 21 days post-injection, with the subsequent decrease at 28 days, may act as an adaptive brain tissue response to reduce the negative effects of the brain damage.

In response to the brain injury, markers of oxidative stress are produced in the brain, while levels of antioxidant defense enzymes (including GSH) are reduced (Rodríguez-Rodríguez et al., 2014). The addition of Tau to the culture medium in a neuron model with trauma led to an increase in GSH levels and, accordingly, a decrease in the effect of oxidative stress and in the levels of pro-inflammatory cytokines (Niu et al., 2018). In our study, it is likely that the increased Tau levels

slowed the decline in GSH levels in the first three weeks after injection. The increase in the Tau level in the injection area led to a slight growth of the level of GSH in 14 days. Then there was a subsequent decrease in Tau levels by 21 and 28 days after the injection, ultimately accompanied by the reduction of GSH levels, which may indicate a completion of the traumatic process in 28 days after the traumatic event.

The Glu level in brain trauma models decreases starting from the first hours after trauma and remains at a low level up to two weeks post-exposure (Xu et al., 2011). This decrease may be due to an increased release of Glu into the synaptic cleft in response to brain injury, its capture by astrocytes, and its subsequent accelerated conversion to Gln, as evidenced by an increase in Gln (Guerriero et al., 2015; Van Horn et al., 2017). This is confirmed by the analysis of the Glu/Gln ratio. Changes in the Glu/Gln ratio may indicate neuronal death or glial cell abnormalities. In our study, the Glu/Gln ratio was the only indicator the change in which in the injected area intensified over time. There are two potential reasons for this. First, it may indicate a disturbance in the conversion of Glu to Gln in astrocytes, caused both by the transport of Glu into the astrocyte and its conversion to Gln under the action of glutamine synthetase. Second, the change in the Glu/Gln ratio may have been caused by hyperactivity of the reverse conversion of Gln to Glu in neurons, likely due to glutaminase hyperactivation (Guerriero et al., 2015; Van Horn et al., 2017). Regardless, both of these processes indicate a breakdown of the Gln to Glu conversion system specific to brain trauma processes.

Analysis of the GABA level showed a two-wave decline: on 7 and 28 days. These results generally agree with those in the literature, where it was shown that in the acute phase of brain trauma, the GABA level decreases from the first day after injury (Harris et al., 2012). We attribute this decrease to the reduced conversion of Glu to GABA. The same effect can be caused by the accelerated uptake and conversion of GABA to Gln by oligodendrocytes (Van Horn et al., 2017). Additionally, the Glu/GABA ratio reflects the balance of excitatory and inhibitory neurotransmitters and its displacement may indicate the development of post-traumatic pathology, for example, to epileptogenesis (Cantu et al., 2015).

In a study on rats with mild traumatic brain injury caused by the blast, it was shown that the NAA level did not change, changes were observed only for mIno, Glu, and Tau, and only the Tau level was restored to the initial values after seven days, while the other metabolites showed an increase (Li Y. et al., 2021). In our study, metabolites characterizing the viability of neurons returned to their original level on day 28. However, the levels of mIno, as an indicator of the glial cells state, and the levels of neutotransmitters and their predictor (Glu, GABA and Gln) did not return to their initial values within 28 days, which may indicate some incomplete processes of tissue repair (Rodríguez-Rodríguez et al., 2014; Guerriero et al., 2015). All this results in longitudinal changes to the functional state of the brain cells obtained during the modeling surgery procedure of glioma.

Many studies have aimed to determine methods of early diagnosis of brain tumor diseases, including the use of *in vivo* ^{1}H MRS (Porcari et al., 2016; Hyare et al., 2017). It is known that during the development of glioblastoma, especially at its later stages, the amount of NAA is significantly reduced. Also, at the early stages of tumor development, the level of mIno increases significantly, decreasing towards the later stage (Bulik et al., 2013). At the same time, our data show that the process of cell xenotransplantation itself leads to a decrease in NAA on the 7th day after the surgery and a decrease in the level of mIno on the 21st day. Besides, glioma cells that secrete Glu lead to an increase in extracellular Glu. Although Gln concentrations in the contralateral brain tissue in patients with glioblastoma were significantly elevated compared with the levels found in normal brain (Chaumeil et al., 2015). In our study, we found an elevated level of Glu and an increased level of Gln.

The data obtained on the dynamics of the level of neurotransmitters in the brain during the simulation of the xenotransplantation process, on the one hand, can be the result of surgical intervention, as well as the result of multiple MRS procedures using anesthesia. At the same time, there is evidence of minimal effects of isoflurane anesthesia on the metabolomic profile in animals (Menshanov, Akulov, 2015; Söbbeler et al., 2018).

Taking together, when using animal models of human glioma, researchers should clearly distinguish between the changes in neurometabolites that are a response to brain injury caused by the injection of cancer cells into the brain, and the processes that may indicate the early development of a brain tumor. Therefore, this is important for understanding how the level of metabolites changes during the process of tumor development.

Conclusion

For modeling orthotopic xenotransplantation of glioma cells into the brain of mice, it is necessary to understand which metabolites are produced by a growing tumor and which are the result of surgery invasion of the tumor cells injection. The dynamic of nine neurometabolites in the mouse brain after needle injection with *in vivo* ^{1}H magnetic resonance spectroscopy was studied. On the 28th day after injection, only metabolic levels of cells reflecting neurons' viability in the area of the injection were restored. However, the levels of neutotransmitters and their predictor (Glu, GABA and Gln) did not return to their initial values within 28 days. So, the recovery processes are largely completed on the 28th day after the traumatic event in the brain tissue, leaving open the question of the neurotransmitter system impairment. It is important to keep in mind when modeling human glioblastoma in mice and monitoring new treatments. In addition, these results may be important at the development of approaches for non-invasive diagnostics of traumatic brain injury as well as recovery and rehabilitation processes of patients after certain brain surgeries.

References

Ashwal S., Holshouser B., Tong K., Serna T., Osterdock R., Gross M., Kido D. Proton spectroscopy detected myoinositol in children with traumatic brain injury. *Pediatr. Res.* 2004;56(4):630-638. DOI 10.1203/01.PDR.0000139928.60530.7D.

Atkinson M., Juhasz C., Shah J., Guo X., Kupsky W., Fuerst D., Johnson R., Watson C. Paradoxical imaging findings in cerebral gliomas. *J. Neurol. Sci.* 2008;269(1-2):180-183. DOI 10.1016/j.jns.2007.12.029.

Barker F.G. 2nd, Chang S.M., Huhn S.L., Davis R.L., Gutin P.H., McDermott M.W., Wilson C.B., Prados M.D. Age and the risk of anaplasia in magnetic resonance-nonenhancing supratentorial cerebral tumors. *Cancer.* 1997;80(5):936-941.

Brooks W.M., Friedman S., Gasparovic C. Magnetic resonance spectroscopy in traumatic brain injury. *J. Head Trauma Rehabil.* 2001; 16(2):149-164. DOI 10.1097/00001199-200104000-00005.

Bulik M., Jancalek R., Vanicek J., Skoch A., Mechl M. Potential of MR spectroscopy for assessment of glioma grading. *Clin. Neurol. Neurosurg.* 2013;115(2):146-153. DOI 10.1016/j.clineuro.2012.11.002.

Cantu D., Walker K., Andresen L., Taylor-Weiner A., Hampton D., Tesco G., Dulla C.G. Traumatic brain injury increases cortical glutamate network activity by compromising GABAergic control. *Cereb. Cortex.* 2015;25(8):2306-2320. DOI 10.1093/cercor/bhu041.

Castillo M., Smith J.K., Kwock L., Wilber K. Apparent diffusion coefficients in the evaluation of high-grade cerebral gliomas. *Am. J. Neuroradiol.* 2001;22(1):60-64.

Chaumeil M.M., Lupo J.M., Ronen S.M. Magnetic resonance (MR) metabolic imaging in glioma. *Brain Pathol.* 2015;25(6):769-780. DOI 10.1111/bpa.12310.

Dass C.R., Choong P.F. GFP expression alters osteosarcoma cell biology. *DNA Cell Biol.* 2007;26(8):599-601. DOI 10.1089/dna.2006.053.

FELASA working group on revision of guidelines for health monitoring of rodents and rabbits; Mähler Convenor M., Berard M., Feinstein R., Gallagher A., Illgen-Wilcke B., Pritchett-Corning K., Raspa M. FELASA recommendations for the health monitoring of mouse, rat, hamster, guinea pig and rabbit colonies in breeding and experimental units. *Lab. Anim.* 2014;48(3):178-192. DOI 10.1177/0023677213516312.

Fisher S.K., Novak J.E., Agranoff B.W. Inositol and higher inositol phosphates in neural tissues: homeostasis, metabolism and functional significance. *J. Neurochem.* 2002;82(4):736-754. DOI 10.1046/j.1471-4159.2002.01041.x.

Ginsberg L.E., Fuller G.N., Hashmi M., Leeds N.E., Schomer D.F. The significance of lack of MR contrast enhancement of supratentorial brain tumors in adults: histopathological evaluation of a series. *Surg. Neurol.* 1998;49(4):436-440. DOI 10.1016/s0090-3019(97)00360-1.

Goodenberger M.L., Jenkins R.B. Genetics of adult glioma. *Cancer Genet.* 2012;205(12):613-621. DOI 10.1016/j.cancergen.2012.10.009.

Gruetter R. Automatic, localized *in vivo* adjustment of all first-and second-order shim coils. *Magn. Reson. Med.* 1993;29(6):804-811. DOI 10.1002/mrm.1910290613.

Guerriero R.M., Giza C.C., Rotenberg A. Glutamate and GABA imbalance following traumatic brain injury. *Curr. Neurol. Neurosci. Rep.* 2015;15(5):27. DOI 10.1007/s11910-015-0545-1.

Gupte R., Christian S., Keselman P., Habiger J., Brooks W.M., Harris J.L. Evaluation of taurine neuroprotection in aged rats with traumatic brain injury. *Brain Imaging Behav.* 2019;13(2):461-471. DOI 10.1007/s11682-018-9865-5.

Haddad A.F., Young J.S., Amara D., Berger M.S., Raleigh D.R., Aghi M.K., Butowski N.A. Mouse models of glioblastoma for the evaluation of novel therapeutic strategies. *Neurooncol. Adv.* 2021; 3(1):vdab100. DOI 10.1093/noajnl/vdab100.

Hall E.D., Sullivan P.G., Gibson T.R., Pavel K.M., Thompson B.M., Scheff S.W. Spatial and temporal characteristics of neurodegeneration after controlled cortical impact in mice: More than a focal brain injury. *J. Neurotrauma.* 2005;22(2):252-265. DOI 10.1089/neu.2005.22.252.

Harris J.L., Yeh H.W., Choi I.Y., Lee P., Berman N.E., Swerdlow R.H., Craciunas S.C., Brooks W.M. Altered neurochemical profile after traumatic brain injury: ^{1}H-MRS biomarkers of pathological mechanisms. *J. Cereb. Blood Flow Metab.* 2012;32(12):2122-2134. DOI 10.1038/jcbfm.2012.114.

Hishii M., Matsumoto T., Arai H. Diagnosis and treatment of early-stage glioblastoma. *Asian J. Neurosurg.* 2019;14(2):589-592. DOI 10.4103/ajns.AJNS_18_19.

Hyare H., Thust S., Rees J. Advanced MRI techniques in the monitoring of treatment of gliomas. *Curr. Treat. Options Neurol.* 2017; 19(3):11. DOI 10.1007/s11940-017-0445-6.

Kierans A.S., Kirov I.I., Gonen O., Haemer G., Nisenbaum E., Babb J.S., Grossman R.I., Lui Y.W. Myoinositol and glutamate complex neurometabolite abnormality after mild traumatic brain injury. *Neurology.* 2014;82(6):521-528. DOI 10.1212/WNL.0000000000000105.

Lescot T., Fulla-Oller L., Po C., Chen X.R., Puybasset L., Gillet B., Plotkine M., Meric P., Marchand-Leroux C. Temporal and regional changes after focal traumatic brain injury. *J. Neurotrauma.* 2010; 27(1):85-94. DOI 10.1089/neu.2009.0982.

Li J., Zhao C., Rao J.S., Yang F.X., Wang Z.J., Lei J.F., Yang Z.Y., Li X.G. Structural and metabolic changes in the traumatically injured rat brain: high-resolution in vivo proton magnetic resonance spectroscopy at 7 T. *Neuroradiology.* 2017;59(12):1203-1212. DOI 10.1007/s00234-017-1915-y.

Li Y., Liu K., Li C., Guo Y., Fang J., Tong H., Tang Y., Zhang J., Sun J., Jiao F., Zhang Q., Jin R., Xiong K., Chen X. 18F-FDG PET combined with MR spectroscopy elucidates the progressive metabolic cerebral alterations after blast-induced mild traumatic brain injury in rats. *Front. Neurosci.* 2021;15:593723. DOI 10.3389/fnins.2021.593723.

Maas A.I., Stocchetti N., Bullock R. Moderate and severe traumatic brain injury in adults. *Lancet Neurol.* 2008;7(8):728-741. DOI 10.1016/S1474-4422(08)70164-9.

Menshanov P.N., Akulov A.E. The neurochemical profile of the hippocampus in isoflurane-treated and unanesthetized rat pups. *Interdiscip. Toxicol.* 2015;8(3):113-117. DOI 10.1515/intox-2015-0017.

Miyai M., Tomita H., Soeda A., Yano H., Iwama T., Hara A. Current trends in mouse models of glioblastoma. *J. Neurooncol.* 2017; 135(3):423-432. DOI 10.1007/s11060-017-2626-2.

Moffett J.R., Ross B., Arun P., Madhavarao C.N., Namboodiri A.M. *N*-Acetylaspartate in the CNS: from neurodiagnostics to neurobiology. *Prog. Neurobiol.* 2007;81(2):89-131. DOI 10.1016/j.pneurobio.2006.12.003.

Nelson S.J., Cha S. Imaging glioblastoma multiforme. *Cancer J.* 2003; 9(2):134-145. DOI 10.1097/00130404-200303000-00009.

Niu X., Zheng S., Liu H., Li S. Protective effects of taurine against inflammation, apoptosis, and oxidative stress in brain injury. *Mol. Med. Rep.* 2018;18(5):4516-4522. DOI 10.3892/mmr.2018.9465.

Ostrom Q.T., Cote D.J., Ascha M., Kruchko C., Barnholtz-Sloan J.S. Adult glioma incidence and survival by race or ethnicity in the United States from 2000 to 2014. *JAMA Oncol.* 2018;4(9):1254-1262. DOI 10.1001/jamaoncol.2018.1789.

Ostrom Q.T., Cioffi G., Gittleman H., Patil N., Waite K., Kruchko C., Barnholtz-Sloan J.S. CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. *Neuro Oncol.* 2019;21(Suppl. 5):v1-v100. DOI 10.1093/neuonc/noz150.

Porcari P., Hegi M.E., Lei H., Hamou M.F., Vassallo I., Capuani S., Gruetter R., Mlynarik V. Early detection of human glioma sphere xenografts in mouse brain using diffusion MRI at 14.1 T. *NMR Biomed.* 2016;29(11):1577-1589. DOI 10.1002/nbm.3610.

Provencher S.W. Automatic quantitation of localized *in vivo* ^{1}H spectra with LCModel. *NMR Biomed.* 2001;14(4):260-264. DOI 10.1002/nbm.698.

Ratai E.M., Gilberto González R. Clinical magnetic resonance spectroscopy of the central nervous system. *Handb. Clin. Neurol.* 2016; 135:93-116. DOI 10.1016/B978-0-444-53485-9.00005-2.

Rodríguez-Rodríguez A., Egea-Guerrero J.J., Murillo-Cabezas F., Carrillo-Vico A. Oxidative stress in traumatic brain injury. *Curr. Med. Chem.* 2014;21(10):1201-1211. DOI 10.2174/09298673216661312 17153310.

Singh K., Trivedi R., Haridas S., Manda K., Khushu S. Study of neurometabolic and behavioral alterations in rodent model of mild traumatic brain injury: a pilot study. *NMR Biomed.* 2016;29(12):1748-1758. DOI 10.1002/nbm.3627.

Söbbeler F.J., Carrera I., Pasloske K., Ranasinghe M.G., Kircher P., Kästner S.B.R. Effects of isoflurane, sevoflurane, propofol and alfaxalone on brain metabolism in dogs assessed by proton magnetic resonance spectroscopy (^{1}H MRS). *BMC Vet. Res.* 2018;14(1):69. DOI 10.1186/s12917-018-1396-1.

Stockhammer F., Plotkin M., Amthauer H., van Landeghem F.K., Woiciechowsky C. Correlation of F-18-fluoro-ethyl-tyrosin uptake with vascular and cell density in non-contrastenhancing gliomas. *J. Neurooncol.* 2008;88(2):205-210. DOI 10.1007/s11060-008-9551-3.

Sugahara T., Korogi Y., Kochi M., Ikushima I., Hirai T., Okuda T., Shigematsu Y., Liang L., Ge Y., Ushio Y., Takahashi M. Correlation of MR imaging-determined cerebral blood volume maps with histologic and angiographic determination of vascularity of gliomas. *AJR Am. J. Roentgenol.* 1998;171(6):1479-1486. DOI 10.2214/ajr.171.6.9843274.

Tan A.C., Ashley D.M., Lopez G.Y., Malinzak M., Friedman H.S., Khasraw M. Management of glioblastoma: State of the art and future directions. *CA Cancer J. Clin.* 2020;70(4):299-312. DOI 10.3322/caac.21613.

Tiwari V., Mashimo T., An Z., Vemireddy V., Piccirillo S., Askari P., Hulsey K.M., Zhang S., de Graaf R.A., Patel T.R., Pan E., Mickey B.E., Maher E.A., Bachoo R.M., Choi C. In vivo MRS measurement of 2-hydroxyglutarate in patient-derived IDH-mutant xenograft mouse models versus glioma patients. *Magn. Reson. Med.* 2020;84(3):1152-1160. DOI 10.1002/mrm.28183.

Tkac I., Starcuk Z., Choi I.-Y., Gruetter R. In vivo ^{1}H NMR spectroscopy of rat brain at 1 ms echo time. *Magn. Reson. Med.* 1999;41(4):649-656. DOI 10.1002/(sici)1522-2594(199904)41:4<649::aid-mrm2>3.0.co;2-g.

Tykocki T., Eltayeb M. Ten-year survival in glioblastoma. A systematic review. *J. Clin. Neurosci.* 2018;54:7-13. DOI 10.1016/j.jocn.2018.05.002.

Van Horn J.D., Bhattrai A., Irimia A. Multimodal imaging of neurometabolic pathology due to traumatic brain injury. *Trends Neurosci.* 2017;40(1):39-59. DOI 10.1016/j.tins.2016.10.007.

Xu S., Zhuo J., Racz J., Shi D., Roys S., Fiskum G., Gullapalli R. Early microstructural and metabolic changes following controlled cortical impact injury in rat: a magnetic resonance imaging and spectroscopy study. *J. Neurotrauma.* 2011;28(10):2091-2102. DOI 10.1089/neu.2010.1739.

Zavjalov E.L., Razumov I.A., Gerlinskaya L.A., Romashchenko A.V. In vivo MRI visualization of U87 glioblastoma development dynamics in the model of orthotopic xenotransplantation to the scid mouse. *Russ. J. Genet. Appl. Res.* 2016;6(4):448-453. DOI 10.1134/S2079059716040225.

ORCID ID

O.B. Shevelev orcid.org/0000-0003-3200-958X
O.P. Cherkasova orcid.org/0000-0002-8035-2422
I.A. Razumov orcid.org/0000-0002-6756-1457
E.L. Zavjalov orcid.org/0000-0002-9412-3874

Acknowledgements. The study was carried out with the support of the Russian Foundation for Basic Research (RFBR) grant 19–52–55004 using the equipment of the Center for Genetic Resources of Laboratory Animals, Institute of Cytology and Genetics, SB RAS, supported by the Ministry of Education and Science of Russia (RFMEFI62119X0023).

Author contributions. Investigation, O.S., I.R. and E.Z.; data analysis, O.S. and E.Z.; writing – original draft preparation, O.S.; writing – review and editing, E.Z.; project administration, E.Z., O.C.; funding acquisition, O.C. All authors have read and agreed to the published version of the manuscript.

Institutional review board statement. The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Inter-Institutional Commission on Biological Ethics at the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences (Permission #78, April 16, 2021).

Data availability statement. The data generated and/or analyzed during this research are presented in this published article. All raw data (MR spectra and MR images) are available from the corresponding author on reasonable request.

Conflict of interest. The authors declare no conflict of interest.

Received May 16, 2022. Revised February 21, 2023. Accepted June 1, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):641-650
DOI 10.18699/VJGB-23-75

Original Russian text https://vavilovj-icg.ru/

Germline-restricted chromosomes of the songbirds

P.M. Borodin

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
borodin@bionet.nsc.ru

Abstract. Germline-restricted chromosomes (GRCs) are present in the genomes of germline cells and absent from somatic cells. A GRC is found in all species of the songbirds (Passeri) and in none of the other bird orders studied to date. This indicates that GRC originated in the common ancestor of the songbirds. The germline-restricted chromosome is permanently absent from somatic cells of the songbird, while female germline cells usually contain two copies of GRC and male ones have one copy. In females, GRCs undergo synapsis and restricted recombination in their terminal regions during meiotic prophase. In males, it is almost always eliminated from spermatocytes. Thus, GRC is inherited almost exclusively through the maternal lineage. The germline-restricted chromosome is a necessary genomic element in the germline cells of songbirds. To date, the GRC genetic composition has been studied in four species only. Some GRC genes are actively expressed in female and male gonads, controlling the development of germline cells and synthesis of the proteins involved in the organization of meiotic chromosomes. Songbird species vary in GRC size and genetic composition. The GRC of each bird species consists of amplified and modified copies of genes from the basic genome of that species. The level of homology between GRCs of different species is relatively low, indicating a high rate of genetic evolution of this chromosome. Transmission through the maternal lineage and suppression of the recombination contribute significantly to the accelerated evolution of GRCs. One may suggest that the rapid coordinated evolution between the GRC genes and the genes of the basic genome in the songbirds might be responsible for the explosive speciation and adaptive radiation of this most species-rich and diverse infraorder of birds.
Key words: germline-restricted chromosomes; avian genome evolution; programmed DNA elimination.

For citation: Borodin P.M. Germline-restricted chromosomes of the songbirds. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):641-650. DOI 10.18699/VJGB-23-75

Хромосомы певчих птиц, ограниченные зародышевой линией

П.М. Бородин

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
borodin@bionet.nsc.ru

Аннотация. Хромосомы, ограниченные зародышевой линией (germline-restricted chromosomes – GRC), присутствуют в геномах герминативных клеток и запрограммированно элиминируются из соматических клеток в ходе развития. Они крайне редко возникают в ходе эволюции. Еще реже они становятся необходимыми элементами геномов герминативных клеток крупных таксонов. Такая хромосома, ограниченная зародышевой линией, была обнаружена у всех исследованных представителей подотряда воробьинообразных певчих птиц. Ни у одного из представителей других отрядов птиц GRC не выявлено. Это свидетельствует о том, что GRC возникла у общего предка воробьинообразных певчих птиц. У всех представителей этого подотряда GRC присутствует, как правило, в двух копиях в герминативных клетках самок и в одной копии у самцов и отсутствует в соматических клетках. У самок GRC синаптируют и рекомбинируют в профазе мейоза. У самцов они почти всегда элиминируются из сперматоцитов. Таким образом, GRC наследуется почти исключительно по материнской линии. Хромосомы, ограниченные зародышевой линией, – необходимый элемент генома герминативных клеток певчих птиц. На сегодняшний день исследованы геномы GRC четырех видов. Некоторые гены GRC активно экспрессируются в гонадах самцов и самок, контролируют развитие герминативных клеток, синтез белков, вовлеченных в организацию мейотических хромосом. GRC разных видов различаются по размерам и генетическому составу. Геном GRC каждого вида птиц состоит из амплифицированных перестроенных копий генов основного генома данного вида. Уровень гомологии между GRC разных видов довольно низок. Это указывает на высокую скорость эволюции генетического состава хромосомы. Значительный вклад в ускорение эволюции GRC вносят передача этой хромосомы по материнской линии и подавление рекомбинации в ней. Можно считать, что быстрая согласованная эволюция генов GRC и генов основного набора воробьинообразных певчих птиц играет важную роль в видообразовании и адаптивной радиации представителей этого самого богатого видами и разнообразного подотряда птиц.
Ключевые слова: хромосомы, ограниченные зародышевой линией; эволюция генома птиц; запрограммированная элиминация ДНК.

Programmed DNA elimination

In the late 19th century, A. Weismann proposed one of the most influential ideas in modern biology. He suggested a fundamental distinction between somatic cells and germ line cells (Weismann, 1890, 1893). According to A. Weismann, germline cells possess a complete set of hereditary material (germ plasm), while somatic cells undergo unequal divisions during their differentiation and specialization, which leads to a fragmentation of the germ plasm, with different daughter somatic cells acquiring its different fragments.

The former part of Weismann's hypothesis became the basis of modern evolutionary biology, genetics, and developmental biology. The latter part underwent significant modifications. Unequal division of somatic cells during early development in the vast majority of modern multicellular organisms does not lead to genome fragmentation (as suggested by Weismann), but rather to the uneven distribution of its products (proteins and RNAs) among the daughter cells. This, in turn, leads to progressive differentiation of their gene activity and, ultimately, to their specialization (Dröscher, 2014).

However, somatic cells in some species behave according to Weismann's hypothesis: they lose a portion of their genomes during the differentiation. This phenomenon was previously referred to as "chromatin diminution". The prevailing term now is "programmed DNA elimination" (Wang, Davis, 2014). The first example of chromatin diminution was discovered by T. Boveri in the nematode *Ascaris megalocephala* (Boveri, 1887). During the early stages of its development, the nuclei of certain cells progressively lose a portion of their chromosomal material. T. Boveri suggested that chromatin diminution is part of a normal developmental program necessary for the formation of specific tissues and organs in the worm (Maderspacher, 2008).

Programmed DNA elimination is observed in various taxa of multicellular organisms. It plays a key role in regulating gene expression and maintaining genomic stability. Various molecular pathways involved in selective DNA removal during chromatin diminution have been identified (Wang, Davis, 2014; Smith, 2018; Smith et al., 2021; Dedukh, Krasikova, 2022). It has been shown that the regions of the genome that are eliminated are the homologous counterparts of those undergoing silencing in other multicellular organisms, such as oncogenes and other early developmental genes (Wang et al., 2012; Streit et al., 2016; Smith, 2018; Smith et al., 2018).

Thus, programmed DNA elimination can be considered a radical form of silencing, a specific way to resolve the problem of antagonistic pleiotropy between early and late genetic programs. Apparently, this mechanism of developmental regulation through programmed DNA elimination has emerged repeatedly in evolution and has proven to be an evolutionary dead end. Its erratic distribution across the phylogenetic tree of multicellular animals supports this hypothesis (Smith et al., 2021).

Whole chromosome elimination

Elimination in somatic cells usually affects chromosome fragments rather than entire chromosomes. Cases of whole chromosome elimination are less frequent (Dedukh, Krasikova, 2022). In several species of mammals and invertebrates from different taxa, chromosome elimination serves as a mechanism of sex determination and/or dosage compensation of genes on heterogametic sex chromosomes (Watson et al., 1998; Wang, Davis, 2014; Smith et al., 2021).

Conflict between host genomes and their genomic parasites sometimes leads to the emergence of additional (i. e., nonessential for survival) chromosomes. They are called B-chromosomes, in contrast to A-chromosomes of the basic karyotype. B-chromosomes vary in size, morphology, and genetic content between species, among individuals within a species, and even among individual cells of an organism (Houben et al., 2014; Camacho, 2022). Several studies detected a tendency for B-chromosomes to accumulate in germ cells and be lost in somatic cells (Camacho et al., 2000; Camacho, 2022).

B-chromosomes may not be present in all populations of a particular species. Species with B-chromosomes also exhibit erratic phylogenetic distribution (D'Ambrosio et al., 2017). For example, B-chromosomes are present in only 16 species from 11 rather species-rich genera out of 1130 species described in the "Atlas of Mammalian Chromosomes" (Graphodatsky et al., 2020). This suggests that B-chromosomes are typical genomic parasites that occasionally arise in some species, sometimes gain pyrrhic victories and spread in populations, but ultimately either become completely eliminated from the genomes of these species or go extinct along with their hosts (Camacho et al., 2000; Houben et al., 2014; Houben, 2017; Camacho, 2022; Chen et al., 2022).

In some cases, however, such parasitic chromosomes may achieve evolutionary success and become an indispensable part of the genome in germ cells of all or most species within larger taxa. Such germline-restricted chromosomes (GRCs) have been discovered in almost all studied species from three dipteran families: fungus gnats (Sciaridae), gall midges (Cecidomyiidae), and non-biting midges (Chironomidae) (Hodson, Ross, 2021), as well as in all studied species of passerine birds (Passeriformes) – the most species-rich and diverse order of birds (Torgasheva et al., 2019; Borodin et al., 2022).

A comprehensive review of current knowledge on GRCs in Diptera was provided by C.N. Hodson and L. Ross (2021). A recent review of GRCs in passerines was published by a group of authors, including nearly all European researchers of this chromosome and the author of this article (Borodin et al., 2022), in a special issue dedicated to non-Mendelian inheritance and meiotic drive (Hanlon, Larracuente, 2022). It paid special attention to hypotheses regarding the mechanisms of GRC transmission in the germline and its elimination from somatic cells. Another review analyzing GRC in the context of genomic conflicts appeared online in September 2023 (Vontzou et al., 2023).

This review will address issues related to the structure, function, and evolution of this enigmatic chromosome.

Discovery of GRC in birds

GRC in birds was first identified by Argentine cytogeneticists M.I. Pigozzi and A.J. Solari in the zebra finch (*Taeniopygia guttata*) (Pigozzi, Solari, 1998). They stumbled upon it entirely by chance during a comparative study of chromosome synapsis and recombination in bird meiosis. Among the numerous bird species they examined (Solari, Moses, 1973; Rahn, Solari, 1986; Solari, 1992; Solari, Pigozzi, 1993; Pigozzi, Solari, 1997), the zebra finch was the first representative of Passeriformes. In the meiotic cells of this species, they observed something that had not been seen in any previously studied species.

The germline cells of both male and female zebra finches contained an additional chromosome that was not present in bone marrow and other somatic cells. This additional chromosome was larger than all other chromosomes. It was euchromatic in oocytes and heterochromatic in spermatocytes (Pigozzi, Solari, 1998). Furthermore, in female germline cells, it was usually present in two copies, whereas in males, it was present in a single copy (Pigozzi, Solari, 2005).

Another unexpected feature of this chromosome was its absence not only in somatic cells but also in spermatozoa. During pachytene, diplotene, and metaphase I of meiosis, the GRC was localized at the periphery of chromosomal plates. However, it was absent in metaphase II of male meiosis. Additionally, round dense DAPI-positive bodies were observed adjacent to some spermatocytes. Electron microscopy revealed that these round bodies were surrounded by a double membrane. M.I. Pigozzi and her colleagues hypothesized that these micronuclei contained GRCs eliminated after the first meiotic division (Pigozzi, Solari, 1998, 2005; Itoh et al., 2009; Goday, Pigozzi, 2010; Schoenmakers et al., 2010).

Subsequently, A.A. Torgasheva et al. (2019) confirmed this hypothesis in a direct experiment. They microdissected these micronuclei, prepared DNA probes from them, and hybridized these probes to preparations of pachytene spermatocytes and oocytes. In all cases, they observed a strong, specific hybridization signal on the GRCs (Torgasheva et al., 2019). Since the GRC was absent from spermatozoa, M.I. Pigozzi and A.J. Solari (2005) and Y. Itoh et al. (2009) suggested that it must be inherited exclusively through the maternal lineage.

Initially, M.I. Pigozzi and A.J. Solari (1998) classified the zebra finch GRC as a B-chromosome. However, they noted its unique characteristics. B-chromosomes usually vary in number in somatic cells and tend to accumulate in germline cells. The zebra finch GRC was absent in somatic cells and always present in every germline cell. From the outset of the GRC investigation, it became clear that this chromosome was not a facultative but an obligate element of the germline cell genome, essential for gametogenesis.

GRC at the phylogenetic tree of birds

The zebra finch GRC has long been considered an intriguing genetic curiosity, the sole instance of the B-chromosome in birds. The discovery of GRC in germline cells of the Bengalese finch (*Lonchura striata domestica*) (del Priore, Pigozzi, 2014) did not alter this opinion. However, the situation changed dramatically in 2018–2019, when three independent research groups published data indicating the presence of GRC in the germline of many species of birds (Biederman et al., 2018; Kinsella et al., 2019; Torgasheva et al., 2019).

A.A. Torgasheva et al. (2019) obtained direct evidence of the antiquity of GRC in birds and its wide distribution among songbirds. They conducted cytogenetic screening of 15 songbird species and 8 bird species outside this suborder. They used immunolocalization of SYCP3, the major protein of the synaptonemal complex (axial scaffold of meiotic prophase chromosomes), centromeric proteins, and MLH1, the mismatch repair protein, which marks sites of homologous recombination (Fig. 1).

A.A. Torgasheva et al. (2019) demonstrated that GRC is present in all pachytene cells of all investigated individuals of songbird species, including the rook (*Corvus frugilegus*), a representative of the Corvida infraorder, the sister group to the Passerida infraorder, which includes all other songbird families (Fig. 2). The size of GRC varies widely, with some species having macrochromosomal GRC of size rank 1 to 3, and others having micro-GRC. No phylogenetic clustering has been observed for this trait: closely related species differ greatly in the size of their GRC (see Fig. 2). However, GRC exhibits considerable conservatism in terms of its morphology; in nearly all investigated species, GRC is an acrocentric chromosome. The exception is the pied flycatcher, whose GRC is metacentric (Torgasheva et al., 2019; Malinovskaya et al., 2022).

GRC has not been detected in any bird species outside the order Passeriformes, including the budgerigar (*Melopsittacus undulatus*), a representative of the order Psittaciformes, which is the sister group to the order Passeriformes (Torgasheva et al., 2019). The list of species examined for the presence of GRC now includes 27 songbird species with GRC and 9 bird species without GRC outside this suborder (see Fig. 2) (Borodin et al., 2022).

Based on cytogenetic screening, A.A. Torgasheva et al. (2019) concluded that GRC is likely present in the germline cell genomes of all songbirds and absent in all other birds. They suggested that GRC originated in the common ancestor of the Passeri suborder, which includes approximately 5000 species. The germline cells of representatives of the Tyranni suborder, consisting of approximately 1200 species, and the Acanthisitti suborder, which includes two species of New Zealand wrens, have not been analyzed yet. Therefore, it is still unknown whether they have GRC. If they have it, this would indicate that GRC emerged much earlier, in the common ancestor of all passerine birds.

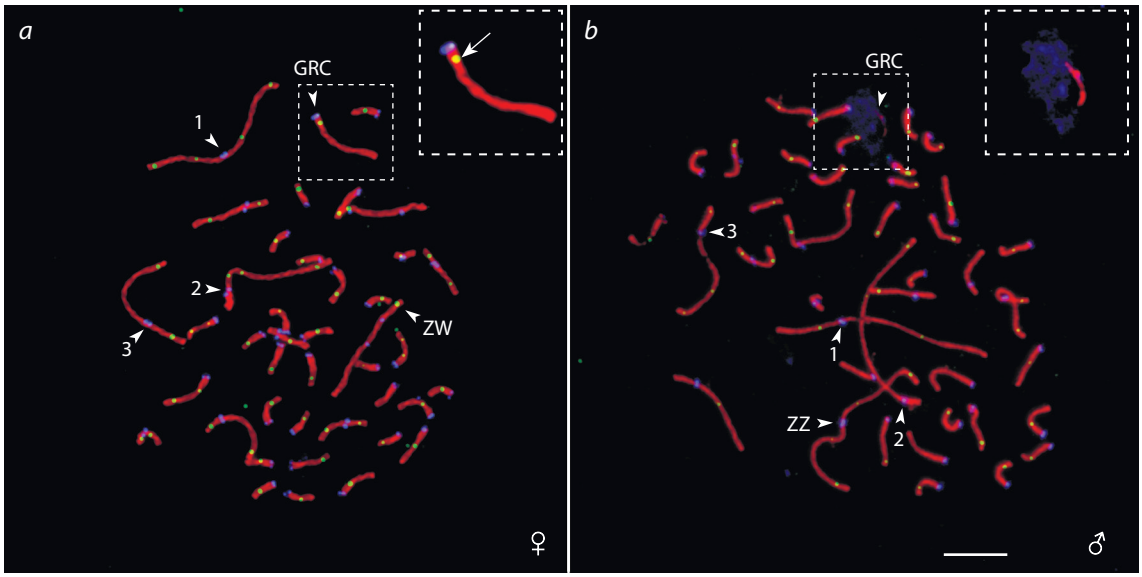

Fig. 1. Germline cells of the great tit at the pachytene stage after immunostaining with antibodies against SYCP3 (red), centromeric proteins (blue), and MLH1 (green).
a – oocyte containing a GRC bivalent; *b* – spermatocyte with a GRC univalent.
The arrowheads indicate the centromeres of the GRC, the three largest macrobivalents, and the heteromorphic ZW bivalent with misaligned centromeres. The arrow points to the MLH1 signal in the pericentromeric region of the GRC bivalent. Scale 5 μm. Photographs from the article (Torgasheva et al., 2019), published in "Proceedings of the National Academy of Sciences", USA, under the CC-BY-SA 4.0 license (modified).

Independent evidence of the evolutionary antiquity of the GRC has been obtained by comparing the results of sequencing the generative and somatic tissues of the zebra finch (Kinsella et al., 2019) and subtractive transcriptomic analysis (Biederman et al., 2018). Over a hundred GRC-specific genes (gametologs) and their somatic paralogs (somatologs) have been identified on at least 19 A-chromosomes.

The results of phylogenetic analysis enabled the categorization of GRC-specific genes into five evolutionary strata based on the degree of divergence from their somatic paralogs (Kinsella et al., 2019).

The level of divergence between gametologs of the youngest stratum 5 and their somatic counterparts is comparable to the level of divergence among the subspecies of the zebra finch. Gametologs of stratum 4 emerged during the divergence of the zebra finch from the long-tailed finch (*Poephila acuticauda*) and the diamond firetail (*Stizoptera bichenovii*). Stratum 3 comprises gametologs that originated from a common ancestor of estrildid species (Estrildidae), stratum 2 corresponds to the ancestor of the Passerida infraorder. The most ancient stratum 1 arose in the common ancestor of the suborder Passeri.

Thus, both cytological and molecular genetic data unambiguously indicate the monophyletic origin of the GRC.

Genetic content of GRC

What does GRC contain that makes it an obligate element of the germline genome of all songbirds? At the time of writing this article, the results of genomic analysis of GRC have been published for only four species of songbirds: the zebra finch (Kinsella et al., 2019), two species of nightingales: western, *Luscinia megarhynchos* and eastern, *L. luscinia* (Schlebusch et al., 2023), and the Eurasian blue tit, *Cyanistes caeruleus* (Mueller et al., 2023).

The most extensively studied zebra finch GRC contains a set of genes crucial for the development of the reproductive system. Contrary to expectations, mobile elements and satellite DNA are not more abundant in GRC compared to A-chromosomes; in fact, they are less abundant (Kinsella et al., 2019; Torgasheva et al., 2019). Transcriptomic analysis revealed the transcription of at least six GRC-specific genes in the testes and 32 genes in the ovaries. Mass spectrometry analysis confirmed the translation of mRNA from some of these genes in the gonads of both sexes. Gene categories related to gonad formation in females and reproduction are overrepresented in the list of GRC-specific genes. Orthologs of many zebra finch GRC-specific genes are predominantly expressed in the chicken gonads of both sexes (Kinsella et al., 2019). Both the macro-GRC of the zebra finch and the micro-GRC of the blue tit contain a functional paralog of the *BMP15* transcription factor, which plays an important role in follicle maturation in the ovaries of birds and other vertebrates (Mueller et al., 2023). In the GRC of the blue tit, J.C. Mueller et al. (2023) revealed the genes controlling the formation of the synaptonemal complex and other genes controlling chromosome synapsis and recombination.

Some of the gametologs found in the zebra finch GRC, including the most ancient ones (*bicc1* and *trim71*), are subjected to purifying selection (Kinsella et al., 2019). The products of their orthologs play an important role in

Fig. 2. Phylogenetic tree of birds with investigated pachytene karyotypes.

Black circles indicate species with macro-GRC, white circles indicate species with micro-GRC, and the absence of circles denotes the absence of GRC. Symbols represent data sources: * Torgasheva et al., (2019); # Malinovskaya et al., (2022); ^ Sotelo-Muñoz et al. (2022); & Poignet et al. (2021). The diagram is adapted from the article by P.M. Borodin et al. (2022) published in "Chromosome Research" under the CC-BY-SA 4.0 license. (modified).

the differentiation of mammalian embryonic cells (Uhlén et al., 2015). The *bicc1* gene encodes an RNA-binding protein that modulates protein translation during embryonic development. The product of the *trim71* gene binds to microRNAs and supports the growth and proliferation of embryonic stem cells. It is also involved in cell cycle control. The gametolog of the *puf60* gene shows signs of recent positive selection. It encodes a protein playing an important role in multiple nuclear processes, including pre-mRNA splicing and transcriptional regulation (Kinsella et al., 2019). These data clearly indicate the functional significance of at least some of the gametologs. However, some genes in the GRC of the blue tit and many genes in both species of nightingales have undergone pseudogenization and presumably lost their functional significance (Schlebusch et al., 2022; Mueller et al., 2023).

Interestingly, the duplication of genes from the somatic genome into GRC did not lead to the loss of their "original" copies. Moreover, in the blue tit, a positive correlation in the

number of copies was found between gametologs and their somatologs ($r = 0.35$, $P = 0.001$) (Mueller et al., 2023).

Thus, GRCs of all four investigated species contain copies of A-chromosome genes, many of which control the development and functioning of germ cells.

Sexual dimorphism in the number of GRCs in the germ cells and its behavior in the meiotic prophase

In almost all male pachytene cells (with rare exceptions, which I will discuss below), at least one GRC is present. It appears as a univalent with a twisted, sometimes fragmented, single axial element of the synaptonemal complex surrounded by chromatin, intensely labeled with anticentromere antibodies. The univalent does not undergo synapsis and recombination neither with itself nor with A-chromosomes. In almost all female pachytene cells (with rare exceptions, which I will discuss below), two GRCs are present. They synapse along their entire length, forming bivalents, and undergo recombination. The only difference between GRC bivalents and A-chromosome bivalents is the distribution of the recombination sites. There are fewer recombination cites on GRC bivalents compared to A-chromosome bivalents of similar size, and they occupy more distal positions on the GRC bivalents (Pigozzi, Solari, 1998, 2005; Torgasheva et al., 2019, 2021; Malinovskaya et al., 2020).

With an increase in the number of species and individuals studied, and the use of molecular probes, many exceptions were identified. Five females (one zebra finch and four sand martins) out of 50 individuals of seven species examined had one GRC (instead of two) in all their pachytene cells (Borodin et al., 2022). Four females of the great tit out of seven examined were mosaic for the number of GRCs: one GRC was observed in a small fraction of their cells (from 2 to 26 %), while the rest contained two GRCs (Torgasheva et al., 2021).

Among 76 males of 26 investigated species, no individuals with more than one GRC in all germ cells have been found so far. However, nine males (seven pale martins, one great tit, one pied flycatcher, and one black-headed munia) exhibited mosaicism in the number of GRCs (Borodin et al., 2022). Some germ cells of mosaic males contained two or even three GRCs (Malinovskaya et al., 2020). The male black-headed munia was mosaic not only in the number but also in the size of GRCs. Some of its cells contained one macro-GRC and one micro-GRC (Sotelo-Muñoz et al., 2022).

Sexual dimorphism in the meiotic behavior of GRC is maintained even in the germ cells, which contain atypical numbers of GRC for their respective sexes. The two axial elements of GRC in male pachytene cells show incomplete synapsis and extremely rare recombination (Malinovskaya et al., 2020). However, even in the case of such recombination occurring in male meiosis, it would not play a role in the evolution of the genetic composition of the GRC due to the extremely low chance of GRC transmission through males (Pei et al., 2022).

The almost exclusive transmission of the GRC through the maternal lineage, coupled with the fact that recombination in most regions of the GRC is suppressed, suggests that the GRC in songbirds represents a new genomic element subject to the action of the Meller's ratchet, in addition to the sufficiently gene-rich W chromosome and the mitochondrial genome. Non-recombining genomic elements are characterized by an exceptionally high rate of fixation of point and structural mutations (Gabriel et al., 1993). It can be hypothesized that the GRC should evolve at a high rate and accelerate the evolution of the entire suborder of songbirds.

Evolution of the GRC

The rate of GRC evolution can be assessed by the degree of divergence in the genetic composition of the GRC among different bird species. A rough estimation of the degree of homology between GRCs of several bird species was obtained using reciprocal fluorescent *in situ* hybridization (FISH). This method revealed an astonishingly low degree of homology between GRCs of different species, which could indicate a rapid evolution of the genetic composition of the GRC (Torgasheva et al., 2019).

The initial results of the genomic analysis of GRCs from different species confirm the assumption of an exceptionally high rate of GRC evolution. The GRCs of the western nightingale (*L. megarhynchos*) and the common nightingale (*L. luscinia*), which have undergone independent evolution for only 1.8 million years, show considerable divergence. Among the 585 gametologs of the western nightingale and the 406 gametologs of the common nightingale, only 192 of them are shared. Among them, only 25 are shared with the GRC of the zebra finch. In other words, only one-third of the identified gametologs is inherited from a common ancestor of the closely related nightingale species. For example, nearly half of the GRC of the common nightingale consists of a large, albeit fragmented, segment homologous to the undivided segment of A-chromosome 2. This genetic material has not been identified in the western nightingale at all (Schlebusch et al., 2023).

Segmental copying of A-chromosome regions into the GRC, followed by the dispersal of these regions within the GRC, is observed in all species that have been sufficiently studied. For example, a major portion of the GRC of the zebra finch consists of dispersed sequences homologous to a segment from the short arm of chromosome 3 (Itoh et al., 2009; Torgasheva et al., 2019). The fragments from the long arm of the same chromosome and from one of the microchromosomes were found to be homologous to the GRC sequences of the siskin. The GRC of the sand martin contains material from chromosomes 4 and W, while the GRC of the great tit is partly homologous to one of the microchromosomes (Torgasheva et al., 2019).

The mechanisms of copying and dispersing fragments of A-chromosomes in the GRC remain unknown. Recombination between the GRC and A-chromosomes is unlikely.

A.A. Torgasheva et al. (2019, 2021) and L.P. Malinovskaya et al. (2020) did not observe in females of four studied songbird species any ectopic contacts between GRC and A-chromosomes, which could lead to recombination and/or conversion. They did not detect a self-synapsis of the GRC synaptonemal complexes, which could lead to deletions, duplications, and dispersal of sequences within the GRC, either.

The remarkable range of inter-species variability in size and heterogeneity in the genetic composition of GRC indicate an extraordinary evolutionary fluidity of this remarkable chromosome. Gametologs constantly accumulate point mutations, gradually diverging from their somatologs and gametologs of other species. At the same time, different gametologs constantly emerge within the GRCs of different species by copying from different regions of different A-chromosomes. Some gametologs amplify in the GRC, increasing its size, while the deletion of others (or the same ones?) results in its reduction. These processes of growth and contraction of the GRC affect different regions in different species, enhancing the divergence of their GRCs.

The distribution of GRC among birds (Torgasheva et al., 2019) and the estimates of the divergence time between gametologs and somatologs of the zebra finch (Biederman et al., 2018; Kinsella et al., 2019) unequivocally indicate a monophyletic origin of GRC in the songbirds.

A.A. Torgasheva et al. (2019) hypothesized that the first GRC could have arisen in the genome of ancestral songbirds through a trisomy of one of the microchromosomes. The ancient nature of GRC in songbirds and its extraordinary genetic fluidity leave no hope of finding A-chromosome paralogs, which would be the last common ancestor of GRC in all songbirds.

An analysis of B-chromosomes in different species suggests that they are derived from fragments of A-chromosomes containing the centromere, which arise during chromosomal rearrangements in the germ line (Camacho et al., 2000; Rubtsov, Borisov, 2018; Poignet et al., 2021). Interchromosomal rearrangements are fixed in bird evolution much less frequently than in other vertebrate taxa. The diploid number ($2n$) of the absolute majority of bird species varies within a very narrow limit: 80 ± 2, and most of the chromosomes in modern birds are syntenic to the reptilian chromosomes (Warren et al., 2010; Griffin, Burt, 2014; Damas et al., 2018). Intrachromosomal rearrangements, particularly inversions, play a significant role in the evolution of bird karyotypes, sometimes acting as one of the mechanisms of speciation (Hooper, Price, 2017; Bravo et al., 2021). Recombination within the inversion loops in heterozygotes for inversion may lead to the formation of chromosome fragments containing functional centromeres. One of such fragments could have become the ancestor of GRC in songbirds.

It can be hypothesized that the proto-GRC became a chromosome restricted to germ cells at the earliest stages of its evolution. Many B-chromosomes exhibit genotaxis, accumulating in generative tissue and being deficient in somatic cells (Camacho, 2022). The evolutionary significance of this phenomenon is evident: it reduces the pressure of natural selection on genes localized on these chromosomes. However, the mechanisms underlying this phenomenon remain unknown.

A.A. Torgasheva et al. (2019) suggested that the proto-GRC might already contain multiple copies of somatic genes controlling the development and functioning of reproductive organs. Therefore, it could be selected for because it provided a higher dosage of these genes. This suggestion appears plausible given the extreme economy of the avian genome (Griffin, Burt, 2014). However, it is more likely that at the time of its origin, the GRC was a typical B chromosome, i. e., an efficient parasite that ensured its transmission without considering the host's interests. Useful genes were probably copied into the GRC at later stages of its evolution. The few B-chromosome variants that persist in host genomes for extended periods and are widespread in populations become gradually "domesticated" acquiring properties beneficial to the hosts (Johnson Pokorná, Reifová, 2021).

The question is what these properties are. By analogy with other examples of programmed DNA elimination discussed above, it can be hypothesized that the GRC may contain unique or amplified copies of the genes controlling early development that exhibit antagonistic pleiotropy. Such genes may be necessary for early development, but their expression in later stages of life can be dangerous. In that case, the GRC can be considered the genetic equivalent of a startup disk or a boot drive. Supporting this hypothesis is the discovery of the evolutionarily ancient gametologs *bicc1*, *trim71*, and *puf60*, the products of which may play important roles in embryonic cell differentiation and cell cycle control (Kinsella et al., 2019).

Another class of genes that could have been evolutionarily advantageous when additional copies were created within the GRC are genes involved in the control of development and proliferation of germ cells (e. g., *prdm1* and *BMP15* in the zebra finch and the blue tit), as well as in the regulation of synapsis and chromosome recombination (*SIX6OS1*, *SYCE2*, *SYCP1*, *TEX12*, *RNF212B* in the blue tit and the nightingale) (Mueller et al., 2023).

It has been suggested that the GRC may participate in sex determination and serve as the basis for a new system (where GRC/0 = X/0 – males, GRC/GRC = X/X – females), which has already emerged or is emerging in the songbirds on top of the typical ZZ-ZW system found in all birds (Stöck et al., 2021). However, I find this proposal highly questionable. As I showed above, females with a single GRC exist and appear to develop normally, while certain males possess two GRCs in a significant portion of their germ cells.

The suppression of recombination between the GRCs, along with the extensive trafficking between the GRCs and A-chromosomes, should lead to rapid incompatibilities between the genomes of closely related species. Following the Dobzhansky–Muller model (Orr, Turelli, 2001), it can

be hypothesized that within each species, the emerging GRC variants are tested for compatibility with A-chromosomes, which already exhibit considerable homology due to constant trafficking. We have already discussed that the mechanisms underlying the massive trafficking of genes from A-chromosomes to the GRC remain unknown.

Furthermore, we do not know the extent to which this trafficking is unidirectional or if there is a reverse movement from the GRC to A-chromosomes. If the trafficking occurs in both directions, the GRC may serve as a generator and incubator of new A-genes. If the trafficking is unidirectional, bird species should rapidly diverge into matrilines based on GRC composition. This could lead to genetic incompatibilities not only between individuals from geographically isolated populations but also within populations, offering broad opportunities for allopatric and sympatric speciation.

Passerines represent the most species-rich suborder of birds, encompassing 6,500 species out of the known 10,500 bird species. All passerines, and only they, possess the GRC. Could this be a contributing factor to their high species number and diversity?

References

Biederman M.K., Nelson M.M., Asalone K.C., Pedersen A.L., Saldanha C.J., Bracht J.R. Discovery of the first germline-restricted gene by subtractive transcriptomic analysis in the zebra finch, *Taeniopygia guttata*. *Curr. Biol.* 2018;28(10):1620-1627. DOI 10.1016/ j.cub.2018.03.067.

Borodin P., Chen A., Forstmeier W., Fouché S., Malinovskaya L., Pei Y., Reifová R., Ruiz-Ruano F.J., Schlebusch S.A., Sotelo-Muñoz M., Torgasheva A., Vontzou N., Suh A. Mendelian nightmares: the germline-restricted chromosome of songbirds. *Chromosom. Res.* 2022;30(2-3):255-272. DOI 10.1007/s10577-022-09688-3.

Boveri T. Ueber Differenzierung der Zellkerne wahrend der Furchung des Eies von *Ascaris megalocephala*. *Anat. Anz.* 1887;2: 688-693.

Bravo G.A., Schmitt C.J., Edwards S.V. What have we learned from the first 500 avian genomes? *Annu. Rev. Ecol. Evol. Syst.* 2021;52:611-639. DOI 10.1146/annurev-ecolsys-012121-085928.

Camacho J.P.M. Non-Mendelian segregation and transmission drive of B chromosomes. *Chromosom. Res.* 2022;30(2-3):217-228. DOI 10.1007/s10577-022-09692-7.

Camacho J.P., Sharbel T.F., Beukeboom L.W. B-chromosome evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2000;355(1394):163-178. DOI 10.1098/rstb.2000.0556.

Chen J., Birchler J.A., Houben A. The non-Mendelian behavior of plant B chromosomes. *Chromosome Res.* 2022;30(2):229-239. DOI 10.1007/s10577-022-09687-4.

D'Ambrosio U., Alonso-Lifante M.P., Barros K., Kovařík A., Mas de Xaxars G., Garcia S. B-chrom: a database on B-chromosomes of plants, animals and fungi. *New Phytol.* 2017;216(3):635-642. DOI 10.1111/nph.14723.

Damas J., Kim J., Farré M., Griffin D.K., Larkin D.M. Reconstruction of avian ancestral karyotypes reveals differences in the evolutionary history of macro- and microchromosomes. *Genome Biol.* 2018; 19(1):155. DOI 10.1186/s13059-018-1544-8.

Dedukh D., Krasikova A. Delete and survive: strategies of programmed genetic material elimination in eukaryotes. *Biol. Rev.* 2022;97(1): 195-216. DOI 10.1111/brv.12796.

del Priore L., Pigozzi M.I. Histone modifications related to chromosome silencing and elimination during male meiosis in Bengalese finch. *Chromosoma*. 2014;123(3):293-302. DOI 10.1007/s00412-014-0451-3.

Dröscher A. Images of cell trees, cell lines, and cell fates: the legacy of Ernst Haeckel and August Weismann in stem cell research. *Hist. Philos. Life Sci.* 2014;36(2):157-186. DOI 10.1007/s40656-014-0028-8.

Gabriel W., Lynch M., Burger R. Muller's ratchet and mutational meltdowns. *Evolution (N. Y).* 1993;47(6):1744-1757. DOI 10.2307/ 2410218.

Goday C., Pigozzi M.I. Heterochromatin and histone modifications in the germline-restricted chromosome of the zebra finch undergoing elimination during spermatogenesis. *Chromosoma*. 2010;119(3): 325-336. DOI 10.1007/s00412-010-0260-2.

Graphodatsky A.S., Perelman P.L., O'Brien S.J. (Eds.) Atlas of Mammalian Chromosomes. John Wiley & Sons, Inc., 2020. DOI 10.1002/ 9781119418061.

Griffin D., Burt D.W. All chromosomes great and small: 10 years on. *Chromosome Res.* 2014;22(1):1-6. DOI 10.1007/s10577-014-9413-0.

Hanlon S.L., Larracuente A.M. When it comes to genetics, cheaters do prosper. *Chromosome Res.* 2022;30(2):137-139. DOI 10.1007/s10577-022-09705-5.

Hodson C.N., Ross L. Evolutionary perspectives on germline-restricted chromosomes in flies (Diptera). *Genome Biol. Evol.* 2021;13(6): 1-19. DOI 10.1093/gbe/evab072.

Hooper D.M., Price T.D. Chromosomal inversion differences correlate with range overlap in passerine birds. *Nat. Ecol. Evol.* 2017;1(10): 1526-1534. DOI 10.1038/s41559-017-0284-6.

Houben A. B chromosomes – a matter of chromosome drive. *Front. Plant Sci.* 2017;8:210. DOI 10.3389/fpls.2017.00210.

Houben A., Banaei-Moghaddam A.M., Klemme S., Timmis J.N. Evolution and biology of supernumerary B chromosomes. *Cell. Mol. Life Sci.* 2014;71(3):467-478. DOI 10.1007/s00018-013-1437-7.

Itoh Y., Kampf K., Pigozzi M.I., Arnold A.P. Molecular cloning and characterization of the germline-restricted chromosome sequence in the zebra finch. *Chromosoma*. 2009;118(4):527-536. DOI 10.1007/s00412-009-0216-6.

Johnson Pokorná M., Reifová R. Evolution of B chromosomes: from dispensable parasitic chromosomes to essential genomic players. *Front. Genet.* 2021;12:1-11. DOI 10.3389/fgene.2021.727570.

Kinsella C.M., Ruiz-Ruano F.J., Dion-Côté A.-M., Charles A.J., Gossmann T.I., Cabrero J., Kappei D., Hemmings N., Simons M.J.P., Camacho J.P.M., Forstmeier W., Suh A. Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* 2019;10(1):5468. DOI 10.1038/s41467-019-13427-4.

Maderspacher F. Theodor Boveri and the natural experiment. *Curr. Biol.* 2008;18(7):R279-R286. DOI 10.1016/j.cub.2008.02.061.

Malinovskaya L.P., Zadesenets K.S., Karamysheva T.V, Akberdina E.A., Kizilova E.A., Romanenko M.V., Shnaider E.P., Scherbakova M.M., Korobitsyn I.G., Rubtsov N.B., Borodin P.M., Torgasheva A.A. Germline-restricted chromosome

(GRC) in the sand martin and the pale martin (Hirundinidae, Aves): synapsis, recombination and copy number variation. *Sci. Rep.* 2020;10(1):1058. DOI 10.1038/s41598-020-58032-4.

Malinovskaya L.P., Slobodchikova A.Y., Grishko E.O., Pristyazhnyuk I.E., Torgasheva A.A., Borodin P.M. Germline-restricted chromosomes and autosomal variants revealed by pachytene karyotyping of 17 avian species. *Cytogenet. Genome Res.* 2022;162(3):148-160. DOI 10.1159/000524681.

Mueller J.C., Schlebusch S.A., Pei Y., Poignet M., Vontzou N., Ruiz-Ruano F.J., Albrecht T., Reifová R., Forstmeier W., Suh A., Kempenaers B. Micro germline-restricted chromosome in blue tits: Evidence for meiotic functions. *Mol. Biol. Evol.* 2023;40(5):msad096. DOI 10.1093/molbev/msad096.

Orr H.A., Turelli M. The evolution of postzygotic isolation: accumulating Dobzhansky–Muller incompatibilities. *Evol. Int. J. Org. Evol.* 2001;55:1085-1094. DOI 10.1111/j.0014-3820.2001.tb00628.x.

Pei Y., Forstmeier W., Ruiz-Ruano F.J., Mueller J.C., Cabrero J., Camacho J.P.M., Alché J.D., Franke A., Hoeppner M., Börno S., Gessara I., Hertel M., Teltscher K., Knief U., Suh A., Kempenaers B. Occasional paternal inheritance of the germline-restricted chromosome in songbirds. *Proc. Natl. Acad. Sci. USA.* 2022;119(4): e2103960119. DOI 10.1073/pnas.2103960119.

Pigozzi M.I., Solari A.J. Extreme axial equalization and wide distribution of recombination nodules in the primitive ZW pair of *Rhea americana* (Aves, Ratitae). *Chromosome Res*. 1997;5(6):421-428. DOI 10.1023/a:1018404610973.

Pigozzi M.I., Solari A.J. Germ cell restriction and regular transmission of an accessory chromosome that mimics a sex body in the zebra finch, *Taeniopygia guttata*. *Chromosome Res*. 1998;6(2):105-113. DOI 10.1023/A:1009234912307.

Pigozzi M.I., Solari A.J. The germ-line-restricted chromosome in the zebra finch: Recombination in females and elimination in males. *Chromosoma*. 2005;114(6):403-409. DOI 10.1007/s00412-005-0025-5.

Poignet M., Johnson Pokorná M., Altmanová M., Majtánová Z., Dedukh D., Albrecht T., Reif J., Osiejuk T.S., Reifová R. Comparison of karyotypes in two hybridizing passerine species: conserved chromosomal structure but divergence in centromeric repeats. *Front. Genet.* 2021;12:76898. DOI 10.3389/fgene.2021.768987.

Rahn M.I., Solari A.J. Recombination nodules in the oocytes of the chicken, *Gallus domesticus*. *Cytogenet. Genome Res.* 1986;43(3-4): 187-193. DOI 10.1159/000132319.

Rubtsov N., Borisov Y. Sequence composition and evolution of mammalian B chromosomes. *Genes (Basel).* 2018;9(10):490. DOI 10.3390/genes9100490.

Schlebusch S.A., Rídl J., Poignet M., Ruiz-Ruano F.J., Reif J., Pajer P., Pačes J., Albrecht T., Suh A., Reifová R. Rapid gene content turnover on the germline-restricted chromosome in songbirds. *Nat. Commun.* 2023;14(1):4579. DOI 10.1038/s41467-023-40308-8.

Schoenmakers S., Wassenaar E., Laven J.S.E., Grootegoed J.A., Baarends W.M. Meiotic silencing and fragmentation of the male germline restricted chromosome in zebra finch. *Chromosoma*. 2010; 119(3):311-324. DOI 10.1007/s00412-010-0258-9.

Smith J.J. Programmed DNA elimination: Keeping germline genes in their place. *Curr. Biol.* 2018;28(10):R601-R603. DOI 10.1016/j.cub. 2018.03.057.

Smith J.J., Timoshevskaya N., Ye C., Holt C., Keinath M.C., Parker H.J., Cook M.E., Hess J.E., Narum S.R., Lamanna F., Kaessmann H., Timoshevskiy V.A., Waterbury C.K.M., Saraceno C., Wiedemann L.M., Robb S.M.C., Baker C., Eichler E.E., Hockman D., Sauka-Spengler T., Yandell M., Krumlauf R., Elgar G., Amemiya C.T. The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nat. Genet.* 2018;50(2):270-277. DOI 10.1038/s41588-017-0036-1.

Smith J.J., Timoshevskiy V.A., Saraceno C. Programmed DNA elimination in vertebrates. *Annu. Rev. Anim. Biosci.* 2021;9(1):173-201. DOI 10.1146/annurev-animal-061220-023220.

Solari A.J. Equalization of Z and W axes in chicken and quail oocytes. *Cytogenet. Genome Res.* 1992;59(1):52-56.

Solari A.J., Moses M.J. The structure of the central region in the synaptonemal complexes of hamster and cricket spermatocytes. *J. Cell Biol*. 1973;56(1):145-152.

Solari A.J., Pigozzi M.I. Recombination nodules and axial equalization in the ZW pairs of the Peking duck and the guinea fowl. *Cytogenet. Cell Genet*. 1993;64(3-4):268-272.

Sotelo-Muñoz M., Poignet M., Albrecht T., Kauzál O., Dedukh D., Schlebusch S.A., Janko K., Reifová R. Germline-restricted chromosome shows remarkable variation in size among closely related passerine species. *Chromosoma*. 2022;131(1-2):77-86. DOI 10.1007/s00412-022-00771-6.

Stöck M., Kratochvíl L., Kuhl H., Rovatsos M., Evans B.J., Suh A., Valenzuela N., Veyrunes F., Zhou Q., Gamble T., Capel B., Schartl M., Guiguen Y. A brief review of vertebrate sex evolution with a pledge for integrative research: towards “sexomics”. *Philos. Trans. R. Soc. B Biol. Sci.* 2021;376(1832):20200426. DOI 10.1098/rstb.2020.0426.

Streit A., Wang J., Kang Y., Davis R.E. Gene silencing and sex determination by programmed DNA elimination in parasitic nematodes. *Curr. Opin. Microbiol.* 2016;32:120-127. DOI 10.1016/j.mib.2016. 05.012.

Torgasheva A.A., Malinovskaya L.P., Zadesenets K.S., Karamysheva T.V., Kizilova E.A., Akberdina E.A., Pristyazhnyuk I.E., Shnaider E.P., Volodkina V.A., Saifitdinova A.F., Galkina S.A., Larkin D.M., Rubtsov N.B., Borodin P.M. Germline-restricted chromosome (GRC) is widespread among songbirds. *Proc. Natl. Acad. Sci. USA*. 2019; 116(24):11845-11850. DOI 10.1073/pnas.1817373116.

Torgasheva A., Malinovskaya L., Zadesenets K., Shnaider E., Rubtsov N., Borodin P. Germline-restricted chromosome (GRC) in female and male meiosis of the great tit (Parus major, Linnaeus, 1758). *Front. Genet.* 2021;12:2008. DOI 10.3389/fgene.2021.768056.

Uhlén M., Fagerberg L., Hallström B.M., Lindskog C., Oksvold P., Mardinoglu A., Sivertsson Å., … Forsberg M., Persson L., Johansson F., Zwahlen M., von Heijne G., Nielsen J., Pontén F. Tissue-based map of the human proteome. *Science.* 2015;347(6220):1260419. DOI 10.1126/science.1260419.

Vontzou N., Pei Y., Mueller J., Reifová R, Ruiz-Ruano F. Schlebusch S., Suh A. Songbird germline-restricted chromosome as a potential arena of genetic conflicts. *Curr. Opin. Genet. Dev*. 2023;83:102113. DOI 10.1016/j.gde.2023.102113.

Wang J., Davis R.E. Programmed DNA elimination in multicellular organisms. *Curr. Opin. Genet. Dev.* 2014;27(1):26-34. DOI 10.1016/ j.gde.2014.03.012.

Wang J., Mitreva M., Berriman M., Thorne A., Magrini V., Koutsovoulos G., Kumar S., Blaxter M.L., Davis R.E. Silencing of germline-expressed genes by DNA elimination in somatic cells. *Dev. Cell*. 2012;23(5):1072-1080. DOI 10.1016/j.devcel.2012.09.020.

Warren W.C., Clayton D.F., Ellegren H., Arnold A.P., Hillier L.W., Kunstner A., Searle S., … Graves T., Fulton L., Nelson J., Chinwalla A., Hou S., Mardis E.R., Wilson R.K. The genome of a songbird. *Nature*. 2010;464(7289):757-762. DOI 10.1038/nature08819.

Watson C.M., Margan S.H., Johnston P.G. Sex-chromosome elimination in the bandicoot Isoodon macrourus using Y-linked markers. *Cytogenet. Genome Res.* 1998;81(1):54-59. DOI 10.1159/000015008.

Weissman A. Prof. Weismann's theory of heredity. *Nature*. 1890; 41(1058):317-323. DOI 10.1038/041317g0.

Weissman A. The Germ-Plasm: A Theory of Heredity. New York: Charles Scribner's Sons, 1893.

ORCID ID

P.M. Borodin orcid.org/0000-0002-6717-844X

Acknowledgements. The study was supported by the Russian Science Foundation, grant 23-14-00182.
The author expresses sincere gratitude to A.A. Torgasheva and L.P. Malinovskaya for assistance in preparing the article.

Conflict of interest. The author declares no conflict of interest.

Received May 25, 2023. Revised June 30, 2023. Accepted June 30, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):651-661
DOI 10.18699/VJGB-23-76

Original Russian text https://vavilovj-icg.ru/

The expression profile of genes associated with behavior, stress, and adult neurogenesis along the hippocampal dorsoventral axis in tame and aggressive foxes

Yu.V. Alexandrovich[1], E.V. Antonov[1, 2], S.G. Shikhevich[1], A.V. Kharlamova[1], L.V. Meister[1], Y.V. Makovka[1], D.V. Shepeleva[1], R.G. Gulevich[1] ✉, Yu.E. Herbeck[1, 3]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Sirius University of Science and Technology, Scientific Center for Translational Medicine, Sochi, Russia
[3] Koret School of Veterinary Medicine, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel
✉ gulevich@bionet.nsc.ru

Abstract. The hippocampus plays the key role in stress response regulation, and stress response appears to be weakened in domesticated animals compared to their wild relatives. The hippocampus is functionally heterogeneous along its dorsoventral axis, with its ventral compartment being more closely involved in stress regulation. An earlier series of experiments was conducted with a unique breeding model of animal domestication, the farm silver fox (*Vulpes vulpes*), which included tame, aggressive, and unselected animals. A decrease in many indices of the hypothalamic–pituitary–adrenal activity was observed in tame animals. Also, adult hippocampal neurogenesis was more intense in tame foxes, and this fact may relate to reduced stress levels in this experimental population of foxes. Nevertheless, the molecular mechanisms responsible for the reduced stress response in tame animals remain obscure. In this study, serum cortisol levels and the mRNA levels of 13 genes in the dorsal and ventral hippocampus have been measured and compared in tame, aggressive, and unselected foxes. At the current stage of domestication, stress-induced cortisol levels in tame, aggressive, and unselected animals differ significantly from each other: tame foxes show the lowest levels, and aggressive ones, the highest. Twelve genes tested demonstrate significant gene expression differences between the dorsal and ventral hippocampi. These differences are mainly consistent with those found in rodents and humans. In tame foxes, significantly elevated mRNA levels were recorded for several genes: *CYP26B1* for cytochrome P450 26B1 and *ADRA1A* for α_{1A} adrenergic receptor in the dorsal hippocampus, whereas the level of *NR3C2* mRNA for mineralocorticoid receptor was higher in the ventral. It is presumed that these genes constitute an important part of the mechanism reducing stress induced by contacts with humans and contribute to linking stress regulation with adult neurogenesis in tame foxes and domesticated animals in general.
Key words: tame behavior; aggression; domestication; silver fox; cortisol.

For citation: Alexandrovich Yu.V., Antonov E.V., Shikhevich S.G., Kharlamova A.V., Meister L.V., Makovka Y.V., Shepeleva D.V., Gulevich R.G., Herbeck Yu.E. The expression profile of genes associated with behavior, stress, and adult neurogenesis along the hippocampal dorsoventral axis in tame and aggressive foxes. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):651-661. DOI 10.18699/VJGB-23-76

Профиль экспрессии генов, связанных с регуляцией стресса, поведения и нейрогенеза, вдоль дорзовентральной оси в гиппокампе у взрослых ручных и агрессивных лисиц

Ю.В. Александрович[1], Е.В. Антонов[1, 2], С.Г. Шихевич[1], А.В. Харламова[1], Л.В. Мейстер[1], Ю.В. Маковка[1], Д.В. Шепелева[1], Р.Г. Гулевич[1] ✉, Ю.Э. Гербек[1, 3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Научно-технологический университет «Сириус», Научный центр трансляционной медицины, Сочи, Россия
[3] Школа ветеринарной медицины им. Й. Корета, Факультет сельского хозяйства, продовольственных ресурсов и окружающей среды им. Р.Г. Смита, Еврейский университет в Иерусалиме, Реховот, Израиль
✉ gulevich@bionet.nsc.ru

Аннотация. Гиппокамп является ключевой структурой в регуляции стресс-ответа, который, по-видимому, снижен у домашних животных по сравнению с их дикими сородичами. Известно, что гиппокамп функционально неоднороден вдоль дорзовентральной оси, и в регуляции стресса в большей мере участвует вентральная часть. В серии экспериментов на уникальной селекционной модели одомашнивания животных – серебристо-черной лисице (*Vulpes vulpes*), включающей ручных, агрессивных и неселекционированных животных,

ранее было показано снижение активности гипоталамо-гипофизарно-надпочечниковой системы во многих звеньях. Кроме того, известно, что уровень нейрогенеза в гиппокампе повышен у взрослых ручных лисиц, что может быть взаимосвязано со снижением уровня стресса. Тем не менее молекулярно-генетические механизмы снижения стресс-ответа у доместицированных животных по-прежнему не ясны. В настоящей работе выполнено сравнение мРНК 13 генов в дорзальном и вентральном гиппокампе и проведен анализ кортизола в крови у ручных, агрессивных и неселекционированных лисиц. Установлено, что на данном этапе доместикации стресс-индуцированный уровень кортизола у ручных, агрессивных и неселекционированных животных достоверно отличается друг от друга, причем у ручных животных он самый низкий, а у агрессивных – самый высокий. Выявлены достоверные различия в экспрессии 12 генов между дорзальной и вентральной частями гиппокампа, что в большинстве случаев соответствует аналогичным различиям, найденным у грызунов и человека. У ручных лисиц обнаружен достоверно повышенный уровень в дорзальном гиппокампе мРНК генов цитохрома P450 26B1 (*CYP26B1*) и адренергического рецептора α_{1A} (*ADRA1A*), а в вентральном гиппокампе – мРНК гена минералокортикоидного рецептора (*NR3C2*). Эти гены могут быть важной частью механизма снижения стресса по отношению к человеку и взаимосвязи регуляции стресса и нейрогенеза у взрослых ручных лисиц в частности и доместицированных животных вообще.
Ключевые слова: ручное поведение; агрессия; доместикация; серебристо-черные лисицы; кортизол.

Introduction

The hippocampus is an important brain region involved in the regulation of stress response, learning, spatial memory, social recognition, and memory consolidation. Hippocampus sizes in mammals and birds and adult neurogenesis in mammals are likely to correlate with the capacity for spatial orientation and memory. Also, changes in hippocampus morphology may be associated with adaptive evolution (Jacobs et al., 1990; Jacobs, Spencer, 1994; Rehkämper et al., 2008; Croston et al., 2015; Sonnenberg et al., 2019). However, other scientists presume that neurogenesis rate and hippocampus size, albeit adaptive, are not related to memory or spatial orientation (Lipp, 2017). The 15-fold increase in neurogenesis rate in adult red foxes as compared to dogs (Amrein, Slomianka, 2010) may be related to spatial memorization, typical of foxes hoarding food (Sklepkovych, Montevecchi, 1996). It is known that the CA2 region takes part in social recognition memory (Tzakis, Holahan, 2019). The CA1 and CA3 volumes in primates appear to be related to social and environmental signals, such as group and home range sizes (Todorov et al., 2019).

The hippocampus is among the key elements of the central regulation of the hypothalamic–pituitary–adrenal (HPA) axis. It is known to be functionally and structurally heterogeneous along its dorsoventral axis. It is believed that the regulation of HPA axis action and stress and emotional responses are governed primarily by the ventral hippocampus, and cognitive functions, by the dorsal (O'Leary, Cryan, 2014; Gulyaeva, 2019). This specialization may be due to the locations of these hippocampus compartments, and, correspondingly, the greater number of dorsal hippocampus projections to the crust and of ventral hippocampus projections to regions belonging to the limbic system (O'Leary, Cryan, 2014).

Along with the lateral ventricle subependymal zone, the hippocampus is a region where neurogenesis occurs constantly, even in adulthood (Ming, Song, 2011). The neurogenesis rate is reduced by stress in most cases, and, vice versa, high neurogenesis rate mitigates the effect of stress on the hippocampus (Levone et al., 2015). Thus, stress and the hippocampus functional response exert reciprocal actions. Apparently, the effect of stress on neurogenesis varies along the dorsoventral axis (O'Leary, Cryan, 2014). This inference may explain the fact that the hippocampus neurogenesis rate in tame foxes, whose stress response is much weaker, is higher than in unselected ones, and the most pronounced differences are recorded in the ventral and intermediate compartments of the hippocampi (Huang et al., 2015). Studies on dogs have shown that differences between the dorsal and ventral hippocampi are at the same level despite considerable variations in the overall neurogenesis rate among individuals (Lowe et al., 2015).

Experiments with hippocampus samples from rats, mice, and humans reveal differences in gene expression between the dorsal and ventral hippocampi. They are likely to reflect the functional and structural heterogeneity along the axis (Cembrowski et al., 2016; Lee et al., 2017; Floriou-Servou et al., 2018; Vogel et al., 2020). However, the expression profiles of these genes along the hippocampus dorsoventral axis have not been studied in other taxa, in spite of their functional and structural features.

These data motivated us to investigate gene expression variation along the dorsoventral axis and seek molecular mechanisms linking neurogenesis and stress by quantitation of mRNAs of 13 genes in the dorsal and ventral hippocampi of silver foxes. This species serves as a model of animal domestication. Its "tame" and "aggressive" populations had been raised by long-term selection for friendly or aggressive attitude to humans, respectively, and foxes not subjected to targeted selection for behavior served as control.

These populations differ significantly in many links of their glucocorticoid stress response and HPA axis activity, which seems to be a common feature of domestic animals (Belyaev, 1979; Price, 2000; Trut et al., 2004, 2009), and in adult neurogenesis in the hippocampus (Huang et al., 2015). Genes of the retinoic acid pathway, associated with neurogenesis activity, which are located in regions presumably affected by the selection, have been found in genome-wide

analysis of the tame foxes (Kukekova et al., 2018; Trut et al., 2021). Also, it has been found that the amount of mRNA of one of these genes, *CYP26B1*, in the dorsal hippocampus of tame foxes differs from aggressive ones (unpublished results). This may be one of the mechanisms altering adult neurogenesis in foxes, which can also modulate stress level, learning, memory, and social behavior. For all that, nothing is known about changes in *CYP26B1* mRNA levels in the ventral hippocampi of foxes with contrasting behaviors, although its elevated expression characterizes the dorsal hippocampus in mice, rats, and humans.

We also measured the dorsal and ventral hippocampus levels of mRNAs of other genes related to neurogenesis, stress, or behavior whose expression is known to vary significantly along the dorsoventral axis in mice, rats, and humans (Vogel et al., 2020) (see Table 1). The genes associated with HPA axis regulation include *NR3C1*, *NR3C2*, and *HSD11B1* for glucocorticoid receptors 1 and 2 and hydroxysteroid 11β-dehydrogenase 1, respectively (de Kloet et al., 2016). *NR2F2* is one of the most reliable markers of the position along the dorsoventral hippocampus axis, and it supposedly acts as a mediator of the transcription activity induced by receptors of glucocorticoids and retinoic acid. This action is likely to be associated with relationships between stress and neurogenesis (de Martino et al., 2004; Vogel et al., 2020). The *ADRA1A* gene for the α_{1A} adrenergic receptor is presumed to act in the regulation of behavior and neurogenesis (Doze et al., 2011; Vogel et al., 2020). *KCND2*, *KCND3*, *CADM2*, and *CPNE2* are associated with K^+- and Ca^{2+}-dependent synaptic and glutamatergic transmissions (Corradini et al., 2014; Truvé et al., 2020; Haddjeri-Hopkins et al., 2021; Xiao et al., 2021), which are likely to play the key role in domestication-driven behavior changes (O'Rourke, Boeckx, 2020; Trut et al., 2021). The expression levels of *TRHR*, on the one hand, and *LCT*, *NTS*, on the other hand, have been used as markers of the ventral and dorsal hippocampi, respectively (Cembrowski et al., 2016; Lee et al., 2017).

Earlier data on the glucocorticoid-mediated stress response in foxes demand re-research at the present phase of selection. After over sixty generations of the selection of Norway rats according to approximately the same criteria as foxes as an alternative experimental domestication model, differences between tame and aggressive animals in the glucocorticoid-mediated stress response vanished. This might result from the adaptation of rats selected to aggression toward humans (Prasolova et al., 2014). Besides, previous studies employed nonsocial restriction stress (Trut et al., 2004, 2009). This procedure is less adequate than the use of social stress in studies of animal–animal and animal–human contacts. It has been shown that restriction stress in rats selected for anxiety-like behavior induces a stronger corticosterone-mediated response than in animals selected for low levels of such behavior, whereas social stress in the resident–intruder test shows quite the opposite (Veenema, Neumann, 2007). Here we studied the stress response to a combined treatment: manual fixation of foxes by a social subject, human. Manual fixation for 15 min was a stress factor for all the three fox populations, because foxes had never been picked in arms during selection and they had had an opportunity to avoid close contacts with humans.

Materials and methods

Experimental animals. Experiments were conducted with three experimental populations of silver foxes (*Vulpes vulpes*): tame, aggressive, and unselected. The first two populations had been selected for friendly and aggressive-fearful reactions to humans, respectively, at the Shared Access Center for Gene Pools of Fur and Farm Animals, Institute of Cytology and Genetics, Novosibirsk, for over 60 years (Belyaev, 1979; Trut et al., 2004, 2009).

Blood was sampled from the vena saphena of 6–7-month-old males prior to experimental stress (manual fixation for 15 min) and immediately after it. Naïve (having experienced no experimental stress) 7–8-month old males were euthanized by injections of 5 % sodium thiopental. Fragments of the dorsal and ventral hippocampi were sampled. All samples were stored at –70 °C. Experimental protocols followed the Guidelines for Accommodation and Care of Laboratory Animals, Species-specific Provisions for Laboratory Predatory Mammals, GOST 33217-2014, and Directive 2010/6106/EU for the Protection of Laboratory Animals.

Chromatography. Blood serum cortisol was assayed by high-performance liquid chromatography with an Agilent 1200 Series LC chromatograph equipped with a diode-array detector and a ZORBAX C18 2 × 150 mm × 5 μm column, as in previous studies (Ovchinnikov et al., 2018). Samples were concentrated by liquid extraction with 1,2-dichloroethane. Elution was done with 30 % aqueous acetonitrile at the rate 1 mL/min. Absorption was measured at $\lambda = 246$ nm. Concentrations were calculated against dexamethasone as an internal standard.

Total RNA isolation and real-time PCR. RNA was isolated with TRI Reagent (Molecular Research Center, Inc.) according to manufacturer's recommendations, as in (Ovchinnikov et al., 2018). Absorption values were measured with a NanoPhotometer N50 (Implen, Germany). For RNA purity assessment, A260/A280 and A260/A230 ratios were calculated. Genomic DNA was removed from RNA samples with a DNase I, RNase-free kit (Thermo Fisher Scientific, Lithuania). cDNA was synthesized in a 20 μL volume with 0.2 μg of DNA-free RNA and a Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Fisher Scientific, Lithuania).

Primers for the genes studied were designed using the online web resource Primer-BLAST (Ye et al., 2012). The sequences are shown in Table 1.

Real-time PCR was conducted in a Roche LightCycler 96 Real-Time PCR System (Roche Diagnostics, Switzerland). The reaction volume 20 μL contained 4 μL of twentyfold

Table 1. Primers for real-time PCR

Gene	Primer sequences
NTS	fNTS-F1: 5′-TGGTGTGCATGATACTTCTGG-3′ fNTS-R1: 5′-ACACTTGCTTTGCTGATCTTTG-3′
NR2F2	fNR2F2-F3: 5′-AGCCAAGGAATGTGTCCAAG-3′ fNR2F2-R3: 5′-CAATTCAGGAACTAAGCGGGA-3′
ADRA1A	fADRA1A-F: 5′-CGTGCTCCAGTCAAGAGTTT-3′ fADRA1A-R: 5′-AAGGTATAGCCCAGGGTGTG-3′
CPNE2	fCPNE2-F6: 5′-CCAGGTCATGTGTTACGACT-3′ fCPNE2-R6: 5′-CCTTTGCTTCTTGGGGTTGA-3′
CYP26B1	fCYP26B1-F1: 5′-TTCTTTGGCCTGGACTCGAA-3′ fCYP26B1-R1: 5′-GGCTAGGCGCAGTTAGAC-3′
CADM2	fCADM2-F3: 5′-GTAGCCATAACAACCAGCCC-3′ fCADM2-R3: 5′-AGAACACAGCGTGACAAATACA-3′
KCND2	fKCND2-F1: 5′-GAAAACCTTCCGCATCCCAAA-3′ fKCND2-R1: 5′-ACATTCTTCCATCTTGGCGTT-3′
KCND3	fKCND3-F6: 5′-CAGAGAGCCGATAAACGCAG-3′ fKCND3-R6: 5′-GTGCAGATAGGCATTGGAGC-3′
TRHR	fTRHR-F1: 5′-GACTCAACCCATCAGAACAAGA-3′ fTRHR-R1: 5′-GGGCATCCATAAAAGGGCAA-3
LCT	fLCT-F1: 5′-AAGAACGGCGATTACAACGA-3′ fLCT-R1: 5′-TGCCATTGATCCTCCTCTTCT-3′
HSD11B1	fHSD11B1-F2: 5′-GCAAGGGGATTGGAGAACAG-3′ fHSD11B1-R2: 5′-GGTGCCAGGAATGTAGTGTG-3′
NR3C1	fNR3C1-F6: 5′-CAAGCTGGGATGAACTTGGA-3′ fNR3C1-R6: 5′-AGTTTCTTGTGACGCTCCTG-3′
NR3C2	fNR3C2-F7: 5′-AAAGGCTACCACAGTCTCCC-3′ fNR3C2-R7: 5′-TCATCGGTCCTCTCTGTAGG-3′
CANX	CANX-F1: 5′-GATGCCCCTGCTAAGATTCC-3′ CANX-R1: 5′-CTTCATCCCAATCCTCTGGC-3′

diluted cDNA, 0.3 μL of primers (10 pmol/μL), 7.4 μL of Milli-Q H_2O, 8 μL of 2.5× reaction mix for Real-Time PCR, and SYBR Green I dye (Mfr. Part No. M-427, Syntol, Russia). Each reaction was performed in two technical replications.

The results were processed by the modified $\Delta\Delta C_t$ method (Livak, Schmittgen, 2001) implemented in GenEx ver.6 software (Multi-D, Sweden). This method allows the reaction efficacy to be estimated. The *CANX* gene for calnexin was used as reference, because its expression is high, little variable among individual foxes, and uniform in the dorsal and ventral hippocampi, as confirmed by analysis with NormFinder software (Andersen et al., 2004). The mean expression of each gene was taken to be an arbitrary unit for the evaluation of relative expression. An additional external reference sample was present in all plates for proper comparison of the results obtained in different plates.

Statistical evaluation. The statistical significance of hormone assays in the experimental groups was assessed by repeated measures factor analysis followed by *post hoc* Fisher's LSD test. The increase in cortisol level was assessed by Student's *t* test.

Real-time PCR results were compared by the Kruskal–Wallis test. Pairwise comparisons were done by the Mann–Whitney test. We applied nonparametric criteria, because the samples did not conform to the Gaussian distribution according to the Kolmogorov–Smirnov test.

Use was made of software packages Statistics 10 (StatSoft, United States) and GenEx ver.6 (Multi-D, Sweden). All differences were considered significant at $p < 0.05$. The results are shown in figures as mean ± SEM.

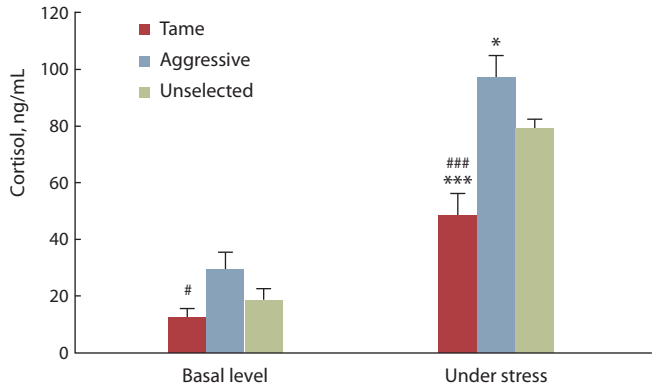

Fig. 1. The basal and stress-induced blood serum cortisol levels in tame, aggressive, and unselected foxes ($n = 11$ in each group).

* $p < 0.05$, *** $p < 0.001$ in comparison with unselected foxes.
\# $p < 0.05$, ### $p < 0.001$ in comparison with aggressive foxes.

Results

Blood serum cortisol in response to stress

Repeated measures factor analysis reveals the effects of genotype ($F_{2.28} = 9.62$, $p < 0.001$) and stress ($F_{2.28} = 179.72$, $p < 0.001$) on blood serum cortisol. The interaction of the *genotype* and *stress* factors was also significant ($F_{2.28} = 9.36$, $p < 0.01$). The basal serum cortisol level in tame foxes was lower than in aggressive (Fig. 1, $p < 0.05$) but did not differ significantly from unselected ones. The increase in serum cortisol level was statistically significant in all genotypes, but it was less pronounced in tame foxes than in aggressive or unselected. Thus, the cortisol level under stress in aggressive and unselected foxes exceeded the value in tame ones ($p < 0.001$, see Fig. 1) and was higher in aggressive foxes than in unselected ($p < 0.05$, see Fig. 1).

Levels of mRNAs in the dorsal and ventral hippocampi of foxes

Quantitative real-time PCR revealed in the dorsal hippocampus significantly higher levels of the genes *HSD11B1*, *CYP26B1*, *CADM2*, *KCND2*, *NR3C1*, *LCT*, and *NR3C2* and in the ventral, *TRHR*, *CPNE2*, *ADRA1A*, and *NR2F2*, whereas *KCND3* showed no difference (Fig. 2, Table 2).

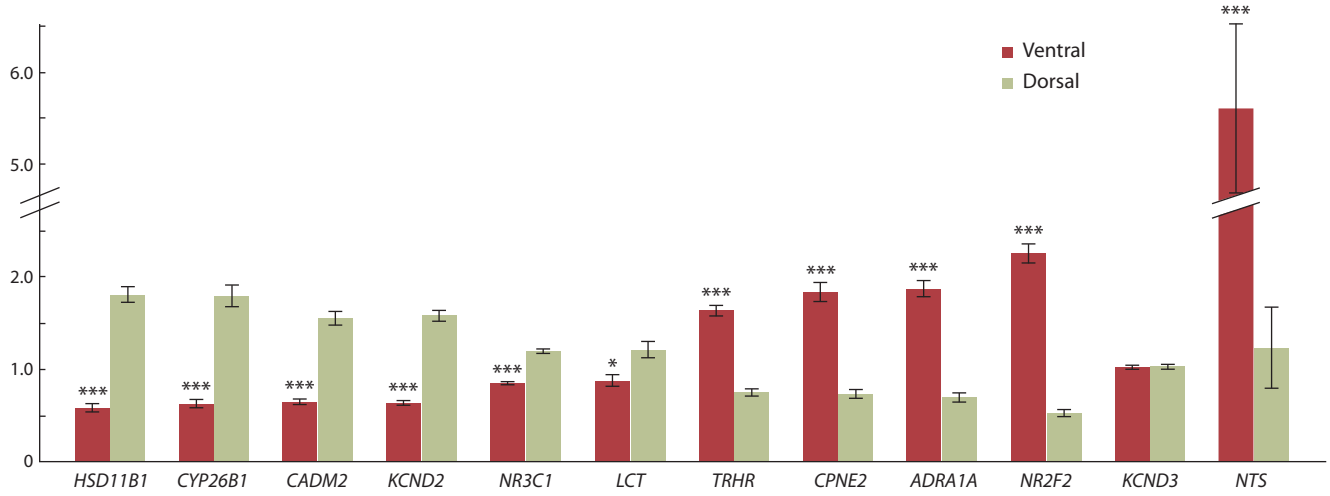

Fig. 2. Differences between the dorsal and ventral hippocampi in relative mRNA levels (n = 18 in each group).

* $p < 0.05$, *** $p < 0.001$ as compared to the dorsal hippocampus.

Table 2. Differences between the dorsal and ventral hippocampi in gene expression in mammals

Species	*HSD11B1*	*CYP26B1*	*CADM2*	*KCND2*	*NR3C1*	*LCT*	*NR3C2*	*TRHR*	*CPNE2*	*ADRA1A*	*NR2F2*	*KCND3*	*NTS*	Reference
Mouse	v (DG) d (CA1)	d (CA3)	d (DG, CA3, CA1)	d (DG, CA3, CA1)	–	d (DG, CA3)	d (DG, CA1)	–	v (CA3, CA1)	v (CA1)	v (CA3, CA1)	v (CA3, CA1)	–	Cembrowski et al., 2016
	d	d	d	d	–	d	d	v	v	v	v	v	v	Floriou-Servou et al., 2018
Rat	–	d	d	d[a]	d[c]	–	d[b]	v	v	–	v	–	d	Lee et al., 2017
Human	v	d	d	–	–	v	–	v	v	v	v	–	d	Vogel et al., 2020
Fox	d	d	d	d	d	d	d	v	v	v	v	–	v	Our data
p	<0.001	<0.001	<0.001	<0.001	<0.001	0.018	0.001	<0.001	<0.001	<0.001	<0.001	>0.05	<0.001	

Note. CA1, CA3, and DG (dentate gyrus) are hippocampus regions; d, mRNA level is higher in the dorsal hippocampus; v, mRNA level is higher in the ventral hippocampus; –, no difference or not known. *p* values are indicated according to the Mann–Whitney test.
[a] 28 and 45 days (growing animals); [b] 28 days; [c] rat *Nr3c1* mRNA according to Kvichansky et al. (2017).

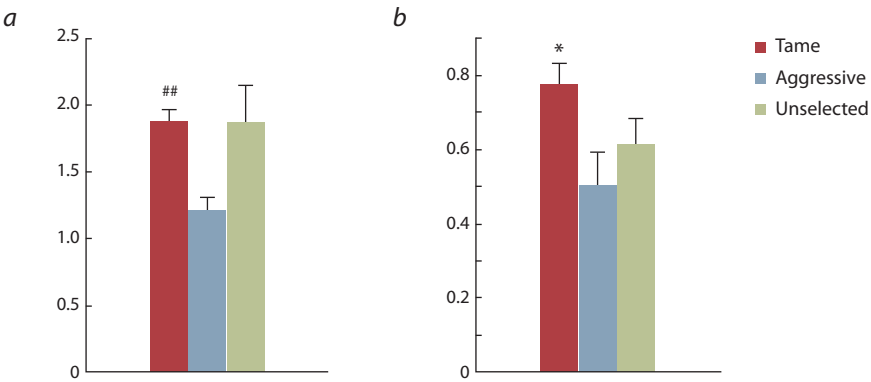

Fig. 3. Relative amounts of *CYP26B1* (*a*) and *ADRA1A* (*b*) mRNA in the dorsal hippocampi of tame (n = 7), aggressive (n = 6), and unselected (n = 6) foxes.

$p < 0.01$ as compared to aggressive foxes; * $p < 0.05$ as compared to aggressive and unselected foxes.

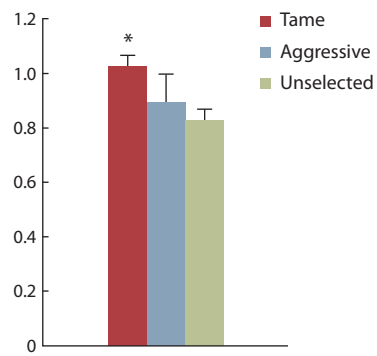

Fig. 4. Relative amounts of *NR3C2* mRNA in the ventral hippocampi of tame (n = 7), aggressive (n = 6), and unselected (n = 6) foxes.

* $p < 0.05$ as compared to unselected foxes.

NTS mRNA levels showed a broad variation among hippocampi of individuals, but the expression in the ventral compartment was always higher (see Fig. 2, Table 2), and so was the neurogenesis rate in a study on dogs reported by Lowe et al. (2015).

Levels of mRNAs in the hippocampi of tame, aggressive, and unselected foxes

We compared the levels of mRNAs indicated in Table 2 in tame, aggressive, and unselected fox groups by the Kruskal–Wallis test and revealed effects of genotype on mRNAs of the genes (1) *CYP26B1* ($H(2, n = 19) = 8.89$; $p = 0.02$) and *ADRA1A* ($H(2, n = 19) = 7.81$; $p = 0.02$) in the dorsal compartment and (2) *NR3C2* for a mineralocorticoid receptor ($H(2, n = 19) = 7.07$; $p = 0.03$) in the ventral compartment. Tame foxes showed a significant increase in the *CYP26B1* mRNA as compared to aggressive animals ($p = 0.003$) (Fig. 3, *a*) and in *ADRA1A* as compared to both aggressive ($p = 0.027$) and unselected ($p = 0.038$) (see Fig. 3, *b*). Also, the *NR3C2* mRNA level in the ventral hippocampus of tame foxes was significantly higher than in unselected ones ($p = 0.011$) (Fig. 4).

Discussion

The presented data point to significant differences between the dorsal and ventral hippocampi in the levels of some mRNAs in foxes, as well as in other species: rats, mice, and humans (Cembrowski et al., 2016; Lee et al., 2017; Floriou-Servou et al., 2018; Vogel et al., 2020). Our results are by and large consistent with literature data on rats and mice (see Table 2). Recent studies of human gene expression (Allen Brain Atlas) demonstrate variation in the expression of about 5,000 genes along the hippocampus dorsoventral axis (Vogel et al., 2020). Some of these genes show linear or near-linear variation; thus, they can be regarded as axial position markers. Others demonstrate nonlinear expression profiles.

The *NR2F2* gene for transcription factor COUP-TFII is one of the markers most precisely indicating axial position. The amount of its mRNA in the ventral hippocampus is much greater than in the dorsal in all studies on rats, mice, and humans and in our study on foxes. Its expression in the adult hippocampus is confined mainly to GABAergic and glutamatergic neurons. The density of GABAergic neurons increases along the dorsoventral axis (Jinno, Kosaka, 2010), and so does *NR2F2* expression. Nevertheless, its expression in the dorsal hippocampus is confined to GABAergic neurons, which is indicative of their high density in the dorsal hippocampus as well (Fuentealba et al., 2010). The different functions of hippocampus compartments may be related to neuron distribution along the dorsoventral axis. It is known that the expression of acetylcholine receptor α_7 (*CHRNA7*) in the hippocampus, also confined to GABAergic neurons, is involved in aggression regulation (Lewis et al., 2018). However, we found no difference in *NR2F2* expression amongst the behavior groups; hence, aggression is controlled by other mechanisms.

Different species demonstrate inverse ratios between mRNA levels of some genes in the dorsal and ventral hippocampus compartments or no variation at all (see Table 2). For example, in our study such genes included *KCND3* and *NTS*. These differences can be explained by nonlinear expression profiles along the dorsoventral axis and putative sampling from nonidentical hippocampus sites in different experiments. In other cases, the species-specific expression of some genes may be related to morphological and functional features of the hippocampus itself. For instance, it is known that hippocampal neurogenesis in foxes considerably surpasses that in many mammals (Amrein, Slomianka, 2010). Morphological and functional features of various species may also stem from their ecology, in particular, spatial behavior in hoarding food (Jacobs et al., 1990; Jacobs, Spencer, 1994; Rehkämper et al., 2008; Amrein, 2015; Crosston et al., 2015; Lipp, 2017; Sonnenberg et al., 2019). It is presumable that the 15-fold hippocampal neurogenesis in red foxes as compared to dogs (Amrein, Slomianka, 2010) is related to this characteristic behavior (Sklepkovych, Montevecchi, 1996).

As the entire hippocampus and, especially, its ventral compartment, is the key region in stress response regulation, we investigated the glucocorticoid-mediated response of foxes at the present stage of selection: tame, aggressive, and unselected. The necessity of studying stress response at different stages of selection has been shown on another model, tame and aggressive Norway rats, which show no significant differences in the glucocorticoid-mediated response at the current stage (Prasolova et al., 2014). The detected significant differences amongst tame, aggressive, and unselected animals are consistent with earlier data (Trut et al., 2009). Thus, we can see that the restraint and combined restraint–emotional stresses induce similar differences in the cortisol-induced stress response in the experimental foxes, and these differences persist at the current stage of selection. However, the unselected animals showed an intermediate level of the glucocorticoid stress response between the tame and aggressive foxes. These discrepancies may be related to both the elevated stress response in the selection for aggressiveness and the unintentional selection of "unselected" foxes towards adaptation to coexistence with humans. Differences can also stem from different experiment designs (restraint stress with limited space in a shed *vs.* manual fixation) and measurement protocols (radioimmunoassay *vs.* HPLC).

It is likely that the weak response to different stress types in tame foxes is the main cause of high adult neurogenesis rate in the hippocampus, as shown in many studies on other species (Levone et al., 2015). Therefore, in search for molecular mechanisms modulating neurogenesis we first considered the levels of mRNAs for glucocorticoid (*NR3C1*, GR) and mineralocorticoid (*NR3C2*, MR) receptors

in the hippocampus. However, we found no differences in the levels of *NR3C1* mRNA amongst animals of different behavior genotypes in neither dorsal nor ventral hippocampus, although some research teams believed that these genes were important in domestication (Oskina et al., 2008; Pörtl, Jung, 2017). It is known that neuron progenitors in the subventricular zone express *NR3C1* but not *NR3C2* (Garcia et al., 2004). Apparently, *NR3C1* activation in these cells is the direct way by which glucocorticoids affect neurogenesis (Saaltink, Vreugdenhil, 2014). It is conceivable that *NR3C1* expression in this subpopulation of hippocampus cells varies amongst foxes differing in behavior.

In contrast, the ventral hippocampi of tame foxes contained more *NR3C2* mRNA than those of aggressive animals. It is likely that the mRNA level in the ventral hippocampi of unselected foxes is intermediate between tame and aggressive, but this difference is below the limits of qPCR accuracy. However, we may expect that in studies of individual splice variants (Three are known in rodents: α, β, and γ.) differences amongst groups in the expression of a particular splice variant will be more pronounced. It has been shown that their expression in cellular stress in a primary culture of cortex cells varies irregularly (Kang et al., 2009). Here we analyze only the total pool of *NR3C2* mRNA, because the fox genome had not been annotated in sufficient detail. It is known that NR3C2 levels in various types of hippocampus cells are different (Le Menuet, Lombès, 2014). Note that the difference was found just in the ventral hippocampus, whose contribution to stress response and emotion regulation is thought to be greater than that of the dorsal (Gulyaeva, 2019).

Studies on rodents demonstrate that elevated MR amount is associated with lower anxiety and active strategy of coping with stress. In females, it is also associated with weaker stress response (Lai et al., 2007; Rozeboom et al., 2007; Kanatsou et al., 2015; de Kloet et al., 2016). In addition, postmortem studies of humans show that depression lowers MR in the frontal (corresponding to rodent ventral) but not occipital (dorsal) hippocampus compartments, with no variation in GR (Medina et al., 2013). Changes in MR expression in rats under stress alter synaptic plasticity in the ventral but not dorsal hippocampus (Maggio, Segal, 2009; O'Leary, Cryan, 2014). It is likely that *NR3C2* expression in neuroglia can increase neurogenesis in the ventral hippocampus under acute stress (Le Menuet, Lombès, 2014; O'Leary, Cryan, 2014). Also, progenitor cell proliferation is lowered in mice with *NR3C2* knockout (Gass et al., 2000), whereas enhanced *NR3C2* expression in the forebrain accelerates progenitor proliferation and increases the population of young neurons in the dentate gyrus (Kanatsou et al., 2017). On the other hand, MR signaling in the hippocampus is involved in the regulation of the start and amplitude of stress response (Ratka et al., 1989; Harris et al., 2013; de Kloet et al., 2016), and stress initiation increases the MR level in the hippocampus (Veenema et al., 2003; de Kloet et al., 2016). The opposite effects of MR and their agonists and antagonists found in various studies may be related to different MR levels in animals at the start of the experiment (de Kloet et al., 2016). We conjecture that the detected high level of mRNA for MR (but not for GR) in the ventral hippocampus of tame foxes is one of the mechanisms that mitigate stress and anxiety in experimental domestication and, probably, indirectly enhance neurogenesis in this compartment. Further studies should be dedicated to the expression of splice variants of MR and their distribution in the hippocampus.

Foxes of different behavior genotypes differed in the contents of *CYP26B1* mRNA in the dorsal hippocampus and *ADRA1A* in the ventral one. As mentioned above, these contents varied along the dorsoventral axis of the hippocampus. The variation in mRNA contents found in the analysis of few samples from the dorsal hippocampi of tame and aggressive foxes by the RNAseq method (unpublished data) also points to the necessity of a comprehensive study of *CYP26B1* expression in hippocampus regions of the three fox populations. The *CYP26B1* gene encodes an enzyme of the cytochrome P450 superfamily. This enzyme catalyzes the degradation of all-trans retinoic acid (atRA), a vitamin A derivative. Changes in *CYP26B1* expression may be associated with different atRA concentrations in the hippocampi of tame and aggressive foxes.

It is known that atRA affects neurogenesis, but its effect looks graphically as an inverted U curve. All-trans retinoic acid deficiency reduces neuron differentiation, whereas higher concentrations enhance neurogenesis by stimulating both proliferation and differentiation of neural stem cells, and still higher atRA concentrations inhibit cell proliferation and affect the cognitive function and behavior (Kane et al., 2010; Hu et al., 2016, 2020; Stoney, McCaffery, 2016; Stoney et al., 2016; Mishra et al., 2018). High *CYP26B1* expression, apparently decreasing the atRA level, is associated with intense neurogenesis in the hippocampus of adult tame rats, earlier demonstrated by Huang et al. (2015). Our results seem to be consistent with the negative effect of atRA on neurogenesis in studies by P. McCaffery's (Stoney et al., 2016) and Zhou's (Hu et al., 2016, 2020) teams. It appears that *CYP26B1* inhibition reduces cell proliferation in the subgranular zone of the hippocampus in mice (Stoney et al., 2016). Probably, the atRA content in the hippocampus of tame foxes is close to the maximum level enhancing neurogenesis. The lower expression of *CYP26B1* in aggressive animals is associated with even higher atRA contents and, probably, lower neurogenesis. The unselected foxes, demonstrating lower neurogenesis than tame ones, seem to be in the middle between tame and aggressive. Note that genome-wide comparisons between village dogs and wolves and between humans and chimpanzees has also demonstrated differences in the atRA system (Theofanopoulou et al., 2017; Pendleton et al., 2018).

The effects of vitamin A and atRA on the HPA axis are complex and controversial. Dexamethasone upregulates the

expression of the *Aldh1a1* gene for an enzyme involved in atRA synthesis (Gil-Ibáñez et al., 2014). Chronic exposure to atRA and other retinoic acid forms enhances HPA axis activity and causes depression (Bremner, McCaffery, 2008; Cai et al., 2015). In particular, atRA induces overexpression of the *Crf*, *Crfr1*, and *Avp* genes in the hypothalamus (Cai et al., 2015). Probably, the lower HPA activity in tame foxes is partly related to the weakening of the atRA system. However, it should be noted that retinoic acid can, in contrast, lower the level of glucocorticoids, and, besides, it can exert an inverse effect on target tissues (Bonhomme et al., 2014; Hélène et al., 2016).

The detected high level of *ADRA1A* mRNA in the dorsal hippocampus of tame foxes is of special interest. This change may reflect the lower anxiety and elevated adult neurogenesis in the hippocampus as a result of lower HPA axis activity in the selection for tame behavior. Although the dorsal hippocampus contributes less to the regulation of stress responses and anxiety, its effects have been described in (Weaver et al., 2004; Gulyaeva, 2019). *ADRA1A* is presumed to play a role in the attention deficit disorder (Elia et al., 2009). The mechanisms mediating the effect of ADRA1A in behavior regulation are poorly understood, but it is known that long-term *ADRA1A* stimulation mitigates depression-like behavior and anxiety, and tricyclic antidepressants increase the ADRA1A receptor density in the forebrain of rodents (Deupree et al., 2007; Doze et al., 2009, 2011). Studies of the subependymal zone of lateral ventricles, which is another adult neurogenesis region along with the subgranular zone of the hippocampus, in transgenic mice demonstrate an association between high ADRA1A expression and high neurogenesis rate (Gupta et al., 2009). It is conceivable that ADRA1A also participates in hippocampal neurogenesis (Doze et al., 2011).

Conclusion

To sum up, we analyzed separately the dorsal and ventral hippocampi of foxes selected for contrasting behaviors and revealed differential expression of the *NR3C2*, *CYP26B1*, and *ADRA1A* genes, associated with both hippocampal neurogenesis and HPA axis regulation. Further studies of the expression of functional groups to which these genes belong are expected to shed light on hitherto unknown molecular mechanisms of domestication in general and on stress response weakening, elevated neurogenesis, and changes of the attitude to humans in domesticated animals in particular.

References

Amrein I. Adult hippocampal neurogenesis in natural populations of mammals. *Cold Spring Harb. Perspect. Biol.* 2015;7(5):a021295. DOI 10.1101/cshperspect.a021295.

Amrein I., Slomianka L. A morphologically distinct granule cell type in the dentate gyrus of the red fox correlates with adult hippocampal neurogenesis. *Brain Res.* 2010;1328:12-24. DOI 10.1016/j.brainres.2010.02.075.

Andersen C.L., Jensen J.L., Ørntoft T.F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 2004;64:5245-5250. DOI 10.1158/0008-5472.CAN-04-0496.

Belyaev D.K. Destabilizing selection as a factor in domestication. *J. Hered.* 1979;70(5):301-308. DOI 10.1093/oxfordjournals.jhered.a109263.

Bonhomme D., Minni A.M., Alfos S., Roux P., Richard E., Higueret P., Moisan M.P., Pallet V., Touyarot K. Vitamin A status regulates glucocorticoid availability in Wistar rats: consequences on cognitive functions and hippocampal neurogenesis? *Front. Behav. Neurosci.* 2014;8:20. DOI 10.3389/fnbeh.2014.00020.

Bremner J.D., McCaffery P. The neurobiology of retinoic acid in affective disorders. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry.* 2008;32(2):315-331. DOI 10.1016/j.pnpbp.2007.07.001.

Cai L., Li R., Zhou J.N. Chronic all-trans retinoic acid administration induces CRF over-expression accompanied by AVP up-regulation and multiple CRF-controlling receptors disturbance in the hypothalamus of rats. *Brain Res.* 2015;1601:1-7. DOI 10.1016/j.brainres.2014.12.054.

Cembrowski M.S., Wang L., Sugino K., Shields B.C., Spruston N. Hipposeq: a comprehensive RNA-seq database of gene expression in hippocampal principal neurons. *eLife*. 2016;5:e14997. DOI 10.7554/eLife.14997.001.

Corradini B.R., Iamashita P., Tampellini E., Farfel J.M., Grinberg L.T., Moreira-Filho C.A. Complex network-driven view of genomic mechanisms underlying Parkinson's disease: analyses in dorsal motor vagal nucleus, locus coeruleus, and substantia nigra. *BioMed Res. Int*. 2014;2014:543673. DOI 10.1155/2014/543673.

Croston R., Branch C.L., Kozlovsky D.Y., Roth T.C., LaDage L.D., Freas C.A., Pravosudov V.V. Potential mechanisms driving population variation in spatial memory and the hippocampus in food-caching chickadees. *Integr. Comp. Biol*. 2015;55(3):354-371. DOI 10.1093/icb/icv029.

de Kloet E.R., Otte C., Kumsta R., Kok L., Hillegers M.H.J., Hasselmann H., Kliegel D., Joëls M. Stress and depression: a crucial role of the mineralocorticoid receptor. *J. Neuroendocrinol*. 2016;28(8). DOI 10.1111/jne.12379.

de Martino M.U., Alesci S., Chrousos G.P., Kino T. Interaction of the glucocorticoid receptor and the chicken ovalbumin upstream promoter-transcription factor II (COUP-TFII): implications for the actions of glucocorticoids on glucose, lipoprotein, and xenobiotic metabolism. *Ann. N. Y. Acad. Sci.* 2004;1024:72-85. DOI 10.1196/annals.1321.006.

Deupree J.D., Reed A.L., Bylund D.B. Differential effects of the tricyclic antidepressant desipramine on the density of adrenergic receptors in juvenile and adult rats. *J. Pharmacol. Exp. Ther.* 2007;321(2):770-776. DOI 10.1124/jpet.106.118935.

Doze V.A., Handel E.M., Jensen K.A., Darsie B., Luger E.J., Haselton J.R., Talbot J.N., Rorabaugh B.R. α1A- and α1B-adrenergic receptors differentially modulate antidepressant-like behavior in the mouse. *Brain Res.* 2009;1285:148-157. DOI 10.1016/j.brainres.2009.06.035.

Doze V.A., Papay R.S., Goldenstein B.L., Gupta M.K., Collette K.M., Nelson B.W., Lyons M.J., Davis B.A., Luger E.J., Wood S.G., Haselton J.R., Simpson P.C., Perez D.M. Long-term α1A-adrenergic receptor stimulation improves synaptic plasticity, cognitive function, mood, and longevity. *Mol. Pharmacol.* 2011;80(4):747-758. DOI 10.1124/mol.111.073734.

Elia J., Capasso M., Zaheer Z., Lantieri F., Ambrosini P., Berrettini W., Devoto M., Hakonarson H. Candidate gene analysis in an on-going genome-wide association study of attention-deficit hyperactivity dis-

order: suggestive association signals in ADRA1A. *Psychiatr. Genet.* 2009;19(3):134-141. DOI 10.1097/YPG.0b013e32832a5043.

Floriou-Servou A., von Ziegler L., Stalder L., Sturman O., Privitera M., Rassi A., Cremonesi A., Thöny B., Bohacek J. Distinct proteomic, transcriptomic, and epigenetic stress responses in dorsal and ventral hippocampus. *Biol. Psychiatry.* 2018;84(7):531-541. DOI 10.1016/j.biopsych.2018.02.003.

Fuentealba P., Klausberger T., Karayannis T., Suen W.Y., Huck J., Tomioka R., Rockland K., Capogna M., Studer M., Morales M., Somogyi P. Expression of COUP-TFII nuclear receptor in restricted GABAergic neuronal populations in the adult rat hippocampus. *J. Neurosci*. 2010;30(5):1595-1609. DOI 10.1523/JNEUROSCI.4199-09.2010.

Garcia A., Steiner B., Kronenberg G., Bick-Sander A., Kempermann G. Age-dependent expression of glucocorticoid- and mineralocorticoid receptors on neural precursor cell populations in the adult murine hippocampus. *Aging Cell*. 2004;3(4):363-371. DOI 10.1111/j.1474-9728.2004.00130.x.

Gass P., Kretz O., Wolfer D.P., Berger S., Tronche F., Reichardt H.M., Kellendonk C., Lipp H.P., Schmid W., Schütz G. Genetic disruption of mineralocorticoid receptor leads to impaired neurogenesis and granule cell degeneration in the hippocampus of adult mice. *EMBO Rep.* 2000;1(5):447-451. DOI 10.1093/embo-reports/kvd088.

Gil-Ibáñez P., Bernal J., Morte B. Thyroid hormone regulation of gene expression in primary cerebrocortical cells: role of thyroid hormone receptor subtypes and interactions with retinoic acid and glucocorticoids. *PLoS One*. 2014;9(3):e91692. DOI 10.1371/journal.pone.0091692.

Gulyaeva N.V. Functional neurochemistry of the ventral and dorsal hippocampus: stress, depression, dementia and remote hippocampal damage. *Neurochem. Res.* 2019;44:1306-1322. DOI 10.1007/s11064-018-2662-0.

Gupta M.K., Papay R.S., Jurgens C.W.D., Gaivin R.J., Shi T., Doze V.A., Perez D.M. α1-Adrenergic receptors regulate neurogenesis and gliogenesis. *Mol. Pharmacol.* 2009;76(2):314-326. DOI 10.1124/mol.109.057307.

Haddjeri-Hopkins A., Tapia M., Ramirez-Franco J., Tell F., Marqueze-Pouey B., Amalric M., Goaillard J.M. Refining the identity and role of Kv4 channels in mouse substantia nigra dopaminergic neurons. *eNeuro*. 2021;8(4):ENEURO.0207-21.2021. DOI 10.1523/ENEURO.0207-21.2021.

Harris A.P., Holmes M.C., De Kloet E.R., Chapman K.E., Seckl J.R. Mineralocorticoid and glucocorticoid receptor balance in control of HPA axis and behaviour. *Psychoneuroendocrinol*. 2013;38(5):648-658. DOI 10.1016/j.psyneuen.2012.08.007.

Hélène R., Julie B., Aloïs L., Marie-Pierre M., Véronique P., Anabelle R., Jean-Benoît C. Retinoids and glucocorticoids have opposite effects on actin cytoskeleton rearrangement in hippocampal HT22 cells. *Int. J. Biochem. Cell Biol.* 2016;71:102-110. DOI 10.1016/j.biocel.2015.12.014.

Hu P., Wang Y., Liu J., Meng F.T., Qi X.R., Chen L., van Dam A.M., Joëls M., Lucassen P.J., Zhou J.N. Chronic retinoic acid treatment suppresses adult hippocampal neurogenesis, in close correlation with depressive-like behavior. *Hippocampus*. 2016;26(7):911-923. DOI 10.1002/hipo.22574.

Hu P., van Dam A.M., Wang Y., Lucassen P.J., Zhou J.N. Retinoic acid and depressive disorders: evidence and possible neurobiological mechanisms. *Neurosci. Biobehav. Rev.* 2020;112:376-391. DOI 10.1016/j.neubiorev.2020.02.013.

Huang S., Slomianka L., Farmer A.J., Kharlamova A.V., Gulevich R.G., Herbeck Yu.E., Trut L.N., Wolfer D.P., Amrein I. Selection for tameness, a key behavioral trait of domestication, increases adult hippocampal neurogenesis in foxes. *Hippocampus*. 2015;25(8):963-975. DOI 10.1002/hipo.22420.

Jacobs L.F., Gaulin S.J., Sherry D.F., Hoffman G.E. Evolution of spatial cognition: sex-specific patterns of spatial behavior predict hippocampal size. *Proc. Natl. Acad. Sci. USA*. 1990;87(16):6349-6352. DOI 10.1073/pnas.87.16.6349.

Jacobs L.F., Spencer W.D. Natural space-use patterns and hippocampal size in kangaroo rats. *Brain Behav. Evol*. 1994;44(3):125-132. DOI 10.1159/000113584.

Jinno S., Kosaka T. Stereological estimation of numerical densities of glutamatergic principal neurons in the mouse hippocampus. *Hippocampus.* 2010;20(7):829-840. DOI 10.1002/hipo.20685.

Kanatsou S., Fearey B.C., Kuil L.E., Lucassen P.J., Harris A.P., Seckl J.R., Krugers H., Joels M. Overexpression of mineralocorticoid receptors partially prevents chronic stress-induced reductions in hippocampal memory and structural plasticity. *PLoS One.* 2015;10(11):e0142012. DOI 10.1371/journal.pone.0142012.

Kanatsou S., Karst H., Kortesidou D., van den Akker R.A., den Blaauwen J., Harris A.P., Seckl J.R., Krugers H.J., Joels M. Overexpression of mineralocorticoid receptors in the mouse forebrain partly alleviates the effects of chronic early life stress on spatial memory, neurogenesis and synaptic function in the dentate gyrus. *Front. Cell. Neurosci*. 2017;11:132. DOI 10.3389/fncel.2017.00132.

Kane M.A., Folias A.E., Wang C., Napoli J.L. Ethanol elevates physiological *all-trans*-retinoic acid levels in select loci through altering retinoid metabolism in multiple loci: a potential mechanism of ethanol toxicity. *FASEB J.* 2010;24(3):823-832. DOI 10.1096/fj.09-141572.

Kang P., Rogalska J., Walker C.A., Burke M., Seckl J.R., Macleod M.R., Lai M. Injury-induced mineralocorticoid receptor expression involves differential promoter usage: a novel role for the rat MRβ variant. *Mol. Cell. Endocrinol.* 2009;305(1-2):56-62. DOI 10.1016/j.mce.2009.02.008.

Kukekova A.V., Johnson J.L., Xiang X., Feng S., Liu S., Rando H.M., Kharlamova A.V., Herbeck Yu., Serdyukova N.A., Xiong Z., Beklemischeva V., Koepfli K.-P., Gulevich R.G., Vladimirova A.V., Hekman J.P., Perelman P.L., Graphodatsky A.S., O'Brien S.J., Wang X., Clark A.G., Acland G.A., Trut L.N., Zhang G. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. *Nat. Ecol. Evol.* 2018;2(9):1479-1491. DOI 10.1038/s41559-018-0611-6.

Kvichansky A.A., Volobueva M.N., Manolova A.O., Bolshakov A.P., Gulyaeva N.V. Neonatal proinflammatory stress alters the expression of genes of corticosteroid receptors in the rat hippocampus: septo-temporal differences. *Neurochem. J.* 2017;11(3):255-258. DOI 10.1134/S1819712417030059.

Lai M., Horsburgh K., Bae S.E., Carter R.N., Stenvers D.J., Fowler J.H., Yau J.L., Gomez-Sanchez C.E., Holmes M.C., Kenyon C.J., Seckl J.R., Macleod M.R. Forebrain mineralocorticoid receptor overexpression enhances memory, reduces anxiety and attenuates neuronal loss in cerebral ischaemia. *Eur. J. Neurosci.* 2007;25(6):1832-1842. DOI 10.1111/j.1460-9568.2007.05427.x.

Le Menuet D., Lombès M. The neuronal mineralocorticoid receptor: from cell survival to neurogenesis. *Steroids*. 2014;91:11-19. DOI 10.1016/j.steroids.2014.05.018.

Lee A.-R., Kim J.-H., Cho E., Kim M., Park M. Dorsal and ventral hippocampus differentiate in functional pathways and differentially associate with neurological disease-related genes during postnatal development. *Front. Mol. Neurosci.* 2017;10:331. DOI 10.3389/fnmol.2017.00331.

Levone B.R., Cryan J.F., O'Leary O.F. Role of adult hippocampal neurogenesis in stress resilience. *Neurobiol. Stress*. 2015;1:147-155. DOI 10.1016/j.ynstr.2014.11.003.

Lewis A.S., Pittenger S.T., Mineur Y.S., Stout D., Smith P.H., Picciotto M.R. Bidirectional regulation of aggression in mice by hippocampal alpha-7 nicotinic acetylcholine receptors. *Neuropsychopharmacology*. 2018;43(6):1267-1275. DOI 10.1038/npp.2017.276.

Lipp H.P. Evolutionary shaping of adult hippocampal neurogenesis in mammals-cognitive gain or developmental priming of personality traits? *Front. Neurosci.* 2017;11:420. DOI 10.3389/fnins.2017.00420.

Livak K.J., Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods*. 2001;25(4):402-408. DOI 10.1006/meth.2001.1262.

Lowe A., Dalton M., Sidhu K., Sachdev P., Reynolds B., Valenzuela M. Neurogenesis and precursor cell differences in the dorsal and ventral adult canine hippocampus. *Neurosci. Lett*. 2015;593:107-113. DOI 10.1016/j.neulet.2015.03.017.

Maggio N., Segal M. Differential modulation of long-term depression by acute stress in the rat dorsal and ventral hippocampus. *J. Neurosci.* 2009;29(27):8633-8638. DOI 10.1523/JNEUROSCI.1901-09.2009.

Medina A., Seasholtz A.F., Sharma V., Burke S., Bunney W., Jr., Myers R.M., Schatzberg A., Akil H., Watson S.J. Glucocorticoid and mineralocorticoid receptor expression in the human hippocampus in major depressive disorder. *J. Psychiatr. Research.* 2013;47(3):307-314. DOI 10.1016/j.jpsychires.2012.11.002.

Ming G., Song H. Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron*. 2011;70(4):687-702. DOI 10.1016/j.neuron.2011.05.001.

Mishra S., Kelly K.K., Rumian N.L., Siegenthaler J.A. Retinoic acid is required for neural stem and progenitor cell proliferation in the adult hippocampus. *Stem Cell Rep*. 2018;10(6):1705-1720. DOI 10.1016/j.stemcr.2018.04.024.

O'Leary O.F., Cryan J.F. A ventral view on antidepressant action: roles for adult hippocampal neurogenesis along the dorsoventral axis. *Trends Pharmacol. Sci.* 2014;35(12):675-687. DOI 10.1016/j.tips.2014.09.011.

O'Rourke T., Boeckx C. Glutamate receptors in domestication and modern human evolution. *Neurosci. Biobehav. Rev.* 2020;108:341-357. DOI 10.1016/j.neubiorev.2019.10.004.

Oskina I.N., Herbeck Yu.E., Shikhevich S.G., Plyusnina I.Z., Gulevich R.G. Alterations in the hypothalamus-pituitary-adrenal and immune systems during selection of animals for tame behavior. *Informatsionnyy Vestnik VOGiS = The Herald of Vavilov Society for Geneticists and Breeders*. 2008;12(1/2):39-49. (in Russian)

Ovchinnikov V.Yu., Antonov E.V., Vasilyev G.V., Shihevich S.G., Shepeleva D.V., Herbeck Yu.E. Hippocampal glucocorticoid receptor and microRNA gene expression and serum cortisol concentration in foxes selected for behavior toward humans. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2018;22(2):230-234. DOI 10.18699/VJ18.352. (in Russian)

Pendleton A.L., Shen F., Taravella A.M., Emery S., Veeramah K.R., Boyko A.R., Kidd J.M. Comparison of village dog and wolf genomes highlights the role of the neural crest in dog domestication. *BMC Biol*. 2018;16(1):e64. DOI 10.1186/s12915-018-0535-2.

Pörtl D., Jung C. Is dog domestication due to epigenetic modulation in brain? *Dog Behav*. 2017;3(2):21-32. DOI 10.4454/db.v3i2.55.

Prasolova L.A., Gerbek Yu.E., Gulevich R.G., Shikhevich S.G., Konoshenko M.Y., Kozhemyakina R.V., Oskina I.N., Plyusnina I.Z. The effects of prolonged selection for behavior on the stress response and activity of the reproductive system of male grey rats (*Rattus norvegicus*). *Russ. J. Genet.* 2014;50(8):846-852. DOI 10.1134/S1022795414080031.

Price E.O. Animal Domestication and Behavior. Oxford; New York: Oxford Univ. Press, 2002.

Ratka A., Sutanto W., Bloemers M., de Kloet E.R. On the role of brain mineralocorticoid (type I) and glucocorticoid (type II) receptors in neuroendocrine regulation. *Neuroendocrinology.* 1989;50(2):117-123. DOI 10.1159/000125210.

Rehkämper G., Frahm H.D., Cnotka J. Mosaic evolution and adaptive brain component alteration under domestication seen on the background of evolutionary theory. *Brain Behav. Evol*. 2008;71(2):115-126. DOI 10.1159/000111458.

Rozeboom A.M., Akil H., Seasholtz A.F. Mineralocorticoid receptor overexpression in forebrain decreases anxiety-like behavior and alters the stress response in mice. *Proc. Natl. Acad. Sci. USA*. 2007;104(11):4688-4693. DOI 10.1073/pnas.0606067104.

Saaltink D.J., Vreugdenhil E. Stress, glucocorticoid receptors, and adult neurogenesis: a balance between excitation and inhibition? *Cell. Mol. Life Sci.* 2014;71:2499-2515. DOI 10.1007/s00018-014-1568-5.

Sklepkovych B.O., Montevecchi W.A. Food availability and food hoarding behaviour by red and arctic foxes. *Arctic*. 1996;49(3):228-234. DOI 10.14430/arctic1199.

Sonnenberg B.R., Branch C.L., Pitera A.M., Bridge E., Pravosudov V.V. Natural selection and spatial cognition in wild food-caching mountain chickadees. *Curr. Biol*. 2019;29(4):670-676. DOI 10.1016/j.cub.2019.01.006.

Stoney P.N., McCaffery P. A vitamin on the mind: new discoveries on control of the brain by vitamin A. *World Rev. Nutr. Diet.* 2016;115:98-108. DOI 10.1159/000442076.

Stoney P.N., Fragoso Y.D., Saeed R.B., Ashton A., Goodman T., Simons C., Gomaa M.S., Sementilli A., Sementilli L., Ross A.W., Morgan P.J., McCaffery P.J. Expression of the retinoic acid catabolic enzyme CYP26B1 in the human brain to maintain signaling homeostasis. *Brain Struct. Funct.* 2016;221:3315-3326. DOI 10.1007/s00429-015-1102-z.

Theofanopoulou C., Gastaldon S., O'Rourke T., Samuels B.D., Messner A., Martins P.T., Delogu F., Alamri S., Boeckx C. Self-domestication in *Homo sapiens*: insights from comparative genomics. *PloS One*. 2017;12(10):e0185306. DOI 10.1371/journal.pone.0185306.

Todorov O.S., Weisbecker V., Gilissen E., Zilles K., De Sousa A.A. Primate hippocampus size and organization are predicted by sociality but not diet. *Proc. Royal Soc. B.* 2019;286(1914):20191712. DOI 10.1098/rspb.2019.1712.

Trut L.N., Plyusnina I.Z., Oskina I.N. An experiment on fox domestication and debatable issues of evolution of the dog. *Russ. J. Genet.* 2004;40(6):644-655. DOI 10.1023/B:RUGE.0000033312.92773.c1.

Trut L., Oskina I., Kharlamova A. Animal evolution during domestication: the domesticated fox as a model. *BioEssays*. 2009;31(3):349-360. DOI 10.1002/bies.200800070.

Trut L.N., Kharlamova A.V., Pilipenko A.S., Herbeck Yu.E. The fox domestication experiment and dog evolution: A view based on modern molecular, genetic, and archaeological data. *Russ. J. Genet.* 2021;57(7):778-794. DOI 10.1134/S1022795421070140.

Truvé K., Parris T.Z., Vizlin-Hodzic D., Salmela S., Berger E., Ågren H., Funa K. Identification of candidate genetic variants and altered protein expression in neural stem and mature neural cells support altered microtubule function to be an essential component in bipolar disorder. *Transl. Psychiatry*. 2020;10(1):e390. DOI 10.1038/s41398-020-01056-1.

Tzakis N., Holahan M.R. Social memory and the role of the hippocampal CA2 region. *Front. Behav. Neurosci*. 2019;13:233. DOI 10.3389/fnbeh.2019.00233.

Veenema A.H., Neumann I.D. Neurobiological mechanisms of aggression and stress coping: a comparative study in mouse and rat selection lines. *Brain Behav. Evol.* 2007;70(4):274-285. DOI 10.1159/000105491.

Veenema A.H., Meijer O.C., de Kloet E.R., Koolhaas J.M. Genetic selection for coping style predicts stressor susceptibility. *J. Neuroendocrinol.* 2003;15:256-267. DOI 10.1046/j.1365-2826.2003.00986.x.

Vogel J.W., La Joie R., Grothe M.J., Diaz-Papkovich A., Doyle A., Vachon-Presseau E., Lepage C., Vos de Wael R., Thomas R.A., Iturria-Medina Y., Bernhardt B., Rabinovici G.D., Evans A.C. A molecular gradient along the longitudinal axis of the human hippocampus informs large-scale behavioral systems. *Nat. Commun.* 2020;11(1):960. DOI 10.1038/s41467-020-14518-3.

Weaver I.C.G., Cervoni N., Champagne F.A., D'Alessio A.C., Sharma S., Seckl J.R., Dymov S., Szyf M., Meaney M.J. Epigenetic programming by maternal behavior. *Nat. Neurosci.* 2004;7:847-854. DOI 10.1038/nn1276.

Xiao F., Zhang X., Ni P., Yu H., Gao Q., Li M., Huo P., Wei Z., Wang S., Zhang Y., Zhao R., Li A., Li Z., Li Yu., Cheng H., Du L., Ren S., Yu Q., Liu Ya., Zhao Yu. Voltage-dependent potassium channel Kv4.2 alleviates the ischemic stroke impairments through activating neurogenesis. *Neurochem. Int*. 2021;50:e105155. DOI 10.1016/j.neuint.2021.105155.

Ye J., Coulouris G., Zaretskaya I., Cutcutache I., Rozen S., Madden T.L. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform.* 2012;13:134. DOI 10.1186/1471-2105-13-134.

Acknowledgements. This work was supported by the Russian Science Foundation, project 21-44-04405. The breeding of experimental animals in the Shared Access Center for Gene Pools of Fur and Farm Animals, Institute of Cytology and Genetics, was supported by State Budgeted Project FWNR-2022-0019.

The authors are grateful to V.V. Ivaykin, A.V. Vladimirova, I.V. Pivovarova, T.I. Semenova, V.I. Vladimirova, T.V. Konovalova, and all the staff of the Shared Access Center for assistance in the study. They also acknowledge the significant contribution of the reviewers to manuscript improvement.

Conflict of interest. The authors declare no conflict of interest.

Received December 1, 2022. Revised June 18, 2023. Accepted June 30, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):662-675
DOI 10.18699/VJGB-23-77

Original Russian text https://vavilovj-icg.ru/

Human-genome single nucleotide polymorphisms affecting transcription factor binding and their role in pathogenesis

E.V. Antontseva ✉, A.O. Degtyareva, E.E. Korbolina, I.S. Damarov, T.I. Merkulova

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ antontseva@bionet.nsc.ru

Abstract. Single nucleotide polymorphisms (SNPs) are the most common type of variation in the human genome. The vast majority of SNPs identified in the human genome do not have any effect on the phenotype; however, some can lead to changes in the function of a gene or the level of its expression. Most SNPs associated with certain traits or pathologies are mapped to regulatory regions of the genome and affect gene expression by changing transcription factor binding sites. In recent decades, substantial effort has been invested in searching for such regulatory SNPs (rSNPs) and understanding the mechanisms by which they lead to phenotypic differences, primarily to individual differences in susceptibility to diseases and in sensitivity to drugs. The development of the NGS (next-generation sequencing) technology has contributed not only to the identification of a huge number of SNPs and to the search for their association (genome-wide association studies, GWASs) with certain diseases or phenotypic manifestations, but also to the development of more productive approaches to their functional annotation. It should be noted that the presence of an association does not allow one to identify a functional, truly disease-associated DNA sequence variant among multiple marker SNPs that are detected due to linkage disequilibrium. Moreover, determination of associations of genetic variants with a disease does not provide information about the functionality of these variants, which is necessary to elucidate the molecular mechanisms of the development of pathology and to design effective methods for its treatment and prevention. In this regard, the functional analysis of SNPs annotated in the GWAS catalog, both at the genome-wide level and at the level of individual SNPs, became especially relevant in recent years. A genome-wide search for potential rSNPs is possible without any prior knowledge of their association with a trait. Thus, mapping expression quantitative trait loci (eQTLs) makes it possible to identify an SNP for which – among transcriptomes of homozygotes and heterozygotes for its various alleles – there are differences in the expression level of certain genes, which can be located at various distances from the SNP. To predict rSNPs, approaches based on searches for allele-specific events in RNA-seq, ChIP-seq, DNase-seq, ATAC-seq, MPRA, and other data are also used. Nonetheless, for a more complete functional annotation of such rSNPs, it is necessary to establish their association with a trait, in particular, with a predisposition to a certain pathology or sensitivity to drugs. Thus, approaches to finding SNPs important for the development of a trait can be categorized into two groups: (1) starting from data on an association of SNPs with a certain trait, (2) starting from the determination of allele-specific changes at the molecular level (in a transcriptome or regulome). Only comprehensive use of strategically different approaches can considerably enrich our knowledge about the role of genetic determinants in the molecular mechanisms of trait formation, including predisposition to multifactorial diseases.
Key words: regulatory single-nucleotide polymorphism; transcription factor-binding sites; gene expression; genome-wide studies.

For citation: Antontseva E.V., Degtyareva A.O., Korbolina E.E., Damarov I.S., Merkulova T.I. Human-genome single nucleotide polymorphisms affecting transcription factor binding and their role in pathogenesis. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):662-675. DOI 10.18699/VJGB-23-77

Однонуклеотидные замены в геноме человека, влияющие на связывание факторов транскрипции, и их роль в развитии патологий

Е.В. Антонцева ✉, А.О. Дегтярева, Е.Е. Корболина, И.С. Дамаров, Т.И. Меркулова

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ antontseva@bionet.nsc.ru

Аннотация. Однонуклеотидные замены, также называемые однонуклеотидными полиморфизмами (single nucleotide polymorphism, SNP), – это наиболее распространенный тип вариаций генома человека. Подавляющая часть выявленных в геноме человека SNP не оказывает какого-либо воздействия на молекулярный фенотип, однако некоторые способны приводить к изменению функции гена или уровня его экспрессии. В то же время большинство SNP, ассоциированных с некими признаками или патологиями, картируются в регуляторных областях генома, изменяют потенциальные сайты связывания транскрипционных факторов и, соответственно, могут влияют на экспрессию генов. В последние десятилетия значительные усилия были направлены на поиск таких регуляторных SNP (rSNP), а так-

же на понимание механизмов, посредством которых они приводят к фенотипическим различиям, в первую очередь к разной предрасположенности к заболеваниям и индивидуальной чувствительности к лекарственным препаратам. Развитие технологии NGS (next generation sequencing) способствовало не только выявлению огромного количества SNP и поиску их ассоциации (genome wide association studies, GWAS) с некими заболеваниями или фенотипическими проявлениями, но и развитию более производительных подходов для их функциональной аннотации. Стоит отметить, что наличие ассоциации не позволяет выделить функциональный, действительно связанный с болезнью вариант последовательности ДНК из множества маркерных, которые выявляются за счет неравновесия по сцеплению. Более того, установление ассоциаций генетических вариантов с заболеванием не дает сведений о функциональности этих вариантов, что необходимо для выяснения молекулярных механизмов развития патологии и разработки эффективных методов ее лечения и профилактики. В связи с этим функциональный анализ SNP, аннотированных в GWAS каталоге, как на полногеномном уровне, так и на уровне отдельных SNP в последние годы стал особенно актуальным. В настоящее время активно развивается полногеномный поиск потенциально регуляторных SNP без каких-либо предварительных знаний об их ассоциации с признаком. Так, картирование локусов количественных признаков экспрессии (eQTL, expression quantitative trait loci) позволяет выявить SNP, для которого в транскриптомах гомозигот по разным его аллелям, а также гетерозигот наблюдаются различия в уровне экспрессии неких генов, причем как близко расположенных, так и на значительном удалении. Для предсказания регуляторных SNP используют также подходы, основанные на поиске аллель-специфических событий в данных RNA-seq, ChIP-seq, DNase-seq, ATAC-seq, MPRA и т. д. Однако для более полной характеристики таких rSNP необходимо устанавливать их ассоциацию с признаком, в частности с предрасположенностью к некой патологии или с чувствительностью к лекарственным препаратам. Таким образом, именно комплексное использование двух основанных на противоположных принципах подходов к поиску значимых для развития признака (патологии) SNP: с одной стороны, исходящего из данных по ассоциации SNP с неким признаком, а с другой стороны, идущего от определения аллель-специфичных изменений на молекулярном уровне (в транскриптоме или регуломе) – существенно обогащает картину наших знаний о роли генетических детерминант в молекулярных механизмах формирования признаков, включая предрасположенность к многофакторным заболеваниям.

Ключевые слова: регуляторный однонуклеотидный полиморфизм; сайты связывания транскрипционных факторов; экспрессия генов; полногеномные исследования.

Introduction

One of the main tasks of human genetics is to clarify the mechanisms by which genome variations lead to phenotypic differences, primarily to individual differences in susceptibility to diseases and in sensitivity to drugs. Single-nucleotide polymorphisms (SNPs) are the most common type of genome variation (Chanock, 2001). Currently, due to the development of next-generation sequencing (NGS) technologies, more than 950 million SNPs of the human genome are registered in database dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi) (Sherry et al., 2001); furthermore, rare SNPs with a frequency (prevalence in the population) of less than 1 % constitute more than 90 % of the total number of SNPs. It seems unlikely that the vast majority of identified variations can be important for the phenotype, but some of them certainly form the genetic basis of phenotypic traits, including predisposition to various diseases. In recent decades, major efforts have been applied to the search for such SNPs.

The most popular approach (which started back in the 1980s) to the identification of trait-related SNPs has been the determination of associations with diseases for SNPs found in candidate genes (Lander, Schork, 1994; Ring, Kroetz, 2002), and in this context, only SNPs affecting the protein-coding part of a gene have been investigated (Cooper, 1998). Somewhat later, there has been some interest in elucidating the functionality of variants located in noncoding regions of genes, i. e., in determining an effect of such variants on a certain molecular phenotype. In particular, it has been shown that such SNPs affect transcription factor-binding sites (TFBSs), thereby leading to changes in the expression of the respective genes (Ludlow et al., 1996; Piedrafita et al., 1996; Knight et al., 1999; Vasiliev et al., 1999). Nonetheless, these few functional studies have remained almost invisible against the backdrop of a huge wave of research aimed at identifying associations.

The productivity of detection of disease-associated SNPs increased dramatically with the advent of the genome-wide association study (GWAS) technology in the mid-2000s (Fig. 1, *a*), which is based on an unbiased – not based on any ideas about the formation of a trait – principle of a genome-wide search for SNPs associated with the trait (Visscher et al., 2012; Tam et al., 2019). To date, more than 72 thousand associations of genetic variants with traits have been found by GWASs (GWAS Catalog, https://www.ebi.ac.uk/gwas/), thus allowing to find many new genes and systems of genes associated with predisposition to various diseases (Buniello et al., 2019; Tam et al., 2019; Claussnitzer et al., 2020).

Nevertheless, GWAS technology does not provide any information about the functionality of the detected variants, thereby making it very difficult to elucidate the molecular mechanisms underlying the development of a pathology and hence to develop effective methods for its treatment and prevention. Additionally, based on results of GWASs, it is almost impossible to distinguish a truly disease-associated variant from the many marker variants that are detected due to linkage disequilibrium (Lappalainen, 2015; Tam et al., 2019; Zhao et al., 2020). It is also known that most SNPs identified by GWASs are located in the noncoding part of the genome and, as a rule, in its regulatory regions (e. g., promoters and enhancers) (Hindorff et al., 2009; Maurano et al., 2012; Bryzgalov et al., 2013; Farh et al., 2015); these data imply an influence of such SNPs on the binding of transcription factors (TFs) and on gene expression. Thus, a need for research on functional interpretation of data from GWASs – both at

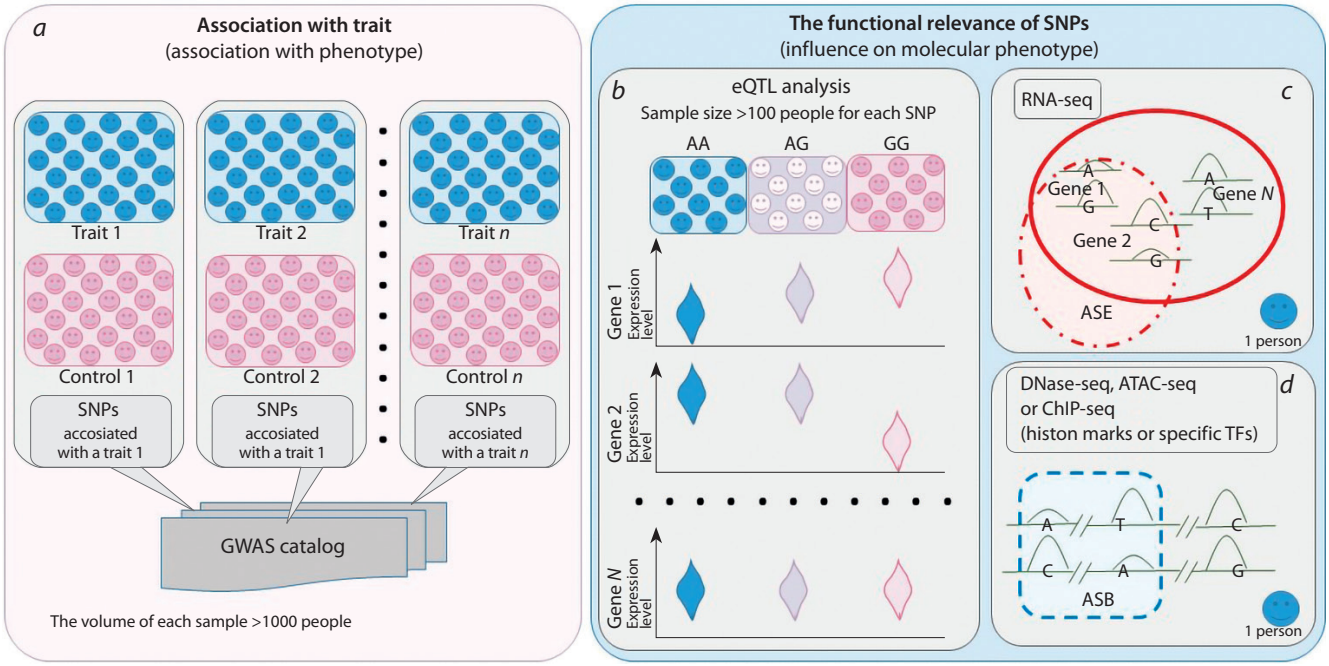

Fig. 1. Genome-wide approaches to the analysis of SNPs.

a, The principle of GWASs; *b*, the scheme of eQTL analysis for each SNP. A search for allele-specific events: *c*, expression (allele-specific expression, ASE) and *d*, binding (allele-specific binding, ASB).

the level of individual potentially regulatory SNPs (rSNPs) and at the level of all such variants collectively – has become obvious. In addition, because the GWAS technology greatly underestimates the actual number of associations owing to the required strict thresholds in statistical processing (Tam et al., 2019), investigators have recognized the need to develop GWAS-unrelated large-scale function-based approaches to the search for rSNPs (Westra, Franke, 2014; Maurano et al., 2015; Cavalli et al., 2016b; Korbolina et al., 2018).

Of the genome-wide function-based approaches, the first in timing of emergence (almost simultaneous with the start of the application of the GWAS technology) is analysis of expression quantitative trait loci (eQTLs) (see Fig. 1, *b*). This analysis, by means of transcriptome data (earlier, microarray data have been employed, then data from high-throughput RNA sequencing [RNA-seq] started to be used), determines for each SNP a difference in the level of expression of individual genes among homozygotes and heterozygotes for different alleles of this SNP (Westra, Franke, 2014; GTEx Consortium, 2020). Somewhat later, techniques were devised for finding allele-specific events both in data from RNA-seq (allele-specific expression [ASE] events) (see Fig. 1, *c*) (Castel et al., 2020; Fan et al., 2020) and in data from ChIP-seq, DNase-seq, and ATAC-seq (allele-specific binding [ASB] events) (see Fig. 1, *d*) (Maurano et al., 2015; Cavalli et al., 2016a, b; Xu et al., 2020; Korbolina et al., 2021). Massive function-based approaches also include massively parallel reporter assay (MPRA), SNP-seq, and SNP-SELEX (Zhang et al., 2018; Lu et al., 2021; Yan et al., 2021).

At the beginning of this review, using the latest studies as examples, we examine rSNPs that are (i) associated with the development of pathologies according to results of GWASs and are (ii) characterized in detail in terms of their influence on an interaction with a TF and on the expression of nearby or distant genes. Next, the review describes the application of functional genomics methods to interpretation of data from GWASs and to the search for new regulatory variants without GWASs. In the course of the presentation, strengths and weaknesses of these approaches are shown, as is the importance of comprehensive use of GWASs and functional genomics methods to reveal the role of SNPs in molecular aberrations underlying the development of a pathology.

Functional interpretation of data from GWASs at the level of individual rSNPs

A wide range of experimental methods are utilized to functionally study individual possible rSNPs. At initial stages, classical methods are usually used: analysis of DNA-probe retardation in a gel by nuclear extract proteins (electrophoretic mobility shift assay, EMSA) and a reporter assay (analysis of reporter gene expression under the control of allelic variants of the SNP region), which allow a researcher to detect an influence of an SNP on the binding of some TF or on reporter gene expression, respectively (Antontseva et al., 2015; Fang et al., 2017). To identify the TFs the binding sites of which are affected by nucleotide substitution, scientists perform EMSA using appropriate antibodies or purified TFs (Piedrafita et al., 1996; Knight et al., 1999; Vasiliev et al., 1999; Jiang et al., 2020), chromatin immunoprecipitation with detection of allele asymmetry in a PCR product (ChIP-AS-qPCR) (Gao et al., 2018; Choi et al., 2020; Thynn et al., 2020; Protze et al., 2022), and mass-spectrometric analysis of proteins isolated from complexes with oligonucleotides containing minor alleles (Fang et al., 2017; Liu D. et al., 2018; Choi et al., 2020).

An impact of a nucleotide substitution in the identified TFBSs on their potential target genes is confirmed by experiments with downregulation or artificial upregulation of genes of the corresponding TFs or by studying the effect of point mutations introduced into the TFBS using CRISPR/Cas9 technology (Prestel et al., 2019; Gutierrez-Arcelus et al., 2020; Pan et al., 2020; Thynn et al., 2020; Wang Y. et al., 2020; Wang X. et al., 2021). Lately, an increasingly popular approach in this field of research on SNPs has also been the determination of allelic imbalance of the expression for SNPs located in transcribed regions of genes. Nonetheless, it is worth noting that such an SNP can be either an rSNP proper (Syddall et al., 2013; Klein et al., 2019) or a marker SNP in a linkage group with an rSNP located in a nontranscribed region (Fang et al., 2017; Li X.-X. et al., 2019; Peng et al., 2020).

Various combinations of these techniques are employed in modern research, as illustrated by the examples below.

rs36115365

A textbook example of a well-studied rSNP is rs36115365 (G/C), which is situated at a locus associated with various types of cancer according to data from GWASs (chr5p15.33: region 2). At this locus, a correlation analysis has revealed nine SNPs ($r^2 > 0.60$, 1000G EUR population) associated with pancreatic cancer, testicular germ cell tumor, lung cancer, and melanoma. To screen all nine SNPs for regulatory activity, EMSA and reporter assay were performed by means of eight human cell lines (Fang et al., 2017). The use of several cell lines is a common practice in such studies (Bryzgalov et al., 2013; Boldes et al., 2020) and is aimed at the highest possible coverage of events of interaction of TFs with their binding sites; the reason is substantial differences in the sets of TFs expressed in different cell types (Tobias et al., 2021). The screening analyses identified only SNP rs36115365 as potentially regulatory, and its C allele showed both much better binding to a certain protein in the EMSA and greater activation of reporter gene expression as compared to the G allele (Fang et al., 2017).

rs36115365 is located ~18 kilobase pairs (kbp) upstream of the start of the *TERT* gene (encoding reverse transcriptase of the telomerase complex) and approximately 5 kbp downstream of the end of the *CLPTM1L* gene. According to the results of ChIP-seq (chromatin immunoprecipitation followed by sequencing of DNA from the precipitates) from project ENCODE (Moore et al., 2020), this SNP's location overlaps with many TFBSs, and this region is enriched with active chromatin histone marks, which is typical for enhancer regions. Inactivation of this region by small-interfering-RNA-mediated transcriptional silencing (Malecová, Morris, 2010) results in downregulation of only the *TERT* gene. To identify the TF the binding site of which changes as a consequence of the substitution of G with C, the binding of proteins from nuclear extracts to oligonucleotides containing minor alleles was implemented, followed by mass-spectrometric analysis of the bound proteins. After an analysis of the obtained peptides, four TF candidates that prefer the C allele were proposed: ZNF148, VEZF1/ZNF161, ZNF281, and ZNF740. In EMSA involving specific antibodies to these TFs, only ZNF148 was confirmed, which was subsequently verified by means of the purified ZNF148 protein in an experiment. A small-interfering-RNA-mediated knockdown of ZNF148 gave a definitive answer because it reduced *TERT* expression and telomerase activity and shortened telomere length. Thus, the C allele corrects the binding site of ZNF148, enhances the expression of *TERT* and, as a consequence, increases the risk of carcinogenesis (Fang et al., 2017).

rs174575

According to GWASs, rs174575 (C/G) correlates with an elevated risk of colorectal cancer (Tian et al., 2020). This SNP is located in the first intron of the delta-6-desaturase gene (*FADS2*) at a distance of +41.5 kbp from its transcription start site and at a distance of –178.8 kbp from the transcription start site of the gene of long noncoding RNA AP002754.2. The rs174575 region in chromatin is enriched with histone modifications characteristic of active regulatory regions (H3K4me1, H3K4me3, or H3K27ac), and judging by DNase-seq data (identification of sites of hypersensitivity to DNase I) and findings of ATAC-seq (assay for transposase-accessible chromatin), corresponds to open chromatin (ENCODE). The hypothesis of a regulatory role of rs174575 is supported by eQTL analysis data, which show an association of the G allele with overexpression of *FADS2* and AP002754.2.

Computer analysis of DNA motifs using Web services Cistrome (Zheng et al., 2019) and JASPAR (Fornes et al., 2020) has revealed that in the case of rs174575, the replacement of G with C damages the binding site of TF E2F1, and the data of ChIP-seq from ENCODE, obtained on colorectal cancer LoVo cells, indicate that E2F1 is mapped to the location of this SNP. A cross-competitive EMSA has confirmed better binding of a certain nuclear extract protein to the G allele, and that this protein is E2F1 has been demonstrated by ChIP-qPCR with appropriate antibodies. For instance, in cell lines with different genotypes – HCT116 (CG), SNU-C1 (CG), and HT115 (CC) – a stronger binding of E2F1 is observed in cells carrying the G allele. A study involving reporter constructs has also confirmed a higher enhancer activity of a DNA fragment containing the G allele (Tian et al., 2020).

Direct contact between the region containing rs174575 and promoters of *FADS2* and AP002754.2 has been detected by the chromosome conformation capture (3C) method, and it turned out that the interaction was much more pronounced in cell lines carrying the G allele. Further experiments indicated that overexpression of AP002754.2 sharply raises the level of *FADS2* expression in HCT116 and LoVo cells, and a knockdown of AP002754.2 by microRNA causes a decrease in the expression of this gene, indicating a stimulatory role of AP002754.2 in the regulation of the *FADS2* gene. On the other hand, overexpression of *FADS2* or AP002754.2 significantly increases the proliferation rate of HCT116 and LoVo cells, whereas a knockdown of *FADS2* or AP002754.2 significantly reduces it. It has also been shown that overexpression of *FADS2* and AP002754.2 accelerates tumor growth in *in vivo* experiments in mice. It is known that the product of the *FADS2* gene is a key enzyme in the biosynthesis of polyunsaturated fatty acids, including arachidonic acid, which in turn is a precursor of prostaglandin E2 (PGE2), which promotes tumor growth and metastasis. Thus, rs174575 acts as an allele-specific enhancer that stimulates the transcription of *FADS2* and AP002754.2, leading to an elevated risk of colorectal cancer in the case of the G allele (Tian et al., 2020).

rs4903064

According to GWASs, rs4903064 (T/C) is associated with renal cell carcinoma, the most common type of kidney cancer (Scelo et al., 2017). rs4903064 is located in the third intron of the *DPF3* gene, which encodes a protein of the BAF subfamily of the SWI/SNF chromatin-remodeling complex. The predominance of the C allele of rs4903064 in tumor tissue samples from patients with clear-cell renal carcinoma has been demonstrated, in contrast to both normal kidney tissues and tumor tissue from individuals with papillary renal cell carcinoma and individuals with chromophobe renal cell carcinoma. In this context, a measurement of the ratio of alleles in *DPF3* pre-mRNA in patients with clear-cell renal carcinoma heterozygous for rs4903064 has confirmed a skew toward the C allele (Protze et al., 2022).

According to ATAC-seq data obtained on primary renal cancer cells, rs4903064 is located in an open chromatin region, in a putative enhancer. Given that the substitution of T with C (rs4903064) creates a potential binding site for TF HIF, a reporter analysis has been performed on HeLa and MCF-7 cell lines, showing that an increase in reporter gene expression takes place only in the case of the C allele and only when cells are treated with a specific stabilizer of HIF: dimethyloxalylglycine. Knockouts of various isoforms of HIF have revealed that the increase in reporter activity upon stimulation with dimethyloxalylglycine depends on HIF-1α. By means of ChIP-qPCR in primary renal tubular cells with different genotypes of rs4903064 (TT, CT, and CC), enhanced binding of HIF-1α and HIF-1β to risk allele C has been confirmed. To elucidate the role of overexpression of *DPF3* in the development of clear-cell renal carcinoma, a knockout of *DPF3* has been performed in cells of proximal renal tubules using CRISPR/Cas9 technology. It was demonstrated in that report that cells with defective expression of *DPF3* grow more slowly than control clones of the corresponding cells, suggesting that increased expression of *DPF3* in proximal tubule cells stimulates proliferation (Protze et al., 2022).

rs17114036

rs17114036 (T/C) is associated with coronary heart disease and ischemic stroke according to GWASs (Dichgans et al., 2014). This SNP is located in the 5th intron of the *PLPP3* gene, encoding phospholipid phosphatase 3, which inhibits inflammation of endothelium and contributes to the integrity of its monolayer (Panchatcharam et al., 2014; Wu et al., 2015). Experiments with ATAC-seq and ChIP-seq (H3K27ac and H3K4me2) performed on human aortic endothelial cells (HAECs) have identified the region containing rs17114036 as a potential enhancer. The enhancer activity of this region was confirmed by reporter analysis, whereas protective allele C (as compared to the T allele) significantly increased the activity of luciferase upon transfection of vector constructs into HAECs. A deletion of a 66-bp region containing rs17114036 by means of CRISPR/Cas9 significantly reduced the expression of *PLPP3* as compared to the unedited genome and enhanced the permeability of the monolayer of edited HAECs. As a result of modeling hemodynamic processes, an increase in the activity of the studied enhancer (containing the C allele) in HAECs was shown with an 18-hour "atheroprotective" flow as compared to an "atherogenic" flow, while no effect of the T allele was found. It turned out that the substitution of the T nucleotide with C creates a binding site (CACC) for the KLF2 protein, as confirmed by ChIP-AS-qPCR analysis in HAECs heterozygous for rs17114036. Cotransfection experiments with a plasmid causing overexpression of KLF2 also showed higher luciferase activity in the case of the tested enhancer carrying the C allele (Krause et al., 2018).

rs4407214

rs4407214 (T/G) is associated with estrogen receptor-negative breast cancer. In the HMEC cell line, an analysis of ChIP-seq data from the ENCODE project regarding locations of marks of active chromatin (H3K4m1, H3K4m2, H3K4m3, H3K9ac, H3K27ac, H3K36m3, H3K79m2, H4K20m1, EZH2, and H2AZ) has helped to find in the rs4407214 region a regulatory locus in intron 1 of the *WDR43* gene (a protein-coding gene associated with rRNA processing and ribosomal biogenesis). By EMSA, the researchers showed ASB of nuclear proteins from MCF10A and CAL-51 cells to DNA probes mimicking the region of this SNP's location in the genome. Reporter gene expression under the control of the identified regulatory region on the same cell lines was also found to be allele-dependent (Couch et al., 2016; Fachal et al., 2020).

Bioinformatic analysis was then performed using the JASPAR database and available ChIP-seq data for this region from the ENCODE project to identify the TFs the binding sites of which are altered by the G-to-T substitution (rs4407214). As a result, USF1 was identified as such a TF. A competitive EMSA with nuclear proteins isolated from CAL-51 and MCF10A cells confirmed the ASB of USF1 in the case of the G allele. It was also demonstrated that CRISPR/Cas9-mediated removal of the presumed regulatory region containing rs4407214 results in underexpression of *PLB1* (phospholipase B1), which is located at a distance of ~400 kbp from *WDR43*.

Mass interpretation of GWAS data by functional genomics methods

Because results of GWASs tend to associate a trait (disease) with multiple loci (Goldstein, 2009; Boyle et al., 2017), most of which in turn contain many often coinherited SNPs (Tak, Farnham, 2015; Schaid et al., 2018), it is very difficult to choose among them the potential rSNPs involved in a disease's pathogenesis. To date, several effective experimental solutions to this problem have been developed. First of all, these are scaled up versions of the approaches utilized to investigate individual rSNPs: a reporter assay (MPRA) and methods for studying protein–nucleic acid interactions (Reel-seq, SNP-seq, and SNP-SELEX) (Zhao et al., 2020; Lu et al., 2021; Yan et al., 2021).

In particular, MPRA has been successfully used to find the rSNPs that play a key role in the genetic predisposition to lupus erythematosus (Lu et al., 2021). According to GWAS findings, 3,073 SNPs in 91 loci are associated with this disease. Those researchers constructed a barcoded library containing 12,396 170-bp oligonucleotides containing in the middle all known variants of these 3,073 SNPs; these oligos were inserted upstream of a minimal promoter placed before the *eGFP* gene. An influence of a minor allele on enhancer activity of inserts for 51 SNPs from 27 loci was demonstrated in the GM12878

cell line. The small number of identified rSNPs can most likely be explained by transfection of only one cell line; if several cell lines had been tested, the number of rSNPs would have been much larger due to expansion of the set of TFs involved in the study. Similarly, in MPRA, 30 out of 832 SNPs (associated with melanoma risk according to GWASs) showed a significant difference in the impact of minor alleles on reporter gene expression in the UACC903 melanoma cell line (Choi et al., 2020). For one of these 30 SNPs (rs398206), which is located in intron 1 of the *MX2* gene, the difference was the largest. It turned out that risk allele A of rs398206 significantly enhances the binding of TF YY1 *in vitro* (in EMSA) and *in vivo* (according to ChIP-AS-qPCR), thereby upregulating *MX2* and thus contributing to the initiation of melanoma (Choi et al., 2020). Other examples can be found in refs (Ulirsch et al., 2016; Liu S. et al., 2017; Kalita et al., 2018; Klein et al., 2019).

By methodically similar Reel-seq and SNP-seq, which are based on a comparative analysis of the binding of a TF to oligonucleotides containing minor alleles, 521 potential rSNPs have been selected out of 4,316 SNPs correlating with breast cancer according to GWASs (Zhao et al., 2020) as well as 403 possible rSNPs out of 903 SNPs associated with prostate cancer (Zhang et al., 2018). The largest study involving the approach from this research group was published in 2021 (Yan et al., 2021). The approach was named SNP-SELEX. To implement it, those authors used a library of 383,544 40-bp oligonucleotides containing in the middle all possible alleles of 95,886 SNPs. SNPs were chosen based on either their association with type 2 diabetes mellitus in GWASs or localization within a 500-kbp window containing a variant associated with this pathology. By means of 270 recombinant TFs, those authors conducted a multiplex analysis of their binding to the oligonucleotides and identified 11,079 SNPs having an appreciable allele effect on binding to at least one TF.

For mass interpretation of the results of GWASs, functional genomics data from available databases are also widely used, such as data on genome-wide profiles (ChIP-seq) of TFs' binding (Li S. et al., 2020) and of histone modifications (Jones et al., 2020), on open-chromatin genome-wide profiles (ATAC-seq) (Corces et al., 2020), on three-dimensional chromatin contacts (Corces et al., 2020), and on locations of enhancer and superenhancer regions (Gong et al., 2018; Sun W. et al., 2018; Nasser et al., 2021) as well as data from eQTL analysis (Gamazon et al., 2018; Zheng et al., 2019; Barbeira et al., 2021).

For example, 8,005 SNPs – either directly associated by GWASs with major depressive disorder or present in the same linkage group – have been mapped to ChIP-seq peaks obtained from a brain tissue or cells of neuronal origin by means of antibodies to 34 various TFs (Li S. et al., 2020). After that, a search was performed for binding sites of the corresponding TFs in the regions of the SNPs using a database containing 7,699 position weight matrices (Whitington et al., 2016), and it was revealed that 34 SNPs disrupt the binding sites of 15 TFs. A reporter assay confirmed the effect of an allele on gene expression for 29 SNPs. One of them, rs3101339, proved to be located at a potential binding site of TF REST in the promoter region of the *NEGR1* gene, whereas the substitution of A with C considerably damaged the structure of this site, as evidenced by a decrease in reporter gene expression under the control of an insert carrying the C allele. The influence of rs3101339 on *NEGR1* gene expression *in vivo* was confirmed by elimination of an appropriate DNA fragment via CRISPR-Cas9 genomic editing. Since the product of *NEGR1* plays an important part in the maintenance of required density of dendritic spines, its underexpression when A is replaced by C may substantially contribute to the development of a depressive state (Li S. et al., 2020).

Data on genome-wide profiles of active-chromatin histone marks are also turning out to be very informative for the functional interpretation of GWAS results. For example, to analyze many SNPs correlating with epithelial ovarian cancer, they have been mapped in the region of ChIP-seq peaks for H3K27Ac; these peaks were obtained by the researchers in a study on 26 tissue samples of this type of cancer (Jones et al., 2020). Then, using the motifbreakR tool (Coetzee et al., 2015), among the mapped SNPs, 469 SNPs were selected in which the nucleotide substitution substantially altered a binding site of some TF. The most frequent was the change in the sequence of the binding site of TF REST, for which there are data on its tumor-suppressive and oncogenic functions (Jones et al., 2020). Besides, the use of ChIP-seq datasets on various histone modifications from relevant databases in combination with data and tools from a database of potential regulatory variants (rVarBase) (Guo et al., 2016) has made it possible to detect in superenhancers 286 and 366 possible rSNPs associated with type 2 diabetes mellitus (Sun W. et al., 2018) and coronary heart disease (Gong et al., 2018), respectively, that alter the predicted TFBSs.

GWAS-unrelated function-based approaches to identifying potential rSNPs

Modern massive function-based approaches to the identification of potential rSNPs are mainly based on the registration of an effect of a nucleotide substitution on some molecular phenotype. This may be (i) determination (in transcriptomes) of a difference in the expression level of individual genes among homozygotes and heterozygotes for different alleles of each SNP (eQTL analysis), (ii) identification of SNPs showing asymmetry of enrichment within transcriptome data (RNA-seq: ASE events) or within epigenomic data (DNase-seq, ChIP-seq, and ATAC-seq: ASB events), or (iii) determination of an influence of an allele on reporter gene expression by MPRA.

eQTL analysis

The term "eQTL" either means that there is a correlation between a variant (eVariant) and the expression level of a certain gene(s) (eGene[s]) (GTEx Consortium, 2017, 2020) or refers directly to an SNP the alleles of which show such a correlation; the term is used much more frequently in the latter sense (Fairfax et al., 2014; Fan et al., 2020; Jiang et al., 2020; Werling et al., 2020). Transcriptomic data obtained using either microarrays (Fairfax et al., 2014; Westra, Franke, 2014) or RNA-seq (GTEx Consortium, 2020) are suitable for finding eQTLs. These data are quite sufficient for the detection of eQTLs located in transcribed regions (Göring et al., 2007), whereas the identification of their entire set also requires genome sequencing data (GTEx Consortium, 2020;

Werling et al., 2020). In contrast to GWASs, which require biological samples from many thousands of individuals (Tam et al., 2019), several hundred participants are sufficient for eQTL analysis (Westra et al., 2013; Fairfax et al., 2014; GTEx Consortium, 2020). Nonetheless, just as in GWASs, in eQTL analyses, the problem of distinguishing an SNP that is indeed relevant to the formation of a trait – among marker variants detected through linkage disequilibrium – is still relevant (Zou et al., 2019; Umans et al., 2021).

The largest-scale project on obtaining transcriptomic data and identifying eQTLs is international consortium GTEx, within which RNA-seq data on 15,201 postmortem samples of 49 tissues collected from 838 donors have been collected, allowing to identify 4,278,636 eQTLs associated with changes in the expression of 18,262 and 5,006 genes encoding proteins and long intergenic noncoding RNAs, respectively (GTEx Consortium, 2020).

There are other datasets of eQTLs, including those obtained not on postmortem but on biopsy materials (Fairfax et al., 2014; Stolze et al., 2020). For example, in a study by Stolze et al., in an analysis of transcriptomes (RNA-seq) of the aortic endothelium of 157 donors, the investigators identified thousands of eQTLs not registered in the GTEx Consortium data (Stolze et al., 2020). Fairfax et al. have employed $CD14^+$ monocytes (derived from healthy individuals) treated *in vitro* with either interferon gamma (for 24 h, 367 individuals) or bacterial cell wall lipopolysaccharide (LPS) (2 and 24 h, 261 and 322 individuals, respectively). $CD14^+$ monocytes from 414 people served as controls there. Through microarray RNA profiling and genotyping, 609,704 SNPs (minor allele frequency > 0.04) were examined and 21,516 eQTLs were detected, 24.6 % of which manifested themselves in control cells, 21.6 % of which manifested themselves after 2 h of treatment with LPS, and 25.4 and 28.3 % after 24 h treatment with LPS and IFN-γ, respectively. The results of this work point to an important role of genomic variants in the nature of transcriptomic response to a drug (Fairfax et al., 2014). In conclusion, it should be noted that any dataset of transcriptomic data obtained from hundreds or more individuals can be used to search for eQTLs, as, for example, has been done by us (Korbolina et al., 2021) with the help of data from RNA-seq analysis of postmortem brain samples from 96 individuals (Ramaker et al., 2017).

At present, quantitative expression trait loci analysis is mainly utilized to identify groups of genes participating in trait formation (Hormozdiari et al., 2016; Morrow et al., 2018; Gamazon et al., 2019; Ratnapriya et al., 2019; Jaffe et al., 2020). Additionally, its results are often used to prioritize GWAS-identified SNPs for their subsequent rigorous experimental investigation. An example is rs13239597, which is located in the *TNPO3* gene promoter and associated with lupus erythematosus and multiple sclerosis according to GWASs. eQTL analysis of transcriptomes of lymphoblastoid cell lines derived from 373 individuals has not revealed any impact of rs13239597 on *TNPO3* expression but uncovered a significant association of the A allele of this SNP with overexpression of the *IRF5* gene [which is located 118 kbp away (Thynn et al., 2020)], in agreement with findings of the GTEx Consortium (GTEx Consortium, 2017). Next, an analysis of available Hi-C data showed that *IRF5* is one of 12 genes that are in direct contact with the rs13239597 region. A computational analysis of motifs that potentially change affinity for a TF as a result of a nucleotide substitution (Coetzee et al., 2015) has revealed four such TFs: EVI1, ERF, GATA1, and TAL1. By ChIP-AS-qPCR, it has been demonstrated that EVI1 binds much better to the rs13239597 region in the case of the A allele as compared to the C allele (Thynn et al., 2020). Similar examples can also be found in refs (Roca-Ayats et al., 2019; Jiang et al., 2020; Tian et al., 2020).

A large-scale search for ASE and ASB events

The development of next-generation-sequencing-based methods of transcriptomic analysis (RNA-seq) and epigenomic analysis (ChIP-seq, DNase-seq, and ATAC-seq) has opened up a unique opportunity for quantifying a difference in the enrichment of two alleles (an allele imbalance) of each heterozygous polymorphic site of a diploid organism within the respective dataset (Maurano et al., 2015; Cavalli et al., 2016a, b; Castel et al., 2020; Fan et al., 2020; Xu et al., 2020; Korbolina et al., 2021). An important feature of these approaches to the identification of potential rSNPs – in contrast to eQTL analysis (and even more so in contrast to GWASs investigating SNPs of many individuals in various genomic contexts and living conditions) – is that allele-asymmetric events are recorded for each individual and the backgrounds are identical. This arrangement enables researchers to obtain reliable data when studying a very small number of individuals, down to one (Harvey et al., 2015). An increase in sample size is required only for involving in the analysis a larger number of SNPs that are in a heterozygous state. For instance, calculations show that data from 20 individuals theoretically allow to determine ASE or ASB events for 65–70 % of SNPs that have a population frequency of ≥5 % (Cavalli et al., 2016a).

For this reason, possibilities of pharmacogenetic and pharmacogenomic projects become much more abundant. The most striking example of such a project is a simultaneous analysis of allele-specific effects of 50 substances (steroid and peptide hormones, nutrients, commonly used drugs, and a number of environmental pollutants) in primary cultures of five cell types (LCLs, PBMCs, HUVECs, SMCs, and melanocytes), each of which is represented by cell samples from three individuals (Moyerbrailean et al., 2016). An analysis of the resultant transcriptomic data helped to identify more than 300 SNPs, an imbalance in the enrichment of the alleles of which within transcriptomes emerged or increased significantly in response to treatment with one or another drug. Via the same approach, inducer (LPS of the bacterial cell wall)-dependent ASE events have been identified in 19 immune response genes by an analysis of transcriptomes of blood mononuclear cells from eight individuals (Edsgärd et al., 2016) as well as 561 ASE events responsive to treatment of $CD4^+$ T cells (from 24 genotyped individuals) with immobilized anti-CD3/CD28 antibodies (Gutierrez-Arcelus et al., 2020). These results open up a new – unrelated to any *a priori* hypothesis – way to elucidate mechanisms of individual sensitivity to drugs.

The largest dataset of ASE events at present, containing 431 million such events, is derived from data on RNA-seq and whole-genome sequencing from the GTEx Consortium (GTEx Consortium, 2020) and is published in a paper by Castel et

al. (Castel et al., 2020). This dataset can serve as a source for mass discovery of rSNPs, for example, in a comparison with data from GWASs or from eQTL analysis. It is worth noting that in the absence of whole-genome sequencing data, relevant information can be acquired by more sophisticated methods of bioinformatic search for allele-specific events directly in RNA-seq data (Harvey et al., 2015; Moyerbrailean et al., 2016; Fan et al., 2020; Korbolina et al., 2021).

ChIP-seq experiments based on antibodies to various TFs make it possible to directly register events of allele-asymmetric interaction of these proteins with their binding sites in the case of a heterozygous state of SNPs at these sites. In a pioneering work in the laboratory of Claus Wadelius, they analyzed all the then-available data from the ENCODE project on binding profiles of TFs in cell lines GM12878 (B cells), H1-hESC, K562, and SK-N-SH, thereby revealing 9,962 SNPs featuring an allelic imbalance in the binding of a TF (ASB) (Cavalli et al., 2016b). By the same approach, 3,713 SNPs have been found showing an allelic imbalance in the binding of TFs in HepG2 and HeLa-S3 cells; testing 39 of them in a luciferase reporter system has confirmed the effect of an allele on reporter gene expression for 27 SNPs (Cavalli et al., 2016a). A detailed analysis of one of them, rs953413, indicates that the A allele disrupts the binding site of TF FOXA, resulting in reduced binding of not only this TF but also of TF HNF4α cooperatively interacting with it, thereby ultimately leading to underexpression of *ELOVL2* and possibly serving as a factor in the pathogenesis of non-alcoholic fatty liver disease (Pan et al., 2020).

Because allele-asymmetric alterations in profiles of histone modifications and of open chromatin can reflect SNP-induced changes in the binding of TFs (Kar et al., 2014; Hatayama, Aruga, 2018; Huang et al., 2018; Yi et al., 2020), these data are also widely utilized to find ASB events. For instance, Maurano et al. (Maurano et al., 2015) have examined 493 open-chromatin profiles (DNase-seq) obtained in various cell lines, where 64,599 SNPs with ASB have been found. Their bioinformatic analysis using position weight matrices for 2,203 motifs of TFBSs from different sources indicates that most of the identified SNPs can affect the binding of TFs and hence the accessibility of the respective DNA regions to DNase I (Maurano et al., 2015). A newer technique for detecting open chromatin, ATAC-seq, is based on the ability of a hyperactive mutant of Tn5 transposase to detect open DNA regions in chromatin (Marinov, Shipony, 2021) and is also used to search for ASB events. In particular, it has been employed to identify 53 rSNPs in breast cancer MCF-7 cells and 125 rSNPs in a line of human mesenchymal stem cells (MSCs); in total, 30 % of the rSNPs found in MCF-7 cells and 43 % of those found in MSCs have been identified as eQTLs in GTEx data, indicating their influence on gene expression (Xu et al., 2020). Examples of use of data from ChIP-seq (involving antibodies to histone marks) for registering ASB events can be found in refs (Sun J. et al., 2016; D'Oliveira Albanus et al., 2021; Li M. et al., 2021).

The combination of searches for ASE events and ASB events can be considered the most productive approach to identifying rSNPs. For example, in our work (Korbolina et al., 2018), at first, ASB events were identified in ChIP-seq data from the ENCODE project for histone modifications H3K27ac, H3K4me1, H3K4me2, H3K4me3, and H3K27me3 as well as for 456 TFs and their associated proteins in human cell lines K562, MCF-7, and HCT-116. Then, by means of RNA-seq data obtained from the same cell lines, SNPs (rSNPs) that are associated with changes in gene expression levels were identified. According to GWASs, out of 1,633 rSNPs found in this way, 27 have shown associations with cancers (Korbolina et al., 2018), and 14, with cognitive disorders (Bryzgalov et al., 2018). Another 30 rSNPs have been implicated in colorectal cancer with the help of data from the International Cancer Genome Consortium (ICGC) (Seshagiri et al., 2012). Genotyping of patients with colorectal cancer and healthy individuals for six of these SNPs has revealed an association of rs590352, rs4796672, and rs2072580 with this disease (Leberfarb et al., 2020). A correlation with breast cancer has been found for rs2072580 (Degtyareva et al., 2020). Later, the same approach has allowed to identify 14,266 rSNPs during the processing of data obtained in a study by Reyes-Palomares et al. (Reyes-Palomares et al., 2020) on H3K4me3 histone marker profiling (ChIP-seq) and RNA-seq data on pulmonary-artery epithelial samples from 19 individuals (Korbolina et al., 2021).

MPRA

Research into the effect of polymorphic-site alleles on reporter gene expression via simultaneous transfection of hundreds and thousands of barcoded plasmid constructs into eukaryotic cells with subsequent transcriptome sequencing is also an informative approach to finding rSNPs (Vockley et al., 2015; Tewhey et al., 2016; Movva et al., 2019). The largest-scale study based on this approach has been conducted in the laboratory of Bas van Steensel (van Arensbergen et al., 2019). Using a promoterless plasmid and fragmented genomes (fragment length 150–500 bp) of four individuals belonging to different ethnic groups, two barcoded libraries were constructed for each individual, where inserts were expected to play the role of a promoter. At the same time, on the basis of data on transcription initiation in enhancer regions (Natoli, Andrau, 2012; van Arensbergen et al., 2017), those authors expected to detect not only promoters but also enhancers. The use of DNA from humans of genetically distant ethnic groups allowed those investigators to hope for an analysis of the largest possible number of polymorphic sites that are homozygous for different alleles in at least two of those people.

After transfection of K562 and HepG2 cells with the resulting libraries, 19 and 14 thousand potential rSNPs, respectively, were found, most of which did not overlap, once again indicating tissue specificity of the supragenomic (protein) regulatory machine. The identified SNPs showed significant enrichment within regulatory regions of the genome. In this case, the enrichment (approximately 15-fold) was three times higher in promoter regions than in enhancer regions (approximately 5-fold); this outcome is obviously due to the design of the reporter constructs. For several rSNPs, by mass-spectrometric analysis of proteins interacting with oligonucleotides containing minor alleles, those authors were able to identify TFs the binding sites of which are altered by a nucleotide substitution. In particular, the A allele of rs623853 was found to disrupt the binding of TFs of the ELF family, whereas the C allele of rs554591 weakens the binding of ZNF787 while enhancing the binding of KLF and SP (van Arensbergen et al., 2019).

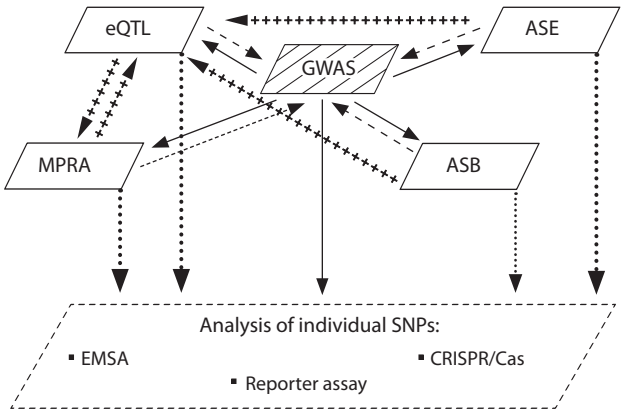

Fig. 2. The scheme of an integrative approach to the search for SNPs functionally important for the development of a trait (pathology) on the basis of two opposite principles: from an association with a trait to its function (solid arrows) and vice versa, from a function to a trait (dashed arrows). Cross-containing arrows show ways to increase the evidence base for functional significance of sets of SNPs, and dotted arrows represent ways to study individual SNPs in detail.

Conclusion

Programs – of coordinated switching on, switching off, and changes of the expression of various genes – that underlie (1) ontogenesis events, (2) the existence of many types of differentiated cells, and (3) the abilities of cells to respond to various factors of the external and internal environment are implemented by the regulatory part of the genome of multicellular organisms. The information encoded in the regulatory regions is converted into a desired pattern of gene expression primarily via the binding of TFs to specific sequences in the regulatory regions (e. g., promoters, enhancers, and silencers) (Lan et al., 2012; Merkulova et al., 2013; Dubois-Chevalier et al., 2018; Chen, Pugh, 2021; Tobias et al., 2021). According to present-day concepts, the SNPs located in regulatory regions of genes, affecting binding sites of TFs, and changing the level of gene expression play a central part in the variation of phenotypic traits, including predisposition/resistance to multifactorial diseases (Maurano et al., 2015; Deplancke et al., 2016; Carrasco Pro et al., 2020). In this regard, there is a strong interest both in functional interpretation of SNPs having an association with various diseases (primarily according to GWAS data) and in the development of massive function-based approaches to the discovery of rSNPs. Interpretation of data from GWASs is carried out either at the level of individual SNPs or for all SNPs collectively by a variety of functional genomics techniques (Fig. 2). Meanwhile, the same methods of functional genomics are used for personal searches for rSNPs, but at the same time, it is necessary to solve the inverse problem: determining a relation between the found rSNPs and a trait (disease). The most popular solution to this problem is to compare the obtained data with available information from GWASs (see Fig. 2). On the other hand, in this way, usually only 1.5–3.0 % of found rSNPs are implicated in various traits (Cavalli et al., 2016b, 2019; Korbolina et al., 2021). In this regard, it seems very promising to take advantage of results of eQTL analysis, enabling an investigator to determine an influence of many specific rSNPs on the expression of fairly large groups of genes, the subsequent analysis of which by modern functional annotation tools (Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, and others) allows to get an idea about a possible affected trait (Korbolina et al., 2021).

To sum up, there are two approaches – based on opposite principles – to finding SNPs that are important for the development of a trait (pathology): on the one hand, starting from data on an association of an SNP with some trait, and on the other hand, starting from determination of allele-specific changes at the molecular level (in the transcriptome or regulome). It can be concluded that comprehensive use of the two approaches appreciably enriches our knowledge about the participation of genetic determinants in molecular mechanisms of trait formation, including predisposition to multifactorial diseases.

References

Antontseva E.V., Matveeva M.Y., Bondar N.P., Kashina E.V., Leberfarb E.Y., Bryzgalov L.O., Gervas P.A., Ponomareva A.A., Cherdyntseva N.V., Orlov Y.L., Merkulova T.I. Regulatory single nucleotide polymorphisms at the beginning of intron 2 of the human *KRAS* gene. *J. Biosci.* 2015;40(5):873-883. DOI 10.1007/s12038-015-9567-8.

Barbeira A.N., Bonazzola R., Gamazon E.R., Liang Y., Park Y., Kim-Hellmuth S., Wang G., Jiang Z., Zhou D., Hormozdiari F., Liu B., Rao A., Hamel A.R., Pividori M.D., Aguet F., Bastarache L., Jordan D.M., Verbanck M., Do R., Stephens M., Ardlie K., McCarthy M., Montgomery S.B., Segrè A.V., Brown C.D., Lappalainen T., Wen X., Im H.K. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol.* 2021;22(1):49. DOI 10.1186/s13059-020-02252-4.

Boldes T., Merenbakh-Lamin K., Journo S., Shachar E., Lipson D., Yeheskel A., Pasmanik-Chor M., Rubinek T., Wolf I. R269C variant of ESR1: high prevalence and differential function in a subset of pancreatic cancers. *BMC Cancer*. 2020;20(1):531. DOI 10.1186/s12885-020-07005-x.

Boyle E.A., Li Y.I., Pritchard J.K. An expanded view of complex traits: from polygenic to omnigenic. *Cell*. 2017;169(7):1177-1186. DOI 10.1016/j.cell.2017.05.038.

Bryzgalov L.O., Antontseva E.V., Matveeva M.Y., Shilov A.G., Kashina E.V., Mordvinov V.A., Merkulova T.I. Detection of regulatory SNPs in human genome using ChIP-seq ENCODE data. *PLoS One*. 2013;8(10):e78833. DOI 10.1371/journal.pone.0078833.

Bryzgalov L.O., Korbolina E.E., Brusentsov I.I., Leberfarb E.Y., Bondar N.P., Merkulova T.I. Novel functional variants at the GWAS-implicated loci might confer risk to major depressive disorder, bipolar affective disorder and schizophrenia. *BMC Neurosci.* 2018;19(S1):22. DOI 10.1186/s12868-018-0414-3.

Buniello A., MacArthur J.A.L., Cerezo M., Harris L.W., Hayhurst J., Malangone C., McMahon A., Morales J., Mountjoy E., Sollis E., Suveges D., Vrousgou O., Whetzel P.L., Amode R., Guillen J.A., Riat H.S., Trevanion S.J., Hall P., Junkins H., Flicek P., Burdett T., Hindorff L.A., Cunningham F., Parkinson H. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 2019;47(D1):D1005-D1012. DOI 10.1093/nar/gky1120.

Carrasco Pro S., Bulekova K., Gregor B., Labadorf A., Fuxman Bass J.I. Prediction of genome-wide effects of single nucleotide variants on transcription factor binding. *Sci. Rep.* 2020;10(1):17632. DOI 10.1038/s41598-020-74793-4.

Castel S.E., Aguet F., Mohammadi P., Aguet F., Anand S., Ardlie K.G., Gabriel S., Getz G.A., Graubert A., Hadley K., … Moore H.M., Nierras C.R., Rao A.K., Vaught J.B., Volpi S., Ardlie K.G., Lappalainen T. A vast resource of allelic expression data spanning human tissues. *Genome Biol.* 2020;21(1):234. DOI 10.1186/s13059-020-02122-z.

Cavalli M., Pan G., Nord H., Wallén Arzt E., Wallerman O., Wadelius C. Allele-specific transcription factor binding in liver and cervix cells unveils many likely drivers of GWAS signals. *Genomics*. 2016a;107(6):248-254. DOI 10.1016/j.ygeno.2016.04.006.

Cavalli M., Pan G., Nord H., Wallerman O., Wallén Arzt E., Berggren O., Elvers I., Eloranta M.-L., Rönnblom L., Lindblad Toh K., Wadelius C. Allele-specific transcription factor binding to common and rare variants associated with disease and gene expression. *Hum. Genet.* 2016b;135(5):485-497. DOI 10.1007/s00439-016-1654-x.

Cavalli M., Baltzer N., Umer H.M., Grau J., Lemnian I., Pan G., Wallerman O., Spalinskas R., Sahlén P., Grosse I., Komorowski J., Wadelius C. Allele specific chromatin signals, 3D interactions, and motif predictions for immune and B cell related diseases. *Sci. Rep.* 2019;9(1):2695. DOI 10.1038/s41598-019-39633-0.

Chanock S. Candidate genes and single nucleotide polymorphisms (SNPs) in the study of human disease. *Dis. Markers*. 2001;17(2): 89-98. DOI 10.1155/2001/858760.

Chen H., Pugh B.F. What do transcription factors interact with? *J. Mol. Biol.* 2021;433(14):166883. DOI 10.1016/j.jmb.2021.166883.

Choi J., Zhang T., Vu A., Ablain J., Makowski M.M., Colli L.M., Xu M., Hennessey R.C., Yin J., Rothschild H., Gräwe C., Kovacs M.A., Funderburk K.M., Brossard M., Taylor J., Pasaniuc B., Chari R., Chanock S.J., Hoggart C.J., Demenais F., Barrett J.H., Law M.H., Iles M.M., Yu K., Vermeulen M., Zon L.I., Brown K.M. Massively parallel reporter assays of melanoma risk variants identify *MX2* as a gene promoting melanoma. *Nat. Commun.* 2020;11(1):2718. DOI 10.1038/s41467-020-16590-1.

Claussnitzer M., Cho J.H., Collins R., Cox N.J., Dermitzakis E.T., Hurles M.E., Kathiresan S., Kenny E.E., Lindgren C.M., MacArthur D.G., North K.N., Plon S.E., Rehm H.L., Risch N., Rotimi C.N., Shendure J., Soranzo N., McCarthy M.I. A brief history of human disease genetics. *Nature*. 2020;577(7789):179-189. DOI 10.1038/s41586-019-1879-7.

Coetzee S.G., Coetzee G.A., Hazelett D.J. motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. *Bioinformatics*. 2015;31(23):3847-3849. DOI 10.1093/bioinformatics/btv470.

Cooper D. The human gene mutation database. *Nucleic Acids Res.* 1998;26(1):285-287. DOI 10.1093/nar/26.1.285.

Corces M.R., Shcherbina A., Kundu S., Gloudemans M.J., Frésard L., Granja J.M., Louie B.H., Eulalio T., Shams S., Bagdatli S.T., Mumbach M.R., Liu B., Montine K.S., Greenleaf W.J., Kundaje A., Montgomery S.B., Chang H.Y., Montine T.J. Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat. Genet.* 2020;52(11): 1158-1168. DOI 10.1038/s41588-020-00721-x.

Couch F.J., Kuchenbaecker K.B., Michailidou K., Mendoza-Fandino G.A., Nord S., Lilyquist J., Olswold C., Hallberg E., Agata S., Ahsan H., … Slager S., Chenevix-Trench G., Pharoah P.D.P., Monteiro A.A.N., García-Closas M., Easton D.F., Antoniou A.C. Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nat. Commun.* 2016;7(1):11375. DOI 10.1038/ncomms11375.

Degtyareva A.O., Leberfarb E.Y., Efimova E.G., Brusentsov I.I., Usova A.V., Lushnikova E.L., Merkulova T.I. *rs2072580T>A* polymorphism in the overlapping promoter regions of the *SART3* and *ISCU* genes associated with the risk of breast cancer. *Bull. Exp. Biol. Med.* 2020;169(1):81-84. DOI 10.1007/s10517-020-04829-2.

Deplancke B., Alpern D., Gardeux V. The genetics of transcription factor DNA binding variation. *Cell*. 2016;166(3):538-554. DOI 10.1016/j.cell.2016.07.012.

Dichgans M., Malik R., König I.R., Rosand J., Clarke R., Gretarsdottir S., Thorleifsson G., Mitchell B.D., Assimes T.L., Levi C., … Willenborg C., Laaksonen R., Voight B.F., Stewart A.F.R., Rader D.J., Hall A.S., Kooner J.S. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*. 2014;45(1):24-36. DOI 10.1161/STROKEAHA.113.002707.

D'Oliveira Albanus R., Kyono Y., Hensley J., Varshney A., Orchard P., Kitzman J.O., Parker S.C.J. Chromatin information content landscapes inform transcription factor and DNA interactions. *Nat. Commun.* 2021;12(1):1307. DOI 10.1038/s41467-021-21534-4.

Dubois-Chevalier J., Mazrooei P., Lupien M., Staels B., Lefebvre P., Eeckhoute J. Organizing combinatorial transcription factor recruitment at *cis*-regulatory modules. *Transcription*. 2018;9(4):233-239. DOI 10.1080/21541264.2017.1394424.

Edsgärd D., Iglesias M.J., Reilly S.-J., Hamsten A., Tornvall P., Odeberg J., Emanuelsson O. GeneiASE: detection of condition-dependent and static allele-specific expression from RNA-seq data without haplotype information. *Sci. Rep.* 2016;6(1):21134. DOI 10.1038/srep21134.

Fachal L., Aschard H., Beesley J., Barnes D.R., Allen J., Kar S., Pooley K.A., Dennis J., Michailidou K., Turman C., … Edwards S.L., Antoniou A.C., Chenevix-Trench G., Simard J., Easton D.F., Kraft P., Dunning A.M. Fine-mapping of 150 breast cancer risk regions identifies 191 likely target genes. *Nat. Genet.* 2020; 52(1):56-73. DOI 10.1038/s41588-019-0537-1.

Fairfax B.P., Humburg P., Makino S., Naranbhai V., Wong D., Lau E., Jostins L., Plant K., Andrews R., McGee C., Knight J.C. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science.* 2014;343(6175):1246949. DOI 10.1126/science.1246949.

Fan J., Hu J., Xue C., Zhang H., Susztak K., Reilly M.P., Xiao R., Li M. ASEP: gene-based detection of allele-specific expression across individuals in a population by RNA sequencing. *PLoS Genet.* 2020; 16(5):e1008786. DOI 10.1371/journal.pgen.1008786.

Fang J., Jia J., Makowski M., Xu M., Wang Z., Zhang T., Hoskins J.W., Choi J., Han Y., Zhang M., … Amos C.I., Iles M.M., Nathanson K.L., Landi M.T., Vermeulen M., Brown K.M., Amundadottir L.T. Functional characterization of a multi-cancer risk locus on chr5p15.33 reveals regulation of *TERT* by ZNF148. *Nat. Commun.* 2017;8(1):15034. DOI 10.1038/ncomms15034.

Farh K.K.-H., Marson A., Zhu J., Kleinewietfeld M., Housley W.J., Beik S., Shoresh N., Whitton H., Ryan R.J.H., Shishkin A.A., Hatan M., Carrasco-Alfonso M.J., Mayer D., Luckey C.J., Patsopoulos N.A., De Jager P.L., Kuchroo V.K., Epstein C.B., Daly M.J., Hafler D.A., Bernstein B.E. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337-343. DOI 10.1038/nature13835.

Fornes O., Castro-Mondragon J.A., Khan A., van der Lee R., Zhang X., Richmond P.A., Modi B.P., Correard S., Gheorghe M., Baranašić D., Santana-Garcia W., Tan G., Chèneby J., Ballester B., Parcy F., Sandelin A., Lenhard B., Wasserman W.W., Mathelier A. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 2020;48(D1):D87-D92. DOI 10.1093/nar/gkz1001.

Gamazon E.R., Segrè A.V., van de Bunt M., Wen X., Xi H.S., Hormozdiari F., Ongen H., Konkashbaev A., Derks E.M., Aguet F., Quan J., Nicolae D.L., Eskin E., Kellis M., Getz G., McCarthy M.I., Dermitzakis E.T., Cox N.J., Ardlie K.G. Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nat. Genet.* 2018;50(7):956-967. DOI 10.1038/s41588-018-0154-4.

Gamazon E.R., Zwinderman A.H., Cox N.J., Denys D., Derks E.M. Multi-tissue transcriptome analyses identify genetic mechanisms underlying neuropsychiatric traits. *Nat. Genet.* 2019;51(6):933-940. DOI 10.1038/s41588-019-0409-8.

Gao P., Xia J.-H., Sipeky C., Dong X.-M., Zhang Q., Yang Y., Zhang P., Cruz S.P., Zhang K., Zhu J., Lee H.-M., Suleman S., Giannareas N., Liu S., Tammela T.L.J., Auvinen A., Wang X., Huang Q., Wang Liguo, Manninen A., Vaarala M.H., Wang Liang, Schleutker J., Wei G.-H. Biology and clinical implications of the 19q13

aggressive prostate cancer susceptibility locus. *Cell*. 2018;174(3): 576-589.e18. DOI 10.1016/j.cell.2018.06.003.

Goldstein D.B. Common genetic variation and human traits. *N. Engl. J. Med.* 2009;360(17):1696-1698. DOI 10.1056/NEJMp0806284.

Gong J., Qiu C., Huang D., Zhang Y., Yu S., Zeng C. Integrative functional analysis of super enhancer SNPs for coronary artery disease. *J. Hum. Genet.* 2018;63(5):627-638. DOI 10.1038/s10038-018-0422-2.

Göring H.H.H., Curran J.E., Johnson M.P., Dyer T.D., Charlesworth J., Cole S.A., Jowett J.B.M., Abraham L.J., Rainwater D.L., Comuzzie A.G., Mahaney M.C., Almasy L., MacCluer J.W., Kissebah A.H., Collier G.R., Moses E.K., Blangero J. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat. Genet.* 2007;39(10):1208-1216. DOI 10.1038/ng2119.

GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204-213. DOI 10.1038/nature 24277.

GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369(6509):1318-1330. DOI 10.1126/science.aaz1776.

Guo L., Du Y., Qu S., Wang J. rVarBase: an updated database for regulatory features of human variants. *Nucleic Acids Res.* 2016;44(D1): D888-D893. DOI 10.1093/nar/gkv1107.

Gutierrez-Arcelus M., Baglaenko Y., Arora J., Hannes S., Luo Y., Amariuta T., Teslovich N., Rao D.A., Ermann J., Jonsson A.H., Navarrete C., Rich S.S., Taylor K.D., Rotter J.I., Gregersen P.K., Esko T., Brenner M.B., Raychaudhuri S. Allele-specific expression changes dynamically during T cell activation in HLA and other autoimmune loci. *Nat. Genet.* 2020;52(3):247-253. DOI 10.1038/s41588-020-0579-4.

Harvey C.T., Moyerbrailean G.A., Davis G.O., Wen X., Luca F., Pique-Regi R. QuASAR: quantitative allele-specific analysis of reads. *Bioinformatics*. 2015;31(8):1235-1242. DOI 10.1093/bioinformatics/btu802.

Hatayama M., Aruga J. Role of Zic family proteins in transcriptional regulation and chromatin remodeling. In: Aruga J. (Ed.) Zic Family. Advances in Experimental Medicine and Biology. Vol. 1046. Singapore: Springer, 2018;353-380. DOI 10.1007/978-981-10-7311-3_18.

Hindorff L.A., Sethupathy P., Junkins H.A., Ramos E.M., Mehta J.P., Collins F.S., Manolio T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA.* 2009;106(23):9362-9367. DOI 10.1073/pnas.0903103106.

Hormozdiari F., van de Bunt M., Segrè A.V., Li X., Joo J.W.J., Bilow M., Sul J.H., Sankararaman S., Pasaniuc B., Eskin E. Colocalization of GWAS and eQTL signals detects target genes. *Am. J. Hum. Genet.* 2016;99(6):1245-1260. DOI 10.1016/j.ajhg.2016.10.003.

Huang Q., Ma C., Chen L., Luo D., Chen R., Liang F. Mechanistic insights into the interaction between transcription factors and epigenetic modifications and the contribution to the development of obesity. *Front. Endocrinol.* (*Lausanne*). 2018;9:370. DOI 10.3389/fendo.2018.00370.

Jaffe A.E., Hoeppner D.J., Saito T., Blanpain L., Ukaigwe J., Burke E.E., Collado-Torres L., Tao R., Tajinda K., Maynard K.R., Tran M.N., Martinowich K., Deep-Soboslay A., Shin J.H., Kleinman J.E., Weinberger D.R., Matsumoto M., Hyde T.M. Profiling gene expression in the human dentate gyrus granule cell layer reveals insights into schizophrenia and its genetic risk. *Nat. Neurosci.* 2020;23(4):510-519. DOI 10.1038/s41593-020-0604-z.

Jiang Z., Huang Y., Zhang P., Han C., Lu Y., Mo Z., Zhang Z., Li X., Zhao S., Cai F., Huang L., Chen C., Shi Z., Zhang Y., Ling F. Characterization of a pathogenic variant in GBA for Parkinson's disease with mild cognitive impairment patients. *Mol. Brain*. 2020;13(1): 102. DOI 10.1186/s13041-020-00637-x.

Jones M.R., Peng P.-C., Coetzee S.G., Tyrer J., Reyes A.L.P., Corona R.I., Davis B., Chen S., Dezem F., Seo J.-H., Kar S., Dareng E., Berman B.P., Freedman M.L., Plummer J.T., Lawrenson K., Pharoah P., Hazelett D.J., Gayther S.A. Ovarian cancer risk variants are enriched in histotype-specific enhancers and disrupt transcription factor binding sites. *Am. J. Hum. Genet.* 2020;107(4):622-635. DOI 10.1016/j.ajhg.2020.08.021.

Kalita C.A., Brown C.D., Freiman A., Isherwood J., Wen X., Pique-Regi R., Luca F. High-throughput characterization of genetic effects on DNA-protein binding and gene transcription. *Genome Res.* 2018;28(11):1701-1708. DOI 10.1101/gr.237354.118.

Kar S., Parbin S., Deb M., Shilpi A., Sengupta D., Rath S.K., Rakshit M., Patra A., Patra S.K. Epigenetic choreography of stem cells: the DNA demethylation episode of development. *Cell. Mol. Life Sci.* 2014;71(6):1017-1032. DOI 10.1007/s00018-013-1482-2.

Klein J.C., Keith A., Rice S.J., Shepherd C., Agarwal V., Loughlin J., Shendure J. Functional testing of thousands of osteoarthritis-associated variants for regulatory activity. *Nat. Commun.* 2019;10(1): 2434. DOI 10.1038/s41467-019-10439-y.

Knight J.C., Udalova I., Hill A.V.S., Greenwood B.M., Peshu N., Marsh K., Kwiatkowski D. A polymorphism that affects OCT-1 binding to the *TNF* promoter region is associated with severe malaria. *Nat. Genet.* 1999;22(2):145-150. DOI 10.1038/9649.

Korbolina E.E., Brusentsov I.I., Bryzgalov L.O., Leberfarb E.Y., Degtyareva A.O., Merkulova T.I. Novel approach to functional SNPs discovery from genome-wide data reveals promising variants for colon cancer risk. *Hum. Mutat.* 2018;39(6):851-859. DOI 10.1002/humu.23425.

Korbolina E.E., Bryzgalov L.O., Ustrokhanova D.Z., Postovalov S.N., Poverin D.V., Damarov I.S., Merkulova T.I. A panel of rSNPs demonstrating allelic asymmetry in both ChIP-seq and RNA-seq data and the search for their phenotypic outcomes through analysis of DEGs. *Int. J. Mol. Sci.* 2021;22(14):7240. DOI 10.3390/ijms22147240.

Krause M.D., Huang R.-T., Wu D., Shentu T.-P., Harrison D.L., Whalen M.B., Stolze L.K., Di Rienzo A., Moskowitz I.P., Civelek M., Romanoski C.E., Fang Y. Genetic variant at coronary artery disease and ischemic stroke locus 1p32.2 regulates endothelial responses to hemodynamics. *Proc. Natl. Acad. Sci. USA.* 2018;115(48):E11349-E11358. DOI 10.1073/pnas.1810568115.

Lan X., Farnham P.J., Jin V.X. Uncovering transcription factor modules using one- and three-dimensional analyses. *J. Biol. Chem.* 2012; 287(37):30914-30921. DOI 10.1074/jbc.R111.309229.

Lander E.S., Schork N.J. Genetic dissection of complex traits. *Science.* 1994;265(5181):2037-2048. DOI 10.1126/science.8091226.

Lappalainen T. Functional genomics bridges the gap between quantitative genetics and molecular biology. *Genome Res.* 2015;25(10): 1427-1431. DOI 10.1101/gr.190983.115.

Leberfarb E.Y., Degtyareva A.O., Brusentsov I.I., Maximov V.N., Voevoda M.I., Autenshlus A.I., Morozov D.V., Sokolov A.V., Merkulova T.I. Potential regulatory SNPs in the *ATXN7L3B* and *KRT15* genes are associated with gender-specific colorectal cancer risk. *Per. Med.* 2020;17(1):43-54. DOI 10.2217/pme-2019-0059.

Li M., Huang H., Li L., He C., Zhu L., Guo H., Wang L., Liu Jiafeng, Wu S., Liu Jingxin, Xu T., Mao Z., Cao N., Zhang K., Lan F., Ding J., Yuan J., Liu Y., Ouyang H. Core transcription regulatory circuitry orchestrates corneal epithelial homeostasis. *Nat. Commun.* 2021;12(1):420. DOI 10.1038/s41467-020-20713-z.

Li S., Li Y., Li X., Liu J., Huo Y., Wang J., Liu Z., Li M., Luo X.-J. Regulatory mechanisms of major depressive disorder risk variants. *Mol. Psychiatry*. 2020;25(9):1926-1945. DOI 10.1038/s41380-020-0715-7.

Li X.-X., Peng T., Gao J., Feng J.-G., Wu D.-D., Yang T., Zhong L., Fu W.-P., Sun C. Allele-specific expression identified rs2509956 as a novel long-distance *cis*-regulatory SNP for *SCGB1A1*, an important gene for multiple pulmonary diseases. *Am. J. Physiol. Lung. Cell. Mol. Physiol.* 2019;317(4):L456-L463. DOI 10.1152/ajplung.00275.2018.

Liu D., Qin S., Ray B., Kalari K.R., Wang L., Weinshilboum R.M. Single nucleotide polymorphisms (SNPs) distant from xenobiotic response elements can modulate Aryl hydrocarbon receptor function: SNP-dependent CYP1A1 induction. *Drug Metab. Dispos.* 2018;46(9):1372-1381. DOI 10.1124/dmd.118.082164.

Liu S., Liu Y., Zhang Q., Wu J., Liang J., Yu S., Wei G.-H., White K.P., Wang X. Systematic identification of regulatory variants associated with cancer risk. *Genome Biol.* 2017;18(1):194. DOI 10.1186/s13059-017-1322-z.

Lu X., Chen X., Forney C., Donmez O., Miller D., Parameswaran S., Hong T., Huang Y., Pujato M., Cazares T., Miraldi E.R., Ray J.P., de Boer C.G., Harley J.B., Weirauch M.T., Kottyan L.C. Global discovery of lupus genetic risk variant allelic enhancer activity. *Nat. Commun.* 2021;12(1):1611. DOI 10.1038/s41467-021-21854-5.

Ludlow L.B., Schick B.P., Budarf M.L., Driscoll D.A., Zackai E.H., Cohen A., Konkle B.A. Identification of a mutation in a GATA binding site of the platelet glycoprotein Ibβ promoter resulting in the Bernard–Soulier syndrome. *J. Biol. Chem.* 1996;271(36):22076-22080. DOI 10.1074/jbc.271.36.22076.

Malecová B., Morris K.V. Transcriptional gene silencing through epigenetic changes mediated by non-coding RNAs. *Curr. Opin. Mol. Ther.* 2010;12(2):214-222.

Marinov G.K., Shipony Z. Interrogating the accessible chromatin landscape of eukaryote genomes using ATAC-seq. In: Shomron N. (Ed.) Deep Sequencing Data Analysis. Methods in Molecular Biology. Vol. 2243. New York: Humana, 2021;183-226. DOI 10.1007/978-1-0716-1103-6_10.

Maurano M.T., Humbert R., Rynes E., Thurman R.E., Haugen E., Wang H., Reynolds A.P., Sandstrom R., Qu H., Brody J., Shafer A., Neri F., Lee K., Kutyavin T., Stehling-Sun S., Johnson A.K., Canfield T.K., Giste E., Diegel M., Bates D., Hansen R.S., Neph S., Sabo P.J., Heimfeld S., Raubitschek A., Ziegler S., Cotsapas C., Sotoodehnia N., Glass I., Sunyaev S.R., Kaul R., Stamatoyannopoulos J.A. Systematic localization of common disease-associated variation in regulatory DNA. *Science.* 2012;337(6099):1190-1195. DOI 10.1126/science.1222794.

Maurano M.T., Haugen E., Sandstrom R., Vierstra J., Shafer A., Kaul R., Stamatoyannopoulos J.A. Large-scale identification of sequence variants influencing human transcription factor occupancy *in vivo*. *Nat. Genet.* 2015;47(12):1393-1401. DOI 10.1038/ng.3432.

Merkulova T.I., Ananko E.A., Ignatieva E.V., Kolchanov N.A. Transcription regulatory codes of eukaryotic genomes. *Russ. J. Genet.* 2013;49(1):29-45. DOI 10.1134/S1022795413010079.

Moore J.E., Purcaro M.J., Pratt H.E., Epstein C.B., Shoresh N., Adrian J., Kawli T., Davis C.A., Dobin A., Kaul R., … Snyder M.P., Bernstein B.E., Wold B., Hardison R.C., Gingeras T.R., Stamatoyannopoulos J.A., Weng Z. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 2020;583(7818):699-710. DOI 10.1038/s41586-020-2493-4.

Morrow J.D., Cho M.H., Platig J., Zhou X., DeMeo D.L., Qiu W., Celli B., Marchetti N., Criner G.J., Bueno R., Washko G.R., Glass K., Quackenbush J., Silverman E.K., Hersh C.P. Ensemble genomic analysis in human lung tissue identifies novel genes for chronic obstructive pulmonary disease. *Hum. Genomics*. 2018;12(1):1. DOI 10.1186/s40246-018-0132-z.

Movva R., Greenside P., Marinov G.K., Nair S., Shrikumar A., Kundaje A. Deciphering regulatory DNA sequences and noncoding genetic variants using neural network models of massively parallel reporter assays. *PLoS One*. 2019;14(6):e0218073. DOI 10.1371/journal.pone.0218073.

Moyerbrailean G.A., Richards A.L., Kurtz D., Kalita C.A., Davis G.O., Harvey C.T., Alazizi A., Watza D., Sorokin Y., Hauff N., Zhou X., Wen X., Pique-Regi R., Luca F. High-throughput allele-specific expression across 250 environmental conditions. *Genome Res.* 2016;26(12):1627-1638. DOI 10.1101/gr.209759.116.

Nasser J., Bergman D.T., Fulco C.P., Guckelberger P., Doughty B.R., Patwardhan T.A., Jones T.R., Nguyen T.H., Ulirsch J.C., Lekschas F., Mualim K., Natri H.M., Weeks E.M., Munson G., Kane M., Kang H.Y., Cui A., Ray J.P., Eisenhaure T.M., Collins R.L., Dey K., Pfister H., Price A.L., Epstein C.B., Kundaje A., Xavier R.J., Daly M.J., Huang H., Finucane H.K., Hacohen N., Lander E.S., Engreitz J.M. Genome-wide enhancer maps link risk variants to disease genes. *Nature*. 2021;593(7858):238-243. DOI 10.1038/s41586-021-03446-x.

Natoli G., Andrau J.-C. Noncoding transcription at enhancers: general principles and functional models. *Annu. Rev. Genet.* 2012;46(1):1-19. DOI 10.1146/annurev-genet-110711-155459.

Pan G., Cavalli M., Carlsson B., Skrtic S., Kumar C., Wadelius C. rs953413 regulates polyunsaturated fatty acid metabolism by modulating *ELOVL2* expression. *iScience*. 2020;23(2):100808. DOI 10.1016/j.isci.2019.100808.

Panchatcharam M., Salous A.K., Brandon J., Miriyala S., Wheeler J., Patil P., Sunkara M., Morris A.J., Escalante-Alcalde D., Smyth S.S. Mice with targeted inactivation of *Ppap2b* in endothelial and hematopoietic cells display enhanced vascular inflammation and permeability. *Arterioscler. Thromb. Vasc. Biol.* 2014;34(4):837-845. DOI 10.1161/ATVBAHA.113.302335.

Peng T., Zhong L., Gao J., Wan Z., Fu W.-P., Sun C. Identification of rs11615992 as a novel regulatory SNP for human *P2RX7* by allele-specific expression. *Mol. Genet. Genomics*. 2020;295(1):23-30. DOI 10.1007/s00438-019-01598-0.

Piedrafita F.J., Molander R.B., Vansant G., Orlova E.A., Pfahl M., Reynolds W.F. An Alu element in the myeloperoxidase promoter contains a composite SP1-thyroid hormone-retinoic acid response element. *J. Biol. Chem.* 1996;271(24):14412-14420. DOI 10.1074/jbc.271.24.14412.

Prestel M., Prell-Schicker C., Webb T., Malik R., Lindner B., Ziesch N., Rex-Haffner M., Röh S., Viturawong T., Lehm M., Mokry M., den Ruijter H., Haitjema S., Asare Y., Söllner F., Najafabadi M.G., Aherrahrou R., Civelek M., Samani N.J., Mann M., Haffner C., Dichgans M. The atherosclerosis risk variant rs2107595 mediates allele-specific transcriptional regulation of *HDAC9* via E2F3 and Rb1. *Stroke*. 2019;50(10):2651-2660. DOI 10.1161/STROKEAHA.119.026112.

Protze J., Naas S., Krüger R., Stöhr C., Kraus A., Grampp S., Wiesener M., Schiffer M., Hartmann A., Wullich B., Schödel J. The renal cancer risk allele at 14q24.2 activates a novel hypoxia-inducible transcription factor-binding enhancer of DPF3 expression. *J. Biol. Chem.* 2022;298(3):101699. DOI 10.1016/j.jbc.2022.101699.

Ramaker R.C., Bowling K.M., Lasseigne B.N., Hagenauer M.H., Hardigan A.A., Davis N.S., Gertz J., Cartagena P.M., Walsh D.M., Vawter M.P., Jones E.G., Schatzberg A.F., Barchas J.D., Watson S.J., Bunney B.G., Akil H., Bunney W.E., Li J.Z., Cooper S.J., Myers R.M. Post-mortem molecular profiling of three psychiatric disorders. *Genome Med.* 2017;9(1):72. DOI 10.1186/s13073-017-0458-5.

Ratnapriya R., Sosina O.A., Starostik M.R., Kwicklis M., Kapphahn R.J., Fritsche L.G., Walton A., Arvanitis M., Gieser L., Pietraszkiewicz A., Montezuma S.R., Chew E.Y., Battle A., Abecasis G.R., Ferrington D.A., Chatterjee N., Swaroop A. Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. *Nat. Genet.* 2019;51(4):606-610. DOI 10.1038/s41588-019-0351-9.

Reyes-Palomares A., Gu M., Grubert F., Berest I., Sa S., Kasowski M., Arnold C., Shuai M., Srivas R., Miao S., Li D., Snyder M.P., Rabinovitch M., Zaugg J.B. Remodeling of active endothelial enhancers is associated with aberrant gene-regulatory networks in pulmonary arterial hypertension. *Nat. Commun.* 2020;11(1):1673. DOI 10.1038/s41467-020-15463-x.

Ring H.Z., Kroetz D.L. Candidate gene approach for pharmacogenetic studies. *Pharmacogenomics*. 2002;3(1):47-56. DOI 10.1517/14622416.3.1.47.

Roca-Ayats N., Martínez-Gil N., Cozar M., Gerousi M., Garcia-Giralt N., Ovejero D., Mellibovsky L., Nogués X., Díez-Pérez A., Grinberg D., Balcells S. Functional characterization of the *C7ORF76* genomic region, a prominent GWAS signal for osteoporosis in 7q21.3. *Bone*. 2019;123:39-47. DOI 10.1016/j.bone.2019.03.014.

Scelo G., Purdue M.P., Brown K.M., Johansson M., Wang Z., Eckel-Passow J.E., Ye Y., Hofmann J.N., Choi J., Foll M., … Deleuze J.-F., McKay J.D., Parker A.S., Wu X., Houlston R.S., Brennan P., Chanock S.J. Genome-wide association study identifies multiple risk loci for renal cell carcinoma. *Nat. Commun.* 2017;8(1):15724. DOI 10.1038/ncomms15724.

Schaid D.J., Chen W., Larson N.B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* 2018;19(8):491-504. DOI 10.1038/s41576-018-0016-z.

Seshagiri S., Stawiski E.W., Durinck S., Modrusan Z., Storm E.E., Conboy C.B., Chaudhuri S., Guan Y., Janakiraman V., Jaiswal B.S., Guillory J., Ha C., Dijkgraaf G.J.P., Stinson J., Gnad F., Huntley M.A., Degenhardt J.D., Haverty P.M., Bourgon R., Wang W., Koeppen H., Gentleman R., Starr T.K., Zhang Z., Largaespada D.A., Wu T.D., de Sauvage F.J. Recurrent R-spondin fusions in colon cancer. *Nature*. 2012;488(7413):660-664. DOI 10.1038/nature11282.

Sherry S.T., Ward M.H., Kholodov M., Baker J., Phan L., Smigielski E.M., Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29(1):308-311. DOI 10.1093/nar/29.1.308.

Stolze L.K., Conklin A.C., Whalen M.B., López Rodríguez M., Õunap K., Selvarajan I., Toropainen A., Örd T., Li J., Eshghi A., Solomon A.E., Fang Y., Kaikkonen M.U., Romanoski C.E. Systems genetics in human endothelial cells identifies non-coding variants modifying enhancers, expression, and complex disease traits. *Am. J. Hum. Genet.* 2020;106(6):748-763. DOI 10.1016/j.ajhg.2020.04.008.

Sun J., Zhao Y., McGreal R., Cohen-Tayar Y., Rockowitz S., Wilczek C., Ashery-Padan R., Shechter D., Zheng D., Cvekl A. Pax6 associates with H3K4-specific histone methyltransferases Mll1, Mll2, and Set1a and regulates H3K4 methylation at promoters and enhancers. *Epigenetics Chromatin*. 2016;9(1):37. DOI 10.1186/s13072-016-0087-z.

Sun W., Yao S., Tang J., Liu S., Chen J., Deng D., Zeng C. Integrative analysis of super enhancer SNPs for type 2 diabetes. *PLoS One*. 2018;13(1):e0192105. DOI 10.1371/journal.pone.0192105.

Syddall C.M., Reynard L.N., Young D.A., Loughlin J. The identification of *trans*-acting factors that regulate the expression of *GDF5* via the osteoarthritis susceptibility SNP rs143383. *PLoS Genet.* 2013;9(6):e1003557. DOI 10.1371/journal.pgen.1003557.

Tak Y.G., Farnham P.J. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin*. 2015;8:57. DOI 10.1186/s13072-015-0050-4.

Tam V., Patel N., Turcotte M., Bossé Y., Paré G., Meyre D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 2019;20(8):467-484. DOI 10.1038/s41576-019-0127-1.

Tewhey R., Kotliar D., Park D.S., Liu B., Winnicki S., Reilly S.K., Andersen K.G., Mikkelsen T.S., Lander E.S., Schaffner S.F., Sabeti P.C. Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay. *Cell*. 2016;165(6):1519-1529. DOI 10.1016/j.cell.2016.04.027.

Thynn H.N., Chen X.-F., Hu W.-X., Duan Y.-Y., Zhu D.-L., Chen H., Wang N.-N., Chen H.-H., Rong Y., Lu B.-J., Yang M., Jiang F., Dong S.-S., Guo Y., Yang T.-L. An allele-specific functional SNP associated with two systemic autoimmune diseases modulates *IRF5* expression by long-range chromatin loop formation. *J. Invest. Dermatol.* 2020;140(2):348-360.e11. DOI 10.1016/j.jid.2019.06.147.

Tian J., Lou J., Cai Y., Rao M., Lu Z., Zhu Y., Zou D., Peng X., Wang H., Zhang M., Niu S., Li Y., Zhong R., Chang J., Miao X. Risk SNP-mediated enhancer–promoter interaction drives colorectal cancer through both *FADS2* and *AP002754.2*. *Cancer Res.* 2020;80(9):1804-1818. DOI 10.1158/0008-5472.CAN-19-2389.

Tobias I.C., Abatti L.E., Moorthy S.D., Mullany S., Taylor T., Khader N., Filice M.A., Mitchell J.A. Transcriptional enhancers: from prediction to functional assessment on a genome-wide scale. *Genome*. 2021;64(4):426-448. DOI 10.1139/gen-2020-0104.

Ulirsch J.C., Nandakumar S.K., Wang L., Giani F.C., Zhang X., Rogov P., Melnikov A., McDonel P., Do R., Mikkelsen T.S., Sankaran V.G. Systematic functional dissection of common genetic variation affecting red blood cell traits. *Cell*. 2016;165(6):1530-1545. DOI 10.1016/j.cell.2016.04.048.

Umans B.D., Battle A., Gilad Y. Where are the disease-associated eQTLs? *Trends Genet.* 2021;37(2):109-124. DOI 10.1016/j.tig.2020.08.009.

van Arensbergen J., FitzPatrick V.D., de Haas M., Pagie L., Sluimer J., Bussemaker H.J., van Steensel B. Genome-wide mapping of autonomous promoter activity in human cells. *Nat. Biotechnol.* 2017;35(2):145-153. DOI 10.1038/nbt.3754.

van Arensbergen J., Pagie L., FitzPatrick V.D., de Haas M., Baltissen M.P., Comoglio F., van der Weide R.H., Teunissen H., Võsa U., Franke L., de Wit E., Vermeulen M., Bussemaker H.J., van Steensel B. High-throughput identification of human SNPs affecting regulatory element activity. *Nat. Genet.* 2019;51(7):1160-1169. DOI 10.1038/s41588-019-0455-2.

Vasiliev G.V., Merkulov V.M., Kobzev V.F., Merkulova T.I., Ponomarenko M.P., Kolchanov N.A. Point mutations within 663–666 bp of intron 6 of the human *TDO2* gene, associated with a number of psychiatric disorders, damage the YY-1 transcription factor binding site. *FEBS Lett.* 1999;462(1-2):85-88. DOI 10.1016/S0014-5793(99)01513-6.

Visscher P.M., Brown M.A., McCarthy M.I., Yang J. Five years of GWAS discovery. *Am. J. Hum. Genet.* 2012;90(1):7-24. DOI 10.1016/j.ajhg.2011.11.029.

Vockley C.M., Guo C., Majoros W.H., Nodzenski M., Scholtens D.M., Hayes M.G., Lowe W.L., Reddy T.E. Massively parallel quantification of the regulatory effects of noncoding genetic variation in a human cohort. *Genome Res.* 2015;25(8):1206-1214. DOI 10.1101/gr.190090.115.

Wang X., Hayes J.E., Xu X., Gao X., Mehta D., Lilja H.G., Klein R.J. Validation of prostate cancer risk variants rs10993994 and rs7098889 by CRISPR/Cas9 mediated genome editing. *Gene*. 2021;768:145265. DOI 10.1016/j.gene.2020.145265.

Wang Y., Ma R., Liu B., Kong J., Lin H., Yu X., Wang R., Li L., Gao M., Zhou B., Mohan M., Yu H., Hou Z., Shen H., Qian B. SNP rs17079281 decreases lung cancer risk through creating an YY1-binding site to suppress DCBLD1 expression. *Oncogene*. 2020;39(20):4092-4102. DOI 10.1038/s41388-020-1278-4.

Werling D.M., Pochareddy S., Choi J., An J.-Y., Sheppard B., Peng M., Li Z., Dastmalchi C., Santpere G., Sousa A.M.M., Tebbenkamp A.T.N., Kaur N., Gulden F.O., Breen M.S., Liang L., Gilson M.C., Zhao X., Dong S., Klei L., Cicek A.E., Buxbaum J.D., Adle-Biassette H., Thomas J.-L., Aldinger K.A., O'Day D.R., Glass I.A., Zaitlen N.A., Talkowski M.E., Roeder K., State M.W., Devlin B., Sanders S.J., Sestan N. Whole-genome and RNA sequencing reveal variation and transcriptomic coordination in the developing human prefrontal cortex. *Cell Rep.* 2020;31(1):107489. DOI 10.1016/j.celrep.2020.03.053.

Westra H.-J., Franke L. From genome to function by studying eQTLs. *Biochim. Biophys. Acta.* 2014;1842(10):1896-1902. DOI 10.1016/j.bbadis.2014.04.024.

Westra H.-J., Peters M.J., Esko T., Yaghootkar H., Schurmann C., Kettunen J., Christiansen M.W., Fairfax B.P., Schramm K., Powell J.E., … Psaty B.M., Ripatti S., Teumer A., Frayling T.M., Metspalu A., van Meurs J.B.J., Franke L. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* 2013;45(10):1238-1243. DOI 10.1038/ng.2756.

Whitington T., Gao P., Song W., Ross-Adams H., Lamb A.D., Yang Y., Svezia I., Klevebring D., Mills I.G., Karlsson R., Halim S., Dun-

ning M.J., Egevad L., Warren A.Y., Neal D.E., Grönberg H., Lindberg J., Wei G.-H., Wiklund F. Gene regulatory mechanisms underpinning prostate cancer susceptibility. *Nat. Genet.* 2016;48(4): 387-397. DOI 10.1038/ng.3523.

Wu C., Huang R.-T., Kuo C.-H., Kumar S., Kim C.W., Lin Y.-C., Chen Y.-J., Birukova A., Birukov K.G., Dulin N.O., Civelek M., Lusis A.J., Loyer X., Tedgui A., Dai G., Jo H., Fang Y. Mechanosensitive PPAP2B regulates endothelial responses to atherorelevant hemodynamic forces. *Circ. Res.* 2015;117(4):e41-e53. DOI 10.1161/CIRCRESAHA.117.306457.

Xu S., Feng W., Lu Z., Yu C.Y., Shao W., Nakshatri H., Reiter J.L., Gao H., Chu X., Wang Y., Liu Y. regSNPs-ASB: a computational framework for identifying allele-specific transcription factor binding from ATAC-seq data. *Front. Bioeng. Biotechnol.* 2020;8:886. DOI 10.3389/fbioe.2020.00886.

Yan J., Qiu Y., Ribeiro dos Santos A.M., Yin Y., Li Y.E., Vinckier N., Nariai N., Benaglio P., Raman A., Li X., Fan S., Chiou J., Chen F., Frazer K.A., Gaulton K.J., Sander M., Taipale J., Ren B. Systematic analysis of binding of transcription factors to noncoding variants. *Nature*. 2021;591(7848):147-151. DOI 10.1038/s41586-021-03211-0.

Yi M., Tan Y., Wang L., Cai J., Li Xiaoling, Zeng Z., Xiong W., Li G., Li Xiayu, Tan P., Xiang B. TP63 links chromatin remodeling and enhancer reprogramming to epidermal differentiation and squamous cell carcinoma development. *Cell. Mol. Life Sci.* 2020;77(21):4325-4346. DOI 10.1007/s00018-020-03539-2.

Zhang P., Xia J.-H., Zhu J., Gao P., Tian Y.-J., Du M., Guo Y.-C., Suleman S., Zhang Q., Kohli M., Tillmans L.S., Thibodeau S.N., French A.J., Cerhan J.R., Wang L.-D., Wei G.-H., Wang L. High-throughput screening of prostate cancer risk loci by single nucleotide polymorphisms sequencing. *Nat. Commun.* 2018;9(1):2022. DOI 10.1038/s41467-018-04451-x.

Zhao Y., Wu D., Jiang D., Zhang X., Wu T., Cui J., Qian M., Zhao J., Oesterreich S., Sun W., Finkel T., Li G. A sequential methodology for the rapid identification and characterization of breast cancer-associated functional SNPs. *Nat. Commun.* 2020;11(1):3340. DOI 10.1038/s41467-020-17159-8.

Zheng R., Wan C., Mei S., Qin Q., Wu Q., Sun H., Chen C.-H., Brown M., Zhang X., Meyer C.A., Liu X.S. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res.* 2019;47(D1):D729-D735. DOI 10.1093/nar/gky1094.

Zou J., Hormozdiari F., Jew B., Castel S.E., Lappalainen T., Ernst J., Sul J.H., Eskin E. Leveraging allelic imbalance to refine fine-mapping for eQTL studies. *PLoS Genet.* 2019;15(12):e1008481. DOI 10.1371/journal.pgen.1008481.

ORCID ID
E.V. Antontseva orcid.org/0000-0002-4214-7153
A.O. Degtyareva orcid.org/0000-0001-8586-2256
E.E. Korbolina orcid.org/0000-0002-7460-5892
I.S. Damarov orcid.org/0000-0002-3883-3054
T.I. Merkulova orcid.org/0000-0002-2707-0127

Acknowledgements. This work was supported by the Russian Science Foundation, grant No. 23-15-00113.
Conflict of interest. The authors declare no conflict of interest.
Received January 17, 2023. Revised March 24, 2023. Accepted March 30, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):676-683
DOI 10.18699/VJGB-23-78

Original Russian text https://vavilovj-icg.ru/

Molecular basis and genetics of hypohidrotic ectodermal dysplasias

V.A. Kovalskaia ✉, T.B. Cherevatova, A.V. Polyakov, O.P. Ryzhkova

Research Centre for Medical Genetics, Moscow, Russia
✉ kovalskaya@med-gen.ru

Abstract. Ectodermal dysplasia (ED) is a heterogeneous group of hereditary diseases of the skin and its appendages, which are characterized by impaired development and/or homeostasis of two or more ectoderm derivatives, including: hair, teeth, nails, sweat glands and their modifications (mammary glands, for instance). The overall prevalence of ectodermal dysplasia remains precisely unknown not only in Russia, but also in the world, nor is known the contribution of individual genes to its structure. This complicates the DNA diagnosis establishment of this disease due to the lack of an accurate diagnostic algorithm and a universal cost-effective method of analysis. To date, the most highly-researched genes involved in the development of anhydrous or hypohidrotic forms of ED are *EDA*, *EDAR*, *EDARADD* and *WNT10A*. The ectodysplasin A (*EDA*) gene is the cause of the most common X-linked form of ED, a gene from the Wnt family (*WNT10A*) is responsible for the autosomal recessive form of the disease, and two other genes (*EDAR* and *EDARADD*) can cause both autosomal recessive and autosomal dominant forms. This review provides the characteristics of the genes involved in ED, their mutation spectra, the level of their expression in human tissues, as well as the interrelation of the aforementioned genes. The domain structures of the corresponding proteins are considered, as well as the molecular genetic pathways in which they are involved. Animal models for studying this disorder are also taken into consideration. Due to the cross-species genes conservation, their mutations cause the disruption of the development of ectoderm derivatives not only in humans, but also in mice, cows, dogs, and even fish. It can be exploited for a better understanding of the etiopathogenesis of ectodermal dysplasias. Moreover, this article brings up the possibility of recurrent mutations in the *EDA* and *WNT10A* genes. The review also presents data on promising approaches for intrauterine ED treatment.
Key words: ectodermal dysplasia; *EDA*; tooth agenesis; Wnt family.

For citation: Kovalskaia V.A., Cherevatova T.B., Polyakov A.V., Ryzhkova O.P. Molecular basis and genetics of hypohidrotic ectodermal dysplasias. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023; 27(6):676-683. DOI 10.18699/VJGB-23-78

Молекулярно-генетическая характеристика гипогидротических эктодермальных дисплазий

В.А. Ковальская ✉, Т.Б. Череватова, А.В. Поляков, О.П. Рыжкова

Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия
✉ kovalskaya@med-gen.ru

Аннотация. Эктодермальные дисплазии – гетерогенная группа наследственных заболеваний кожи и ее придатков, которые характеризуются нарушением развития и/или гомеостаза двух и более производных эктодермы, включая: волосы, зубы, ногти, потовые железы и их модификации (например, молочные железы). Общая распространенность эктодермальных дисплазий остается точно неизвестной не только в России, но и в мире, так же как и вклад отдельных генов в ее структуру. Это затрудняет ДНК-диагностику данного заболевания ввиду отсутствия строгого алгоритма диагностики и универсального, экономически выгодного метода анализа. На сегодняшний день наиболее изученными генами, вовлеченными в развитие ангидротической или гипогидротической форм эктодермальной дисплазии являются *EDA*, *EDAR*, *EDARADD* и *WNT10A*. Ген эктодисплазина А (*EDA*) служит причиной самой частой Х-сцепленной формы эктодермальной дисплазии, ген из семейства Wnt (*WNT10A*) отвечает за аутосомно-рецессивную форму заболевания, а два других гена (*EDAR* и *EDARADD*) могут быть причиной как аутосомно-рецессивных, так и аутосомно-доминантных форм. В настоящем литературном обзоре приведены характеристика генов, вовлеченных в эктодермальную дисплазию, спектры их мутаций, уровень их экспрессии в тканях человека, а также взаимосвязь вышеупомянутых генов друг с другом. Обсуждается также доменная структура соответствующих белков, рассмотрены молекулярно-генетические пути, в которые они преимущественно вовлечены, и описаны животные модели для изучения данной патологии. Ввиду межвидовой консервативности упомянутых генов, мутации в них вызывают

нарушения развития производных эктодермы не только у человека, но и у мышей, коров, собак и даже рыб, что может быть использовано для лучшего понимания этиопатогенеза эктодермальных дисплазий. Более того, в статье поднимаются вопросы о возможных частых мутациях в генах *EDA* и *WNT10A*. Приведены также данные касаемо разрабатываемых перспективных подходов к внутриутробному лечению эктодермальной дисплазии.
Ключевые слова: эктодермальная дисплазия; *EDA*; агенезия зубов; семейство Wnt.

Introduction

Ectodermal dysplasia (ED) refers to a diverse set of molecular genetic disorders that share a common feature of developmental abnormalities or imbalances affecting two or more ectodermal structures (Wright et al., 2019). Despite the fact that the ectoderm determines the development of many organs and tissues, such as: the central and peripheral nervous system, pituitary gland, olfactory neuroepithelium, melanocytes, tooth enamel, epidermis, including sweat glands, hair, nails, ED primarily affects the latter group of ectodermal derivatives.

The exact prevalence of ED is uncertain due to limited research and differing classification criteria across countries. Nonetheless, some estimate that it may affect as many as 70 out of 100,000 newborns (Itin, Fistarol, 2004). Reported data for the Danish population, collected from 1995 to 2010, indicate that the prevalence of ectodermal dysplasia corresponds to 21.9 per 100,000, and molecularly confirmed X-linked ectodermal dysplasia – 1.6 per 100,000 (Nguyen-Nielsen et al., 2013). Thus, ectodermal dysplasias, although they are not among the most common hereditary diseases, have a widespread presence and make a significant contribution to the structure of dental, dermatological, and genetic pathologies.

This group of pathological conditions may have been known since the end of the 18th century, however, the first documented case of ectodermal dysplasia dates back to 1838, when Wedderburn, in a letter to Charles Darwin, described 10 men from an Indian family suffering from partial absence of teeth, baldness and excessive dry skin (Felsher, 1944). In 1848, two additional patients were reported by Thurman, and another case was documented by Guilford in 1883. However, only in 1929 the term "hereditary ectodermal dysplasia" was introduced by Weech, who also introduced the term "anhidrotic" to describe individuals with ectodermal dysplasia who exhibit a diminished ability to sweat. This term was subsequently replaced by "hypohidrotic" (Weech, 1929). Following an analysis of 19 affected families in 1937, Siemens concluded that the genetic etiology of ectodermal dysplasia could not be explained by a single gene and one mode of inheritance. Thus, it was realized that both dominant and recessive forms of the disease were present, along with sex-linked forms that exhibited phenotypic overlap but did not entirely replicate it (Siemens, 1937). In 1939, Clouston noticed that despite similar clinical data patients with manifestations of ectodermal dysplasia may differ significantly from each other in the extent of sweat gland development. As a result, he classified them into two broad categories: hypohidrotic type, which was limited to only 4 cases, and hidrotic ectodermal dysplasia, which encompassed over 50 patients (Clouston, 1939).

Further study of this nosological unit resulted in the development of the initial clinical classification by Freire-Maia and Pinheiro (Freire-Maia, 1971; Freire-Maia, Pinheiro, 1988), which has been widely employed in routine medical practice. Their classification was based on the principle of involvement of certain ectodermal structures in the pathological process. They assigned to group A all conditions in which at least two classical derivatives of the ectoderm were affected, such as: hair, teeth, nails and sweat glands. Diseases assigned to group B included deviations in only one of the four structures mentioned above and one additional ectodermal defect, such as abnormalities of the ears, lips, or palmar and plantar hyperkeratosis. The condition, which was characterized by the presence of only ectodermal signs, they called pure ectodermal dysplasia. The combination of ectodermal signs with other anomalies was called ectodermal dysplasia syndrome by the authors. In addition, all the classical structures of the ectoderm were numbered, where 1 – hair, 2 – teeth, 3 – nails, 4 – sweat glands, in order to distinguish further the main groups of ED: ED1 – trichodysplasia, ED2 – dental dysplasia, ED3 – onychodysplasia, ED4 – dyshidrosis (Deshmukh, Prashanth, 2012). It is worth mentioning that the classification proposed by Freire-Maia and Pinheiro did not consider the molecular and genetic aspects of ectodermal dysplasia and necessitated revision with the emergence of next-generation sequencing techniques and advancements in genomic medicine.

At the end of 2019, a group of international experts associated with the National Foundation for Ectodermal Dysplasias (NFED) published a revised classification of ectodermal dysplasia (ED) in the American Journal of Medical Genetics. This updated classification system is based on the molecular pathways involved in the development of ED and provides a more precise list of pathologies than previous classifications. Specifically, the new classification identifies 102 syndromes that fall under the definition of ectodermal dysplasia, reflecting a more comprehensive understanding of this complex condition (Wright et al., 2019). In addition to non-syndromic ectodermal dysplasia, it included such heterogeneous syndromes as: Coffin–Siris, Dubowitz, Hallermann–Streiff, Gorlin–Goltz, Johanson–Blizzard and others, which does not meet the criteria set by domestic terminology. In the Russian Federation, the term "ectodermal dysplasia" typically refers only to non-syndromic forms of the condition, specifically anhidrotic (hypohidrotic) and hidrotic forms (Kozlova, Demikova, 2007). However, it is worth noting that these isolated forms are fully consistent with the molecular etiology proposed by experts from NFED.

The current understanding of ED suggests the involvement of four major signaling pathways: EDA-mediated, WNT-, NF-κB-, and TP63-mediated pathways. Of these, only the first three have been implicated in the development of anhidrotic forms of ectodermal dysplasia (Fig. 1) (Mikkola, 2009; Sadier et al., 2015; Wright et al., 2019).

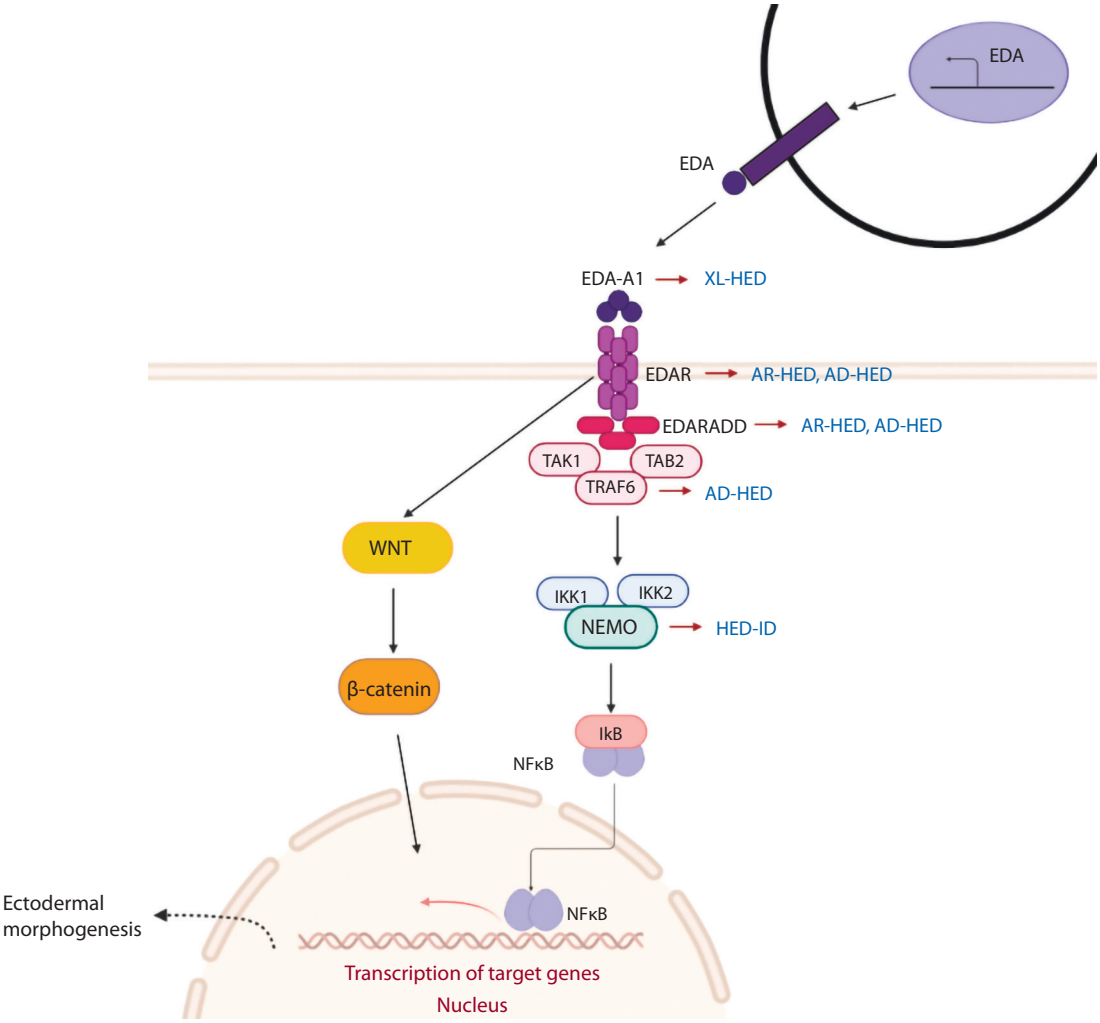

Fig. 1. The major proteins involved in the development of ectodermal structures.

EDA is expressed on the cell surface, but its extracellular domain can be proteolytically cleaved to form a soluble signaling molecule that binds to the ectodysplasin receptor (EDAR). EDAR interacts with the EDARADD protein, and further downstream signaling via activation of the NFκB pathway leads to the expression of genes specific to the epidermis, hair, teeth, and nails. XL-HED – X-linked hypohidrotic ectodermal dysplasia, AD-HED – autosomal dominant hypohidrotic ectodermal dysplasia, AR-HED – autosomal recessive hypohidrotic ectodermal dysplasia, HED-ID – hypohidrotic ectodermal dysplasia with immunodeficiency.

Major genes involved in the development of ectodermal dysplasia

The human *EDA* gene (also known as *ED1*, *HED*, *EDA1*, *EDA2*, *HED1*, *ODT1*, *XHED*, *ECTD1*) is a protein-coding gene responsible for the synthesis of ectodysplasin A, a type II transmembrane protein belonging to the tumor necrosis factor (TNF) family, which is involved in the transmission of epithelial-mesenchymal signals during the morphogenesis of ectodermal structures in humans (Bayés et al., 1998; Mikkola, Thesleff, 2003).

The *EDA* (ectodysplasin A) gene is mapped on the X chromosome, at the Xq13.1 locus, according to the main transcript (NM_001399.5), it contains 8 exons, with start and stop codons in the first and last exons, respectively. A total of 8 protein-coding isoforms have been described, differing in length and function, but isoform 1 (EDA-A1), consisting of 391 amino acids, is the main one and is the ligand for the EDAR receptor. Another isoform, known as EDA-2, is distinguished by the absence of Val307 and Glu308 in the TNF domain and binds only to EDA2R, ensuring the subsequent correct postembryonic functioning of various structures and tissues (Kere et al., 1996). Both isoforms, EDA1 and EDA2, through the EDAR and EDA2R receptors activate the NFκB signaling pathway, but only the EDA1/EDAR interaction is important in the development of ectoderm derivatives and the disease (Newton et al., 2004). The precise reason why the disruption of EDA-A2/XEDAR interaction does not lead to the ectodermal dysplasia phenotype is currently unknown and requires further investigation, however, studies have shown that EDA-A2 is expressed primarily in aging adipose tissue, arteries, heart, lungs, muscle, and skin, and may also regulate glucose metabolism, and serves as a predictor of steatosis aggravation in patients with non-alcoholic fatty liver disease (Yang et al., 2015; Cai et al., 2021).

Besides the C-terminal TNF domain (249–383 aa), ectodysplasin A contains a collagen domain (180–229 aa), a furin cleavage site (153–160 aa), and a transmembrane N-terminal domain (42–62 aa) (Chen et al., 2001; www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi; www.uniprot.org/uniprot/Q92838) (Fig. 2). Ectodysplasin A, as a member of the TNF-ligand family, can function locally through direct intercellular contacts as a complete membrane form. However, it predominantly acts in its secreted form, which is generated through proteolytic cleavage at a furin consensus site that releases the C-terminal part of the protein as a soluble trimeric ligand. This ligand then initiates downstream signaling by activating various proteins (Elomaa, 2001) (see Fig. 1).

EDAR is another key protein in this molecular genetic pathway, encoded by the gene of the same name at the chr2q12.3 locus. As for the topology, the ectodysplasin A receptor has an extracellular part, including a ligand-binding domain (LBD) (13–148 aa), encoded by exons 2–5 and a cytoplasmic part, represented by a death domain, encoded by exon 12 (354–428 aa) for interaction with the β-isoform of EDARADD (Sadier et al., 2015; Zhang et al., 2020; www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). The latter, in turn, through the EDAR-associated death domain (124–189 aa) and the TRAF6-binding site (27–31 aa) leads to subsequent activation of the NF-κB pathway (Morlon et al., 2005; Asano et al., 2021) (see Fig. 1).

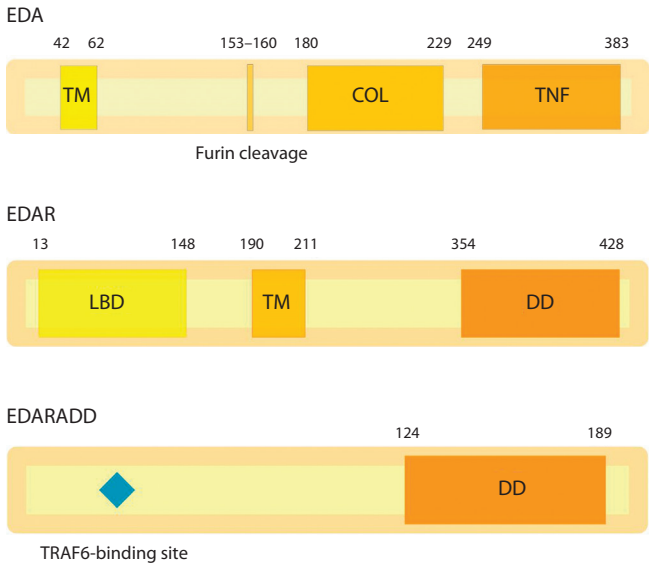

Fig. 2. Domain structure of the main proteins involved in the EDA-mediated pathway.

TM – transmembrane domain; Furin cleavage – Furin cleavage site; COL – collagen domain; TNF – tumor necrosis factor domain; LBD – ligand-binding domain; DD – death domain.

The high degree of similarity between the human and mouse genes facilitated the generation of several mutant mouse lines in the 1990s (Headon, Overbeek, 1999; Trzeciak, Koczorowski, 2015). It was observed that subjects lacking *EDA2R*(*XEDAR*) did not exhibit any symptoms of ectodermal dysplasia (Newton et al., 2004). However, individuals harboring recessive (*downless*) and dominant (*Sleek*) mutations in the *EDAR* gene demonstrated significant disruption in the development of ectodermal structures, such as sparse hair, absence of cover behind the ear, and presence of abnormal teeth, particularly incisors (Crocker, Cattanach, 1979). In the phenotype of *Tabby* mutant (analog of human *EDA*) male mice, there were found alopecia areata behind the ears, tail alopecia, an absence of some vibrissae, abnormal fur texture due to the absence of zigzag-shaped and protective hairs, and an absence of sweat glands normally found on the paw pads (Ferguson et al., 1997; Srivastava et al., 1997). In *crinkled* and *swh/swh* mice, due to homozygous variants in the *EDARADD* gene, a similar ectodermal dysplasia phenotype was observed, and females were unable to feed offspring due to underdevelopment of the mammary glands (Yan et al., 2002; Kuramoto et al., 2005, 2011).

The highly conserved nature of the EDA-mediated pathway has enabled the identification of a common phenotype in other vertebrates, as well as distinct mutations in genes of interest (Pantalacci et al., 2008). Notably, a long deletion spanning exon 3 of the *EDA* gene was identified in four male calves, resulting in a significant reduction in hair density on the head, auricles, neck, back, and tail. These areas of the body exhibited sparse hair growth, while teeth abnormalities such as partial adentia and conical teeth were also observed (Drögemüller et al., 2001, 2002). In male dogs with a hemizygous mutation in the splicing acceptor site of exon 8 of the *EDA* gene, all sweat glands were absent, they were completely devoid of hair in the frontal part and in the pelvic region on the back. Most premolars and some incisors were missing, and the present teeth were mostly conical. Moreover, affected dogs showed increased morbidity and mortality from pulmonary infectious diseases compared to other dogs in the same environment (Casal et al., 2005). Zebrafish and medaka mutants with disturbances in EDA signaling (*EDA* and *EDAR* genes) were observed to have lost fins and scales, lacked teeth, or had abnormally-shaped teeth (Harris et al., 2008; Atukorala et al., 2010). The situation was similar with marine and freshwater sticklebacks: marine representatives of *Gasterosteus aculeatus*, in which the expression of the *EDA* gene is much higher, demonstrated a more developed cover with 32 lateral plates, while freshwater individuals were limited to 0–9 lateral plates (O'Brown et al., 2015). Based on these data, the EDA pathway probably controls the development of ectoderm derivatives in all vertebrates (Sadier et al., 2014).

EDA is mostly expressed in endocrine organs (adrenals, thyroid, ovaries), various parts of the brain and heart, the lowest level of expression is observed in blood cells (TPM 0.18). For cultured fibroblasts, this indicator is 0.89, which makes them the most accessible object for studying *EDA*-transcripts (www.gtexportal.org). Expression of *EDAR* and *EDARADD* predominantly occurs in the bladder, esophageal mucosa, and skin. However, it is also more efficient to study *EDARADD* expression patterns on a culture of fibroblasts, while studying the structure of mRNA and splicing disorders of *EDAR*-transcripts is possible mainly only when using blood leukocytes (www.gtexportal.org).

According to the HGMD database, mutations in *EDA*, *EDAR* and *EDARADD* are relatively evenly distributed throughout the genes and affect all significant domains (www.hgmd.cf.ac.uk). To date, 371 pathogenic variants have been described in the *EDA* gene, 83 in *EDAR,* and 19 in *EDARADD*, including missense and nonsense mutations that make up the major part, deletions and insertions, including gross ones, as well as splicing variants, affecting both canonical splicing sites and leading to activation of cryptic ones (www.hgmd.cf.ac.uk).

To date, there is no record of classical recurrent mutations in the genes related to the EDA-mediated pathway in any population. However, some researchers have reported that variants in *EDA* affecting amino acids R155 and R156 at overlapping furin cleavage sites may amount to from 7 to 30 % (Vincent et al., 2001; Chaudhary et al., 2022). This phenomenon can, in particular, be explained by the presence of a CpG rich region in exon 3 in arginine codons 155, 156, which, when methylated, causes the so-called C-T transition (Chen et al., 2001).

Another notable observation is that the majority of mutations occurring in the *EDA* gene lead to X-linked hypohidrotic ectodermal dysplasia (OMIM 305100), characterized by classic symptoms such as scalp hypotrichosis, nail dystrophy, oligodontia with conical incisors, and hypohidrosis (www.omim.org). While male patients exhibit a more severe phenotype, clinical manifestations can also be observed in females, even in the absence of an unequal pattern of X-chromosome inactivation (Vincent et al., 2001). Up to 70 % of heterozygous female carriers of pathogenic variants in the *EDA* gene demonstrate one or more disorders: some degree of hypotrichosis, reduced sweating, missing one or more teeth, underdevelopment of the mammary glands, or problems with breastfeeding – the latter, however, can only be fully assessed after puberty or pregnancy, respectively (Wahlbuhl-Becker et al., 2017; Wohlfart et al., 2020). Moreover, even within the same family, there is a certain variability in the phenotype (Cañueto et al., 2011; Han et al., 2020). Cases of selective tooth agenesis (OMIM 313500) of the X-linked mode of inheritance, which were also caused by pathogenic variants in the *EDA* gene, are described. Despite the fact that mutations leading to this phenotype have been described in different protein domains of ectodysplasin A, the presence of residual activity of the protein and the possibility of its binding to the *EDAR* receptor is probably the key factor (Mues et al., 2010).

For the *EDAR* and *EDARADD* genes, both autosomal dominant and autosomal recessive forms of anhidrotic ectodermal dysplasia have been described. Some authors believe that dominant mutations are mainly localized in the domains of protein-protein interactions, which leads to disruption of oligomerization and a dominant negative effect (Sadier et al., 2014). However, this assumption is not fully justified. Thus, functional analysis of missense mutations p.D120Y, p.L122R, and p.D123N located near the *EDARADD* death domain proved not only their dominant nature, but also the ability to significantly reduce the interaction with TRAF6 and suppress the subsequent activation of NF-κB. The p.E152K mutation in the heterozygous state, located directly in the *EDAR*-associated death domain, on the contrary, was recessive and showed only a slight decrease in affinity for TRAF6 (Asano et al., 2021).

The V370A variant is a conservative amino acid substitution in the *EDAR* gene identified in Asian and Latin American populations by whole genome sequencing (Park et al., 2012). It is believed to be a gain-of-function mutation that leads to a 2-fold increase in the activation of the NF-κB pathway (Kataoka et al., 2021) and, accordingly, correlates with increased hair thickness and special tooth morphology in representatives of Asia and indigenous peoples of the USA (Bryk et al., 2008). An interesting fact is that this variant was selected, presumably in Central China, about 30,000 years ago, and the presence in the genotype of pathogenic variants in the *EDA* gene in the presence of V370A reduces the severity of clinical manifestations of anhidrotic ectodermal dysplasia (Cluzeau et al., 2011).

NEMO is another protein involved in the pathogenetic cascade. Due to the fact that NF-κB also controls the immune response and apoptosis, the clinical manifestations of mutations in the *NEMO* gene are not only limited to damage to ectodermal structures, but also include immune system disorders, with the development, in particular, of anhidrotic ectodermal dysplasia with immunodeficiency 1 (OMIM 300291) (Smahi et al., 2002).

Mutations in the *WNT10A* gene are the most common cause of non-syndromic selective tooth agenesis (Xu et al., 2017; Yu et al., 2019), but are also associated with the development of hypohidrotic ectodermal dysplasia, odonto-onychodermal dysplasia, and Schöpf–Schulz–Passarge syndrome. The *WNT10A* gene encodes a protein of the same name, a component of the canonical Wnt/β-catenin signaling pathway that plays an important role in several stages of dental morphogenesis, including activation of the mesenchymal odontogenic potential during early tooth development, as well as the induction and maintenance of primary and secondary enamel nodes (Xu et al., 2017). Conversations are currently in progress regarding the contribution of Wnt signaling to the development and formation of hair follicles and skin structures (Adaimy et al., 2007). The human *Wnt*-family includes genes that show significant similarity to mouse wingless genes, and therefore alopecia is consistently observed in *WNT10A*-deficient mice (*WNT10A–/–*), besides growth retardation, kyphosis, and reproductive dysfunction (Wang et al., 2018).

Currently, 94 variants for *WNT10A* are described in the HGMD database, however, only p.Cys107Ter (rs121908119) and p.Phe228Ile (rs121908120) variants have been suggested to be located in hotspots. These mutations were the most common among patients of Polish and Italian origin with *WNT10A*-mediated ectodermal dysplasia (Castori et al., 2010; Mostowska et al., 2012).

Treatment approaches of ectodermal dysplasia

At present, treatment options for ED are primarily aimed at managing the symptoms to prevent complications. A targeted and effective treatment approach is yet to be developed. However, a phase 2 clinical trial (NCT04980638) involving the intra-amniotic administration of ER004 to male fetuses with confirmed X-linked ectodermal dysplasia is currently

underway. ER004 is a novel signaling protein replacement molecule that has been specifically designed to bind with high affinity to the endogenous EDAR receptor. ER004 is believed to function by providing a replacement for the deficient ectodysplasin A protein in patients who have pathogenic variants in the *EDA* gene. This replacement aims to facilitate the normal development of essential ectodermal structures. The proposed administration method is intra-amniotic, with a suggested dose of 100 mg/kg fetal weight per injection. Treatment would consist of a total of three injections administered at intervals of three weeks, beginning at 26 weeks of gestation. In order to evaluate the long-term efficacy and safety of the treatment, individuals will be monitored for a period of 5 years. The end of testing is scheduled for April 2029 (www.clinicaltrials.gov/ct2/show/NCT04980638).

EDI200 is an additional drug currently being developed. It is expected to be administered postnatally, between days 2 and 14 of life, in male patients diagnosed with X-linked ectodermal dysplasia. The treatment would consist of five injections, each containing 3 mg/kg of the human ectodysplasin A molecule. Similar to ER004, EDI200 also targets the activation of the EDA-mediated pathway. *In vivo* experiments conducted on XLHED-affected animals have demonstrated that a course of EDI200 therapy, whether administered prenatally or postnatally, can correct *EDA* deficiency. To evaluate long-term efficacy and safety, individuals treated with EDI200 will be monitored until they reach the age of 10 years (until March of 2025) (www.clinicaltrials.gov/ct2/show/NCT01992289).

Conclusion

In Russia, molecular genetic studies of ectodermal dysplasia have not yet been carried out; the contribution of mutations of various ED genes remains unknown. The study of the full spectrum of mutations in the ED genes will allow developing an algorithm for the molecular genetic diagnosis of ectodermal dysplasia.

References

Adaimy L., Chouery E., Megarbane H., Mroueh S., Delague V., Nicolas E., Belguith H., de Mazancourt P., Megarbane A. Mutation in *WNT10A* is associated with an autosomal recessive ectodermal dysplasia: the odonto-onycho-dermal dysplasia. *Am. J. Hum. Genet*. 2007;81(4):821-828. DOI 10.1086/520064.

Asano N., Yasuno S., Hayashi R., Shimomura Y. Characterization of *EDARADD* gene mutations responsible for hypohidrotic ectodermal dysplasia. *J. Dermatol*. 2021;48(10):1533-1541. DOI 10.1111/1346-8138.16044.

Atukorala A., Inohaya K., Baba O., Tabata M., Ratnayake R., Abduweli D., Kasugai S., Mitani H., Takano Y. Scale and tooth phenotypes in medaka with a mutated ectodysplasin-A receptor: implications for the evolutionary origin of oral and pharyngeal teeth. *Arch. Histol. Cytol.* 2010;73(3):139-148. DOI 10.1679/aohc.73.139.

Bayés M., Hartung A., Ezer S., Pispa J., Thesleff I., Srivastava A., Kere J. The anhidrotic ectodermal dysplasia gene (EDA) undergoes alternative splicing and encodes ectodysplasin-A with deletion mutations in collagenous repeats. *Hum. Mol. Genet.* 1998;7(11):1661-1669. DOI 10.1093/hmg/7.11.1661.

Bryk J., Hardouin E., Pugach I., Hughes D., Strotmann R., Stoneking M., Myles S. Positive selection in East Asians for an *EDAR* allele that enhances NF-κB activation. *PLoS One.* 2008;3(5):e2209. DOI 10.1371/journal.pone.0002209.

Cai Z., Deng X., Jia J., Wang D., Yuan G. Ectodysplasin A/ectodysplasin A receptor system and their roles in multiple diseases. *Front. Physiol*. 2021;12:788411. DOI 10.3389/fphys.2021.788411.

Cañueto J., Zafra-Cobo M., Ciria S., Unamuno P., González-Sarmiento R. A novel *EDA* gene mutation in a Spanish family with X-linked hypohidrotic ectodermal dysplasia. *Actas Dermosifiliogr.* 2011;102(9):722-725. DOI 10.1016/j.adengl.2011.11.012.

Casal M., Scheidt J., Rhodes J., Henthorn P., Werner P. Mutation identification in a canine model of X-linked ectodermal dysplasia. *Mamm. Genome*. 2005;16(7):524-531. DOI 10.1007/s00335-004-2463-4.

Castori M., Castiglia D., Brancati F., Foglio M., Heath S., Floriddia G., Madonna S., Fischer J., Zambruno G. Two families confirm Schöpf–Schulz–Passarge syndrome as a discrete entity within the WNT10A phenotypic spectrum. *Clin. Genet*. 2011;79(1):92-95. DOI 10.1111/j.1399-0004.2010.01513.x.

Chaudhary A.K., Gholse A., Nagarajaram H.A., Dalal A.B., Gupta N., Dutta A.K., Danda S., Gupta R., Sankar H.V., Bhavani G.S., Girisha K.M., Phadke S.R., Ranganath P., Bashyam M.D. Ectodysplasin pathogenic variants affecting the furin-cleavage site and unusual clinical features define X-linked hypohidrotic ectodermal dysplasia in India. *Am. J. Med. Genet. A*. 2022;188(3):788-805. DOI 10.1002/ajmg.a.62579.

Chen Y., Molloy S., Thomas L., Gambee J., Bächinger H., Ferguson B., Zonana J., Thomas G., Morris N. Mutations within a furin consensus sequence block proteolytic release of ectodysplasin-A and cause X-linked hypohidrotic ectodermal dysplasia. *Proc. Natl. Acad. Sci. USA.* 2001;98(13):7218-7223. DOI 10.1073/pnas.131076098.

Clouston H. The major forms of hereditary ectodermal dysplasia. (With an autopsy and biopsies on the anhydrotic type). *Can. Med. Assoc. J.* 1939;40(1):1-7.

Cluzeau C., Hadj-Rabia S., Bal E., Clauss F., Munnich A., Bodemer C., Headon D., Smahi A. The EDAR370A allele attenuates the severity of hypohidrotic ectodermal dysplasia caused by EDA gene mutation. *Br. J. Dermatol*. 2012;166(3):678-681. DOI 10.1111/j.1365-2133.2011.10620.x. Epub 2011 Dec. 5.

Crocker M., Cattanach B.M. The genetics of Sleek: a possible regulatory mutation of the tabby-crinkled-downless syndrome. *Genet. Res.* 1979;34(3):231. DOI 10.1017/s0016672300019479.

Deshmukh S., Prashanth S. Ectodermal dysplasia: a genetic review. *Int. J. Clin. Pediatr. Dent*. 2012;5(3):197-202. DOI 10.5005/jp-journals-10005-1165.

Drögemüller C., Distl O., Leeb T. Partial deletion of the bovine *ED1* gene causes anhidrotic ectodermal dysplasia in cattle. *Genome Res.* 2001;11(10):1699-1705. DOI 10.1101/gr.182501.

Drögemüller C., Kuiper H., Peters M., Guionaud S., Distl O., Leeb T. Congenital hypotrichosis with anodontia in cattle: a genetic, clinical and histological analysis. *Vet. Dermatol.* 2002;13(6):307-313. DOI 10.1046/j.1365-3164.2002.00313.x.

Elomaa O. Ectodysplasin is released by proteolytic shedding and binds to the *EDAR* protein. *Hum. Mol. Genet.* 2001;10(9):953-962. DOI 10.1093/hmg/10.9.953.

Felsher Z. Hereditary ectodermal dysplasia: report of a case, with experimental study. *Arch. Derm. Syphilol*. 1944;49(6):410-414. DOI 10.1001/archderm.1944.01510120024005.

Ferguson B., Brockdorff N., Formstone E., Ngyuen T., Kronmiller J., Zonana J. Cloning of *Tabby*, the murine homolog of the human EDA gene: evidence for a membrane-associated protein with a short collagenous domain. *Hum. Mol. Genet*. 1997;6(9):1589-1594. DOI 10.1093/hmg/6.9.1589.

Freire-Maia N. Ectodermal dysplasias. *Hum. Hered.* 1971;21(4):309-312. DOI 10.1159/000152419.

Freire-Maia N., Pinheiro M. Ectodermal dysplasias – some recollections and a classification. *Birth Defects Orig. Artic. Ser*. 1988;24:3-14.

Han Y., Wang X., Zheng L., Zhu T., Li Y., Hong J., Xu C., Wang P., Gao M. Pathogenic *EDA* mutations in Chinese Han families with hy-

pohidrotic ectodermal dysplasia and genotype-phenotype: a correlation analysis. *Front. Genet.* 2020;11:21. DOI 10.3389/fgene.2020.00021.

Harris M., Rohner N., Schwarz H., Perathoner S., Konstantinidis P., Nüsslein-Volhard C. Zebrafish *eda* and *edar* mutants reveal conserved and ancestral roles of ectodysplasin signaling in vertebrates. *PLoS Genet*. 2008;4(10):e1000206. DOI 10.1371/journal.pgen.1000206.

Headon D., Overbeek P. Involvement of a novel Tnf receptor homologue in hair follicle induction. *Nat. Genet*. 1999;22(4):370-374. DOI 10.1038/11943.

Itin P., Fistarol S. Ectodermal dysplasias. *Am. J. Med. Genet. C Semin. Med. Genet*. 2004;131C(1):45-51. DOI 10.1002/ajmg.c.30033.

Kataoka K., Fujita H., Isa M., Gotoh S., Arasaki A., Ishida H., Kimura R. The human *EDAR* 370V/A polymorphism affects tooth root morphology potentially through the modification of a reaction-diffusion system. *Sci. Rep*. 2021;11(1):5143. DOI 10.1038/s41598-021-84653-4.

Kere J., Srivastava A.K., Montonen O., Zonana J., Thomas N., Ferguson B., Munoz F., Morgan D., Clarke A., Baybayan P., Chen E.Y., Ezer S., Saarialho-Kere U., de la Chapelle A., Schlessinger D. X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. *Nat. Genet.* 1996;13(4):409-416. DOI 10.1038/ng0895-409.

Kozlova S.I., Demikova N.S. Hereditary Disorders and Genetic Counseling: Guide Atlas. Moscow: KMK–Avtorskaya Akademiya Publ., 2007. (in Russian)

Kuramoto T., Morimura K., Nomoto T., Namiki C., Hamada S., Fukushima S., Sugimura T., Serikawa T., Ushijima T. Sparse and wavy hair: a new model for hypoplasia of hair follicle and mammary glands on rat chromosome 17. *J. Hered.* 2005;96(4):339-345. DOI 10.1093/jhered/esi053.

Kuramoto T., Yokoe M., Hashimoto R., Hiai H., Serikawa T. A rat model of hypohidrotic ectodermal dysplasia carries a missense mutation in the *Edaradd* gene. *BMC Genet*. 2011;12(1):91. DOI 10.1186/1471-2156-12-91.

Mikkola M. Molecular aspects of hypohidrotic ectodermal dysplasia. *Am. J. Med. Genet. A*. 2009;149A(9):2031-2036. DOI 10.1002/ajmg.a.32855.

Mikkola M., Thesleff I. Ectodysplasin signaling in development. *Cytokine Growth Factor Rev*. 2003;14(3-4):211-224. DOI 10.1016/s1359-6101(03)00020-0.

Morlon A., Munnich A., Smahi A. TAB2, TRAF6 and TAK1 are involved in NF-κB activation induced by the TNF-receptor, Edar and its adaptator Edaradd. *Hum. Mol. Genet*. 2005;14(23):3751-3757. DOI 10.1093/hmg/ddi405.

Mostowska A., Biedziak B., Zadurska M., Dunin-Wilczynska I., Lianeri M., Jagodzinski P. Nucleotide variants of genes encoding components of the Wnt signalling pathway and the risk of non-syndromic tooth agenesis. *Clin. Genet*. 2012;84(5):429-440. DOI 10.1111/cge.12061.

Mues G., Tardivel A., Willen L., Kapadia H., Seaman R., Frazier-Bowers S., Schneider P., D'Souza R. Functional analysis of Ectodysplasin-A mutations causing selective tooth agenesis. *Eur. J. Hum. Genet*. 2010;18(1):19-25. DOI 10.1038/ejhg.2009.127.

Newton K., French D., Yan M., Frantz G., Dixit V. Myodegeneration in EDA-A2 transgenic mice is prevented by XEDAR deficiency. *Mol. Cell. Biol*. 2004;24(4):1608-1613. DOI 10.1128/MCB.24.4.1608-1613.2004.

Nguyen-Nielsen M., Skovbo S., Svaneby D., Pedersen L., Fryzek J. The prevalence of X-linked hypohidrotic ectodermal dysplasia (XLHED) in Denmark, 1995–2010. *Eur. J. Med. Genet.* 2013;56(5):236-242. DOI 10.1016/j.ejmg.2013.01.012.

O'Brown N., Summers B., Jones F., Brady S., Kingsley D. A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene *EDA*. *eLife.* 2015;4:e05290. DOI 10.7554/eLife.05290.

Pantalacci S., Chaumot A., Benoît G., Sadier A., Delsuc F., Douzery E., Laudet V. Conserved features and evolutionary shifts of the *EDA* signaling pathway involved in vertebrate skin appendage development. *Mol. Biol. Evol*. 2008;25(5):912-928. DOI 10.1093/molbev/msn038.

Park J.H., Yamaguchi T., Watanabe C., Kawaguchi A., Haneji K., Takeda M., Kim Y.I., Tomoyasu Y., Watanabe M., Oota H., Hanihara T., Ishida H., Maki K., Park S.B., Kimura R. Effects of an Asian-specific nonsynonymous *EDAR* variant on multiple dental traits. *J. Hum. Genet.* 2012;57(8):508-514. DOI 10.1038/jhg.2012.60.

Sadier A., Viriot L., Pantalacci S., Laudet V. The ectodysplasin pathway: from diseases to adaptations. *Trends Genet.* 2014;30(1):24-31. DOI 10.1016/j.tig.2013.08.006.

Sadier A., Lambert E., Chevret P., Décimo D., Sémon M., Tohmé M., Ruggiero F., Ohlmann T., Pantalacci S., Laudet V. Tinkering signaling pathways by gain and loss of protein isoforms: the case of the EDA pathway regulator EDARADD. *BMC Evol. Biol.* 2015;15:129. DOI 10.1186/s12862-015-0395-0.

Siemens H.W. Studien über Vererbung von Hautkrankheiten. XII. Anhidrosis hypotrichotica. *Arch. Dermat. Syph.* 1937;175:565-577. DOI 10.1007/BF02058385.

Smahi A., Courtois G., Rabia S., Döffinger R., Bodemer C., Munnich A., Casanova J., Israël A. The NF-κB signalling pathway in human diseases: from incontinentia pigmenti to ectodermal dysplasias and immune-deficiency syndromes. *Hum. Mol. Genet.* 2002;11(20):2371-2375. DOI 10.1093/hmg/11.20.2371.

Srivastava A., Pispa J., Hartung A., Du Y., Ezer S., Jenks T., Shimada T., Pekkanen M., Mikkola M., Ko M., Thesleff I., Kere J., Schlessinger D. The Tabby phenotype is caused by mutation in a mouse homologue of the *EDA* gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. *Proc. Natl. Acad. Sci. USA.* 1997;94(24):13069-13074. DOI 10.1073/pnas.94.24.13069.

Trzeciak W., Koczorowski R. Molecular basis of hypohidrotic ectodermal dysplasia: an update. *J. Appl. Genet*. 2015;57(1):51-61. DOI 10.1007/s13353-015-0307-4.

Vincent M., Biancalana V., Ginisty D., Mandel J., Calvas P. Mutational spectrum of the ED1 gene in X-linked hypohidrotic ectodermal dysplasia. *Eur. J. Hum. Genet*. 2001;9(5):355-363. DOI 10.1038/sj.ejhg.5200635.

Wahlbuhl-Becker M., Faschingbauer F., Beckmann M., Schneider H. Hypohidrotic ectodermal dysplasia: breastfeeding complications due to impaired breast development. *Geburtshilfe Frauenheilkd*. 2017;77(4):377-382. DOI 10.1055/s-0043-100106.

Wang K., Yamada S., Izumi H., Tsukamoto M., Nakashima T., Tasaki T., Guo X., Uramoto H., Sasaguri Y., Kohno K. Critical *in vivo* roles of WNT10A in wound healing by regulating collagen expression/synthesis in *WNT10A*-deficient mice. *PLoS One*. 2018;13(3):e0195156. DOI 10.1371/journal.pone.0195156.

Weech A. Hereditary ectodermal dysplasia (congenital ectodermal defect): a report of two cases. *Am. J. Dis. Child*. 1929;37(4):766-790. DOI 10.1001/archpedi.1929.01930040075005.

Wohlfart S., Meiller R., Hammersen J., Park J., Menzel-Severing J., Melichar V., Huttner K., Johnson R., Porte F., Schneider H. Natural history of X-linked hypohidrotic ectodermal dysplasia: a 5-year follow-up study. *Orphanet J. Rare Dis*. 2020;15(1):7. DOI 10.1186/s13023-019-1288-x.

Wright J.T., Fete M., Schneider H., Zinser M., Koster M.I., Clarke A.J., Hadj-Rabia S., Tadini G., Pagnan N., Visinoni A.F., Bergendal B., Abbott B., Fete T., Stanford C., Butcher C., D'Souza R.N., Sybert V.P., Morasso M.I. Ectodermal dysplasias: classification and organization by phenotype, genotype and molecular pathway. *Am. J. Med. Genet. A.* 2019;179(3):442-447. DOI 10.1002/ajmg.a.61045.

Xu M., Horrell J., Snitow M., Cui J., Gochnauer H., Syrett C.M., Kallish S., Seykora J.T., Liu F., Gaillard D., Katz J.P., Kaestner K.H., Levin B., Mansfield C., Douglas J.E., Cowart B.J., Tordoff M.,

Liu F., Zhu X., Barlow L.A., Rubin A.I., McGrath J.A., Morrisey E.E., Chu E.Y., Millar S.E. *WNT10A* mutation causes ectodermal dysplasia by impairing progenitor cell proliferation and KLF4-mediated differentiation. *Nat. Commun*. 2017;8:15397. DOI 10.1038/ncomms15397.

Yan M., Zhang Z., Brady J., Schilbach S., Fairbrother W., Dixit V. Identification of a novel death domain-containing adaptor molecule for ectodysplasin-A receptor that is mutated in crinkled mice. *Curr. Biol*. 2002;12(5):409-413. DOI 10.1016/s0960-9822(02)00687-5.

Yang J., Huang T., Petralia F., Long Q., Zhang B., Argmann C., Zhao Y., Mobbs C., Schadt E., Zhu J., Tu Z., GTEx Consortium. Synchronized age-related gene expression changes across multiple tissues in human and the link to complex diseases. *Sci. Rep*. 2015;5:15145. DOI 10.1038/srep15145. Erratum in: *Sci. Rep*. 2016;6:19384. [corrected to Goldmann J.].

Yu M., Wong S., Han D., Cai T. Genetic analysis: Wnt and other pathways in nonsyndromic tooth agenesis. *Oral Dis.* 2019;25(3):646-651. DOI 10.1111/odi.12931.

Zhang L., Yu M., Wong S.W., Qu H., Cai T., Liu Y., Liu H., Fan Z., Zheng J., Zhou Y., Feng H., Han D. Comparative analysis of rare *EDAR* mutations and tooth agenesis pattern in *EDAR*- and *EDA*-associated nonsyndromic oligodontia. *Hum. Mutat*. 2020;41(11):1957-1966. DOI 10.1002/humu.24104.

ORCID ID

V.A. Kovalskaia orcid.org/0000-0002-8728-8574
O.P. Ryzkova orcid.org/0000-0003-1285-9093

Conflict of interest. The authors declare no conflict of interest.

Received August 19, 2022. Revised February 6, 2023. Accepted March 24, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):684-693
DOI 10.18699/VJGB-23-79

Original Russian text https://vavilovj-icg.ru/

Problems with studying directional natural selection in humans

S.V. Mikhailova

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
mikhail@bionet.nsc.ru

Abstract. The review describes the main methods for assessing directional selection in human populations. These include bioinformatic analysis of DNA sequences via detection of linkage disequilibrium and of deviations from the random distribution of frequencies of genetic variants, demographic and anthropometric studies based on a search for a correlation between fertility and phenotypic traits, genome-wide association studies on fertility along with genetic loci and polygenic risk scores, and a comparison of allele frequencies between generations (in modern samples and in those obtained from burials). Each approach has its limitations and is applicable to different periods in the evolution of *Homo sapiens*. The main source of error in such studies is thought to be sample stratification, the small number of studies on nonwhite populations, the impossibility of a complete comparison of the associations found and functionally significant causative variants, and the difficulty with taking into account all nongenetic determinants of fertility in contemporary populations. The results obtained by various methods indicate that the direction of human adaptation to new food products has not changed during evolution since the Neolithic; many variants of immunity genes associated with inflammatory and autoimmune diseases in modern populations have undergone positive selection over the past 2–3 thousand years owing to the spread of bacterial and viral infections. For some genetic variants and polygenic traits, an alteration of the direction of natural selection in Europe has been documented, e.g., for those associated with an immune response and cognitive abilities. Examination of the correlation between fertility and educational attainment yields conflicting results. In modern populations, to a greater extent than previously, there is selection for variants of genes responsible for social adaptation and behavioral phenotypes. In particular, several articles have shown a positive correlation of fertility with polygenic risk scores of attention deficit/hyperactivity disorder.
Key words: natural selection; *Homo sapiens*; fertility; adaptation; polygenic index; genome-wide association study.

For citation: Mikhailova S.V. Problems with studying directional natural selection in humans. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):684-693. DOI 10.18699/VJGB-23-79

Проблемы изучения направленного естественного отбора у человека

С.В. Михайлова

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
mikhail@bionet.nsc.ru

Аннотация. В обзоре описаны основные методы оценки направленного отбора в популяциях человека. В их числе биоинформатический анализ последовательностей ДНК, основанный на выявлении неравновесия по сцеплению и отклонения от случайного распределения частот генетических вариантов; демографические и антропометрические исследования, заключающиеся в поиске корреляции рождаемости с фенотипическими признаками; полногеномные оценки ассоциаций фертильности с генетическими локусами и полигенными индексами, а также сравнение частот аллелей между поколениями (как в современных выборках, так и в полученных из захоронений ДНК). Каждый из этих подходов имеет свои ограничения и применим к разным периодам эволюции *Homo sapiens*. Основными источниками ошибок в таких исследованиях считаются стратификация выборок, ограниченное число исследований на неевропеоидном населении, невозможность полного сопоставления найденных ассоциаций и функционально значимых каузативных вариантов, а также сложность учета всех негенетических факторов, определяющих фертильность в современных популяциях. Полученные с помощью разных методов результаты свидетельствуют о том, что направление адаптации человека к новым для него пищевым продуктам не меняется в ходе эволюции с неолита; многие из вариантов генов иммунитета, ассоциированных в современных популяциях с воспалительными и аутоиммунными заболеваниями, подверглись положительному отбору в период последних 2–3 тыс. лет в связи с распространением бактериальных и вирусных инфекций. По некоторым генетическим вариантам и полигенным признакам показана смена направления естественного отбора на территории Европы, среди них связанные с иммунным ответом и когнитивными способностями. Анализ корреляции фертильности и уровня образо-

вания дает противоречивые результаты. В современных популяциях в большей степени, чем ранее, наблюдается отбор по вариантам генов, отвечающих за социальную адаптацию и поведенческие фенотипы. В том числе в нескольких работах показана положительная корреляция фертильности с полигенными индексами синдрома дефицита внимания/гиперактивности.
Ключевые слова: естественный отбор; *Homo sapiens*; фертильность; адаптация; полигенный индекс; полногеномный анализ ассоциаций.

Introduction

It has been shown that the prevalence of some diseases is increasing in human populations around the world. These include obesity (along with related disorders: type 2 diabetes mellitus and coronary heart disease), attention deficit/hyperactivity disorder (ADHD), autism spectrum disorders, and allergic diseases (Charpin, Gouitaa, 2001; Saklayen, 2018; Zeidan et al., 2022; Wolf et al., 2023). This phenomenon requires new approaches in medicine, social services, education, leisure, and nutrition of children and adolescents.

At the same time, there is no consensus on the causes of the observed changes; numerous studies are being conducted to find a relation of the diseases with the lifestyle, maternal stress, and environmental pollution. The human environment has changed dramatically in the last century with the advent of new chemical compounds, the growth of anthropogenic pollution, and the progress of medicine, including reproductive technologies. Urbanization has led to an increase in population density and social stress (Suvorov, 2021). The demographic transition has caused the aging of populations; the extended life expectancy affected the incidence of metabolic disorders and cardiovascular pathologies. In addition, research on the contribution of human genetics to the increased disease incidence remains relevant. Standard approaches of medical genetics aimed at finding associations of genetic variants with diseases do not allow to answer the question of the direction of selection in modern people; this is because (i) these diseases – especially in the postreproductive period – are not directly related to the number of children an individual has, (ii) homozygous or heterozygous carriage of a genetic variant can affect fitness in different ways, and (iii) most diseases are multifactorial, and the same is true for personality traits. Currently, several techniques are used to assess natural selection in human populations.

Bioinformatic methods

The analysis of DNA sequences in large human cohorts allows to estimate the direction of natural selection over the past millennia. For this purpose, methods, algorithms, and software packages have been developed based on estimates of the frequency of polymorphic variants and linkage disequilibrium in the areas of their location (Grossman et al., 2013; Field et al., 2016; Palamara et al., 2018; Speidel et al., 2019; Abondio et al., 2022). It is assumed that (a) longer haplotypes (areas that have not undergone recombination) in any region of the genome, (b) increased frequencies of derived alleles in adjacent loci, (c) a greater difference in allele frequencies between populations than could be expected during genetic drift, and (d) a higher rate of substitutions at each position (calculated from a genomic alignment of 100 vertebrate species) may be indicative of directional selection in the genomic region in question.

In a study performed on the 1000 Genomes database, 412 regions with signs of directional selection were analyzed (Grossman et al., 2013). These regions proved to be enriched with coding variants: 235 regions containing one or more protein-coding genes and 177 regions without genes encoding known proteins were found. In addition, 48 genes of long intergenic noncoding RNAs were revealed there. Among the 33 genes containing the most highly ranked single-nucleotide variants (SNVs), there are *SCL24A5* and *MATP* (associated with reduced skin pigmentation), *EDAR* (affecting the formation of hair, sweat glands, and teeth), *ARHGEF3* (affecting bone mineral density), *BTLA* (associated with rheumatoid arthritis), *CTNS* (affecting cysteine metabolism), *ITPR3* (associated with type 1 diabetes mellitus), innate-immunity receptor gene *TLR5*, *ITGAE* (involved in cell adhesion and lymphocyte activation), and *AP4B1* (associated with cerebral palsy). A total of 11 loci were associated with height and pigmentation, and 79 with a predisposition to infectious and autoimmune diseases. The *SLC24A5* gene, in addition to pigmentation, determines resistance to leprosy. Genes *ALMS1*, *CCR9*, *CXCR4*, and *VDR* are also located in the loci associated with resistance to infections and having signs of selection.

Another study, performed on 2,478 people of various backgrounds from the 1000 Genomes database, has identified several genomic regions with positive selection for multiallelic traits, including red and white blood cell counts, hair color, and the body mass index (BMI) (Speidel et al., 2019). Among the targets of positive selection, enrichment with SNVs in functional regions of the genome was observed in that paper. Signs of positive selection were found, in the major histocompatibility complex (MHC) locus and genes *LCT* (associated with lactose tolerance), *EDAR*, *EDARADD*, *HERC1* (associated with syndromic mental retardation), and *ATXN2* (associated with type 1 diabetes mellitus, obesity, and hypertension). In the same article, the proposed method of looking for signs of positive selection was applied to British Biobank data, resulting in a signal for SNVs associated with lighter hair in Europeans. The most significant signal identified is related to SNVs associated with a reduced BMI in white Americans. The largest number of loci with signs of selection was found in Europeans, whereas East Asians have the least number of signals (presumably owing to the "bottleneck" effect in their history) (Speidel et al., 2019).

In an examination of data from the British Biobank, 12 directional selection signals were detected, including those located near immunity genes (*TLR1-6-10*, *HLA*, *IGHG*, *STAT4*, *MUC5B*, *FAM19A5*, and *ANXA*), near the genes that determine pigmentation (*GRM5* and *MC1R*), and near *LCT* (Palamara et al., 2018). As a result of the analysis of 3,195 genomes from British project UK10K, signs of selection over the past 2–3 thousand years have been found in MHC locus and in *LCT* and *WDFY4* (associated with activation of T cells during

viral infections) genes as well as in the genes responsible for skin and hair pigmentation (*KITLG*, *OCA2/HERC2*, *ASIP*, and *SLC24A4*). Polygenic selection has been identified for variants predisposing to higher growth and lower total cholesterol in both sexes and to a reduced BMI in males (Field et al., 2016). By the same method applied to data of the British Biobank, other researchers (Song et al., 2021) have analyzed selection signals for 870 disease-associated polygenic traits and 15 non-disease-associated phenotypic traits. It was noted that during the last 2–3 thousand years, 88 % of polygenic traits have undergone selection. The highest scores were shown by genes that determine the ease of skin tanning and lighter hair. Selection for most disease-related loci, including those associated with autism spectrum disorders and elevated cholesterol, was negative; an exception was genes predisposing to Crohn's disease and ADHD. For polygenic scores of intelligence and insomnia, a change in the direction of adaptation was predicted at ~133 generations ago.

Twenty-nine genetic loci with signs of directional selection were identified in biological samples from the Japan Biobank (Yasumizu et al., 2020). The highest statistical significance was registered for the alcohol dehydrogenase (*ADH*) gene cluster and for genes *CIAO2A* (metal ion binding), *MYOF* (cell membrane regeneration), *GRIA2* (glutamate receptor), and *ASAP2* (vesicular transport). That work also confirmed previously reported selection pressure on *EDAR* genes and on the MHC gene cluster in the Japanese population.

Several studies have assessed the evolution of genes associated with inflammatory diseases separately. In an evaluation of the length of haplotypes containing 588 SNVs associated with 10 inflammatory diseases in European populations, signs of relatively recent (within the last 1,200–2,600 years) positive selection were identified in 21 loci, while variants associated with diseases, not protective alleles, were being selected (Raj et al., 2013). As a result of a comparison of several common multifactorial diseases and phenotypic traits in terms of the number of SNVs under positive selection associated with them, most of these SNVs were implicated specifically in inflammatory diseases. The genetic variants found are mainly involved in the molecular pathways participating in the activation of T helper 17 (Th17) lymphocytes (*STAT1*, *STAT3*, *STAT5*, *IRF1*, *CSF2*, *IL2*, *IL3*, *IL12A*, *IL2RA*, and *SOCS1*) (Raj et al., 2013). On data from the Estonian Biobank (2,300 whole-genome sequences), a search for signs of directional selection was performed at 535 loci associated with 21 autoimmune diseases. As a result, 153 loci showed signs of selection, while 29 of them were found to be selected due to linkage with other variants (Pankratov et al., 2022). The largest number of loci found in this work was associated with leukocyte activation and cytokine synthesis.

The search for genetic variants that have been under selection is difficult because a substantial number of identified genome loci with signs of directional selection have been found in intergenic regions (Grossman et al., 2013; Yasumizu et al., 2020), for which no functional significance can be predicted. As a consequence, there is no answer to the question whether this effect is due only to the insufficient level of knowledge about noncoding regions of the genome or to the inaccuracy of the algorithms used.

Comparison of genomes of ancient and modern humans

The accumulation of data on DNA sequences from human remains from burial sites makes it possible to compare them with genotypes of modern humans. A comparison of genotypes among people who lived at different times in the same geographic area is of particular interest.

When comparing ancient and modern genomes of people from Britain, investigators found seven loci to have signs of directional selection over the past 4,500 years (Mathieson, Terhost, 2022). Most of them are associated with vitamin D or calcium metabolism. It was demonstrated that the strength of selection for individual loci has changed over time, suggesting that some factors have appeared that have softened it. Among the 28 complex anthropometric and metabolic traits analyzed in that work, evidence of polygenic selection was detected only for skin pigmentation.

Polymorphic sites in innate-immunity genes associated with predisposition to mycobacterial infections (*SLC11A1*, *MBL2*, *TLR2*, *P2RX7*, *IL10*, and *TNFA*) have been studied in remains of 151 people in time series data (1st–18th centuries AD) from Northern and Eastern Poland (Lewandowska et al., 2018). This DNA analysis indicated that genetic drift has played the main role in the evolution of people in this region; however, two SNVs (rs17235409 of the divalent cation transporter *SLC11A1* gene and rs1800896 of interleukin *IL10* gene) manifested signs of nonrandom evolution.

In a collection of 1,013 genomes of Europeans born from the Mesolithic Age to the Middle Ages, the frequency of rs34536443 of the *TYK2* gene has been estimated (Kerner et al., 2021). This gene encodes a tyrosine kinase that is involved in signal transduction from cytokine receptors. The minor allele of rs34536443 has previously been shown to be associated with susceptibility to tuberculosis. In that work (Kerner et al., 2021), it was found that this rare allele emerged as a single mutational event during the Early Neolithic ~8,500 years ago on the Anatolian Peninsula, and then spread to Central Europe, where its frequency had remained within 3 % until ~5,000 years ago. In the Bronze Age, the frequency peaked (10 %) ~3,000 years ago, and after the Iron Age, it began to decline sharply until it reached 2.9 %. The observed changes in the frequency of rs34536443 are associated with the spread of tuberculosis in Europe. rs1800562 of the *HFE* gene (iron metabolism regulator) was also investigated by those authors (Kerner et al., 2021). Its maximum frequency in Europe (~10 %) was reached in the Middle Ages, and then it decreased in Europe to an average of 4 %.

Other work researchers have explored 827 ancient samples of European origin (from 25,000 years BC to the present) (Kuijpers et al., 2022). In accordance with the available data on genome-wide polygenic risk scores of inherited traits in Europeans, the genomes were compared at different time intervals. It was demonstrated that after the Neolithic, height and intelligence increased in the European population, skin pigmentation diminished, and the risk of coronary heart disease went up due to decreasing concentrations of high-density lipoprotein cholesterol. It was suggested that the latter trend is related to cognitive functions because variations in lipoprotein levels are associated with intelligence, learning, and memory.

To identify the loci that were selected during the bubonic plague epidemic (Black Death) in Europe (1,347–1,351 AD), an association analysis has been performed on immunity genes in 206 DNA samples originating from burials of two European populations before, during, and after this epidemic (London: years ~1,000–1,250 and 1,350–1,539, and Denmark: ~850–1,350 and ~1,350–1,800) (Klunk et al., 2022). Four positively selected loci (rs2549794, which affects *ERAP2* mRNA splicing, and three SNVs in noncoding regions: rs11571319, rs17473484, and rs1052025) proved to be common between British and Danish burials. It was shown that genes located near these SNVs (*ERAP1*, *ERAP2*, *LNPEP*, *CTLA4*, *ICOS*, *TICAM2*, and *TMED7*) are differentially expressed in human macrophages in response to *Yersinia pestis* infection.

An analysis of 187 polygenic traits in three sets of ancient human genomes (from 8,000–4,200, ~14,000–3,400, and ~45,000–7,000 years ago) indicates that in Near East genomes, selection signals for tan-determining genes varied depending on latitude; signs of positive selection were observed at low latitudes, and signs of negative one at high latitudes (Song et al., 2021). Positive selection signals were also demonstrated in that work for 13 disease loci, including Crohn's disease, atopic dermatitis, and periodontitis.

During an analysis of a study population consisting of the 1000 Genomes database and 230 remains of inhabitants of Eurasia (dated from 6,500 to 1,000 BC), the strongest selection signal was detected near the *LCT* gene (rs4988235) (Mathieson et al., 2015). Two other independent signals were found: near *FADS1* (rs174546) and *DHCR7/NADSYN1* (rs7940244). FADS1 and FADS2 are fatty acid desaturases taking part in the synthesis of long-chain polyunsaturated fatty acids from short precursors; variants at this locus correlate with plasma fatty-acid concentrations. The most statistically significant SNV at this locus (rs174546) is associated with a reduced level of triglycerides. 7-Dehydrocholesterol reductase (encoded by the *DHCR7* gene) participates in the metabolism of cholesterol and vitamin D. Additionally, directional selection was documented for the MHC and *TLR1-6-10* immunity gene loci, genes *HERC2* and *SLC45A2* (responsible for pigmentation), loci near genes *ATXN2*, *GRM5* (glutamate receptor), *ZKSCAN3* (transcriptional regulation), and *SLC22A4* (organic cation transporter).

The limitations of this approach to the analysis of directional selection are as follows: data from burial sites are fragmentary, and the observed alterations in genomes can be due to migrations from regions where the genetic pool was formed, for example, under the influence of bottleneck effects (Kerner et al., 2021) and assortative mating (Mills, Mathieson, 2022). In the work performed on Estonian Biobank data and on the Allen Ancient DNA Resource (AADR) V44.3 (Marnetto et al., 2022), it was shown that in a mixed population formed from previously isolated populations of pastoralists, hunter-gatherers, and farmers, some phenotypes are associated with the carriage of DNA sequences specific for ancestral populations in certain regions of the genome. Among other things, it was demonstrated that blood cholesterol levels among modern Estonians are positively correlated with the similarity of their genomes to those of carriers of the Yamnaya culture at loci associated with cholesterol levels. Thus, during the formation of a population, the observed difference in allele frequencies relative to an ancestral population may not necessarily indicate directional selection.

Estimation of the number of children in carriers of different phenotypes and genotypes

In addition to changes in the prevalence of genetic variants or phenotypic traits over time, the presence of directional natural selection is indicated by differing numbers of offspring in carriers of certain genotypes or traits. After the demographic transition that has taken place in many countries and has featured a decline of infant and child mortality with a simultaneous decrease in fertility, it is the estimation of the number of children that is the most accurate method for researching modern directional selection in human populations. Cross-sectional studies on populations of young people do not reflect the real number of offspring because there is a tendency to postpone the birth of the first and hence subsequent children (Balbo et al., 2013). In an analysis of a sample older than 50 years, dead individuals, which nevertheless had offspring, fall out of sight. Therefore, long-term studies are the most promising, especially those that enable a comparison of successive generations. The most commonly used indicators of fertility are (i) the total number of children ever born (NEB) in women over 45 and in men over 55, (ii) the age at first birth, which shows a strong inverse correlation with NEB (Tropf et al., 2015; Kong et al., 2017; Sanjak et al., 2018; Arkhangelskiy et al., 2020), (iii) relative reproductive success (the number of children per individual divided by the average number of children per individual in the population), and (iv) childlessness.

Demographic and anthropometric studies

Fertility is the subject of demographic research. Special attention is paid to problems of population reproduction in countries that have made the demographic transition because the observed decline of the birth rate leads to the "aging" of the population, to a decrease in the proportion of working-age people, and as a result, to a decline of economic growth rates. Demographers seek to predict population size and fertility and to detect its possible association with socioeconomic factors, which would allow to regulate population reproduction. Thus, demography essentially investigates the association of complex phenotypic traits with natural selection in a population.

A search for relations between fertility and income, educational attainment, religiosity, and anthropometric data has yielded mixed results (Table 1).

Some articles point to an inverse correlation between the NEB and the income of individuals and households (Arkhangelskiy et al., 2021; Turner, Robbins, 2023), and others have revealed correlations of different directions when men are compared with women (Fieder, Huber, 2022); it has also been noted that after the division of a cohort into subgroups by income level, the direction of the correlation changes (Cohen et al., 2013; Arkhangelskiy et al., 2021). Demographers pay special attention to the analysis of correlations between the birth rate and educational attainment. In the vast majority of papers, indicators of education and cognitive abilities show an inverse correlation with the number of offspring (Beauchamp, 2016; Reeve et al., 2018; Sanjak et al., 2018; Arkhangelskiy et al., 2020; Fieder, Huber, 2022) (see Table 1). Of note, among men, educational attainment correlates with

Table 1. Described correlations of the number of children ever born (NEB) with demographic and anthropometric characteristics, if sex is not specified: for both sexes

Phenotypic trait	Sex	Direction of correlation with NEB	Region	Reference
BMI		+	USA	Beauchamp, 2016
Educational attainment		–		
Height	Women	–		
BMI, weight, waist circumference	Men	+	UK	Sanjak et al., 2018
Educational attainment		–		
Body fat	Men	+		
Fluid intelligence score	Women	–		
Height	Women	–		
Income		–	USA	Turner, Robbins, 2023
Income		–	Russia	Arkhangelskiy et al., 2021
Educational attainment	Women	–	Russia	Arkhangelskiy et al., 2020
Body fat	Women	+	UK	Arner et al., 2021
Years of education		–	USA	Fieder, Huber, 2022
Attendance of religious services		+		
Educational attainment	Men	+		
Income	Men	+		
Income	Women	–		
General cognitive ability		–	Asia, Europe, and USA	Reeve et al., 2018
Height	Women	–	Europe	Byars et al., 2010
Weight		+		
Total cholesterol level		–		

income, the above effect may change (Fieder, Huber, 2022). A survey of 9,452 women from 27 EU countries indicates that the estimated NEB is higher among women with higher education, but this study deals with intentions, not actually born offspring (Testa, 2014).

Anthropometric studies have uncovered a slow change in some parameters in populations. For example, since the 1950s, there has been a 0.5 % increment in the number of cases of fetal disproportion during childbirth, and this phenomenon is associated with a weakening of natural selection as a result of the massive use of cesarean section (Mitteroecker et al., 2016). A positive correlation of the NEB with physical characteristics has been documented, including weight, the BMI, and body fat (Byars et al., 2010; Beauchamp, 2016; Sanjak et al., 2018; Arner et al., 2021) (see Table 1). Modern data on selection for height contradict those obtained from ancient specimens; it has been reported that the number of children is negatively correlated with females' height (Byars et al., 2010; Beauchamp, 2016). The Framingham two-generation study revealed reductions in total cholesterol and systolic blood pressure in a European population (Byars et al., 2010).

Nevertheless, the NEB depends on historical, cultural, economic, and social environments (for example, the availability of contraceptives and child care) (Barban et al., 2016). In demographic research, it is often impossible to establish causal relations between observed phenomena and to take into account all nongenetic factors, and this state of affairs makes it difficult to employ such data. In this regard, proposals are being made for the integration of genetic and demographic studies (Hugh-Jones, Abdellaoui, 2022).

Analysis of the association of fertility with genetic markers

The NEB has one of the highest levels of polygenicity of any trait (Mathieson et al., 2023). At the same time, NEB heritability in different articles is estimated at 14–46 % (Barban et al., 2016). In a sample of women from the UK and the Netherlands, it has been shown that up to 10 % of the NEB variance is determined by common genetic variants (Tropf et al., 2015). The main way to assess the effect of the genotype on fertility is a genome-wide analysis of the association of fertility rates in populations of modern people. The results of such research are given in Tables 2 and 3.

A meta-analysis of NEB genome-wide association studies (343,072 people) has revealed three main loci: rs10908474 (near genes *SLC27A3* and *GATAD2B*), rs13161115 (between genes *EFNA5* and *FBXL17*), and rs2415984 in the intron of long intergenic noncoding RNA gene *LINC00871* (the strongest signal). An additional study identified a locus near rs2415984 containing genes *ARHGAP27*, *PLEKHM1*, and

Table 2. Genome-wide associations identified for the total number of children ever born, if sex is not specified: for both sexes

Genetic locus and its functional significance	Phenotype associated with nearby genes according to https://www.ncbi.nlm.nih.gov/gene/	Sex	Reference
SLC27A3, *GATAD2B*; lipid metabolism, transcription repressor	Autism, intellectual disability		Barban et al., 2016
EFNA5, *FBXL17*; central nervous system development, blood vessel growth	Degenerative disorder of the nervous system	Men	
LINC00871 (*ARHGAP27*, *PLEKHM1*, and *MIR4315-1*); endocytosis, vesicular transport			
MC1R; melanocortin 1 receptor	Hair pigmentation		Mathieson et al., 2023
FADS1/2; fatty acid synthesis enzymes	Serum triglyceride level		
ARHGAP27; clathrin-mediated endocytosis			
PLEKHM1; vesicular transport	Autosomal recessive osteopetrosis type 6		
PIK3IP1; cell cycle regulator	Inhibition of T cell activation		
ZFP82; transcription regulator	Tumor suppression		
LRP4; WNT signaling regulator	Senani–Lenz syndrome		
GLDN; formation of peripheral neurons	Lethal congenital contracture syndrome		
RPS11; ribosomal complex protein			
PGGHG; carbohydrate metabolism			

Table 3. Polygenic associations found for the number of children ever born (NEB), if sex is not specified: for both sexes

Polygenic trait	Direction of correlation with NEB	Sex	Reference
ADHD risk	+		Demontis et al., 2019
PolyEduc, educational attainment polygenic score	–	Men	Fieder, Huber, 2022
POLYEDU, educational attainment polygenic score	–		Kong et al., 2017
Height	+	Men	Song et al., 2021
Skin color	+	Men	
High intelligence	–		
Income	–		Hugh-Jones, Abdellaoui, 2022
Educational attainment	–		
ADHD risk, depressive-disorder risk	+		
Coronary heart disease risk	+		
BMI	+		
Extraversion	+		
Years of education	–		Barban et al., 2016
Risk of autism spectrum disorder	–		

MIR4315-1. An analysis of potential functional significance of these variants indicates linkage disequilibrium of rs13161115 with a methylation site near the *EFNA5* gene (Barban et al., 2016) (see Table 2). For the loci associated with age at first birth, an inverse correlation with educational attainment has been shown. The NEB polygenic score inversely correlates with the number of years of education (see Table 3). Most of the correlations found in the above meta-analysis for the NEB and age at first birth polygenic scores are related to behavioral and reproductive phenotypes.

A recent genome-wide search for NEB and childlessness associations (785,604 Europeans) yielded 43 genomic loci

associated with age of puberty, age at first birth, sex hormone regulation, endometriosis, and age at menopause (Mathieson et al., 2023) (see Table 2). Among them, 28 are common between men and women and six are gender-specific to the NEB; nine, including a gender-specific one, are associated with childlessness. rs12949256 of *ARHGAP27* (p.Ala117Thr) is associated with a higher NEB but a shorter reproductive period. NEB-associated coding variants were found in genes *PIK3IP1* (rs2040533, p.Thr251Ser), *ZFP82* (rs17206365, p.Leu59Met), and *LRP4* (rs6485702, p.Ile1086Val). A comparison of data between modern (Mathieson et al., 2023) and ancient (Mathieson et al., 2015) Europeans shows that at the *FADS1/2* locus (biosynthesis of ω-3 and ω-6 lipids), directional selection has continued for several thousand years. A yet unknown role of the melanocortin 1 receptor (*MC1R*) gene in reproductive biology has been discovered. Its effect on the number of offspring was stronger in women. Although variants of this gene determine ~73 % of the heritability of the red hair color, phenotypically the red color in the study population was not associated with NEB. After exclusion of red-haired women from the analysis, observed *MC1R* effect on the number of offspring in the British Biobank persisted. No relation was detected between effects of specific SNVs on hair color and on NEB. Intron insertion/deletion polymorphism of the *CADM2* gene manifested the strongest association with childlessness. For this gene, which codes for a cell adhesion molecule and is expressed in the brain, a strong balancing selection signal and an association with risky behavior have been shown (Boutwell et al., 2017). A female-specific association with childlessness was found for transcription regulator gene *PPP3R1* (Mathieson et al., 2023). The results obtained in that paper were validated for 35 identified loci in a sample of 34,367 FinnGen women (Mathieson et al., 2023). None of the signals identified in that study showed significant genome-wide associations with educational attainment, church attendance, or indices of social deprivation. In an assessment of potential functional significance of the found variants, it was noted that most of the found signals of modern directional selection are related to the hypothalamic–pituitary–gonadal axis, which regulates fertility and reproductive aging.

In several articles, authors separately studied the relation between educational attainment and the NEB. In a sample of 129,808 Icelanders born between 1910 and 1990, the POLYEDU polygenic index of educational attainment showed a negative correlation with the NEB, with gradual diminution in several observed generations (Kong et al., 2017). In particular, an increase in the frequency of the rs62056842 variant in an intron of the *MAPT* gene, which is expressed in the nervous system and is associated with reduced educational attainment, was detected. When children born to mothers aged 21 or younger (18 % of all children in the sample) and children born to men aged 22 or younger (13 %) were excluded from the analysis, the correlations disappeared (Kong et al., 2017).

Another study points to a borderline inverse correlation between the NEB and the polygenic educational attainment score among men, with the number of children regressing positively in the interaction of income and the PolyEduc polygenic score, thus indicating a positive correlation between income and the number of offspring in men with genetic predisposition to higher education (Fieder, Huber, 2022). In an analysis of the association of the polygenic score of fertility with 33 polygenic traits in two generations of Europeans from the British Biobank (348,595 people of European descent, taking into account the number of siblings and the number of offspring), it was shown that polygenic scores that predict higher income and education attainment are correlated with reduced fertility, whereas polygenic risk scores for ADHD, depressive disorder, and coronary heart disease as well as a higher BMI and extraversion predict more offspring (Hugh-Jones, Abdellaoui, 2022). That paper indicates that educational attainment and the risk of ADHD and a depressive disorder are selected among young mothers (age at first birth before 22), but the natural selection was reversed among older mothers. On the other hand, several anthropometric polygenic scores are selected only among older ones. By contrast, the largest analysis in terms of sample size (Mathieson et al., 2023) did not reveal selection against educational attainment.

In modern society, out of 15 categories of genes that determine phenotypic traits, those responsible for metabolism, nutritional habits, psychiatric indicators, dermatological signs (in men), social cognition, and reproduction correlate with the NEB (Song et al., 2021).

A positive correlation between polygenic risk scores for ADHD and fertility was also demonstrated in a comparison of a sample of 20,183 people with ADHD and a control sample of 35,191 residents of Europe and the United States (Demontis et al., 2019).

A limitation of genome-wide association studies is that they detect only a genetic locus, not a specific gene or polymorphic site associated with a given trait. Just as in a bioinformatic analysis, a high proportion of the associations falls into intergenic regions and cannot always be interpreted. A lot of controversy is caused by the stratification of a population sample; this approach can significantly affect the results (Sohail et al., 2019; Mills, Mathieson, 2022). Researchers emphasize the nonrandom compilation of cohorts as a source of bias because the very consent to participate in the study correlates, for example, with educational attainment not only in the frequently used American commercial sample 23andMe but also in the British Biobank (Mills, Mathieson, 2022; Schork et al., 2022). It is reported that a meta-analysis of small populations of Europeans from small geographic areas can give incorrect results due to the different proportions of ancestral populations of pastoralists, hunter-gatherers, and farmers who participated in the formation of the modern population of Europe in different regions; on the other hand, when examining large populations settled in different climatic regions, it is necessary to take into account the influence of climatic factors on the phenotype. In the UK, geographic clustering of genetic variants was found that affects complex traits, including alleles associated with educational attainment, thereby proving the influence of demographic factors on the correlation between genes and the environment (Abdellaoui et al., 2019). The vast majority of genome-wide association studies have been performed on populations of European origin, and therefore the findings cannot be automatically extrapolated to humanity as a whole; in addition, the number of studies with sufficient statistical power is small.

Analysis of frequencies of genetic variants in close generations

The direction of selection can be assessed by comparison of genotypes among people born in the same population during periods adjacent to significant natural, socioeconomic, or political events in the region of their birth. The difference in allele frequencies in this case may reflect increased perinatal, prenatal, and infant mortality or a difference in the number of children among reproductive-age people carrying different genotypes during the period of the events under study.

Similar work was done on data from the British Biobank (Wu et al., 2022). A genome-wide association study of infant mortality rates by place and year of birth was performed. Cohorts born between 1936 and 1970 in England and Wales experienced a decline of infant mortality with spikes during World War II. Several statistically significant loci were found, including missense variant rs1446585 of the *R3HDM1* gene near the *LCT* gene and missense variant rs5743618 of the *TLR1* gene as well as rs2852853 in an intron of 7-dehydrocholesterol reductase gene *DHCR7* (vitamin D metabolism), rs9944197 in an intron of the gene of ribosomal protein *EFL1*, and intergenic rs10521293. Those authors were especially interested in the *LCT* and *TLR1/6/10* loci, which had previously shown natural selection among Europeans (Mathieson et al., 2015). The frequency of these alleles did not differ by year of birth in regions with low infant mortality but differed in this manner in regions with high infant mortality. The biggest difference was noted in 1942 (a year after the maximum of the German bombing), but the density of the bombings by region did not match the level of infant mortality; accordingly, those authors attribute the observed effect to harsh living conditions and food shortages.

A work comparing genotypes among three groups of adolescents who born before, during, and after the socioeconomic crisis of the 1990s in Russia was previously published by us (Mikhailova et al., 2022). We analyzed frequencies of common genetic variants previously found to be associated with stress resistance and stress-induced disorders in populations of Novosibirsk city schoolchildren aged 14–17 years. A statistically significant increase in frequencies of stress-protective variant rs4680 G of the *COMT* gene and long (7R+8R) tandem repeats in exon 3 of the *DRD4* gene was found in the "stress" group. Both genes belong to the dopaminergic regulatory system. We hypothesized that under the conditions of prolonged social stress, carriers of certain genotypes have more offspring due to better adaptation to the conditions of socioeconomic deprivation because social stress affects fertility.

A limitation of such research is that alterations of allele frequencies occur within some short periods and may not affect the genetic pool of the population, especially owing to a drop of the birth rate in populations during of major destructive events. Furthermore, it is difficult to take into consideration all the factors that can potentially alter allele frequencies.

Conclusion

Despite the limitations of each approach and the lack of information about directional selection in African and Asian populations, several dozen genetic variants and a number of polygenic traits have been found that have undergone natural selection during human evolution. Genetic loci and phenotypic traits have been identified whose direction of natural selection has not changed from the Neolithic to our time, although the intensity of the selection has varied (*LCT* and *FAD1/2*). For some genetic variants, the direction of adaptation changed, probably as a result of an encounter with pathogens (e. g., rs34536443 of the *TYK2* gene). In modern populations, there has been a reversal of directional selection relative to previous evolution for polygenic scores of height in women and for the BMI in men. A considerable number of genetic variants that are reported to be associated with inflammatory diseases (including Crohn's disease and atopic dermatitis) have been positively selected in the past, but it is not clear whether these variants are now under selection pressure or the observed genotype ratio is already a consequence of balancing selection due to antagonistic pleiotropy. Conflicting data have been obtained about selection of relatively recent complex polygenic traits: income and educational attainment. It has been shown that the targets of selection – to a greater extent than in previous centuries – are genes responsible for social adaptation and behavioral phenotypes. For instance, the positive association of the ADHD polygenic risk score with fertility as documented by several researchers is indicative of selection for this phenotype in modern populations.

References

Abdellaoui A., Hugh-Jones D., Yengo L., Kemper K.E., Nivard M.G., Veul L., Holtz Y., Zietsch B.P., Frayling T.M., Wray N.R., Yang J., Verweij K.J.H., Visscher P.M. Genetic correlates of social stratification in Great Britain. *Nat. Hum. Behav*. 2019;3(12):1332-1342. DOI 10.1038/s41562-019-0757-5.

Abondio P., Cilli E., Luiselli D. Inferring signatures of positive selection in whole-genome sequencing data: an overview of haplotype-based methods. *Genes*. 2022;13(5):926. DOI 10.3390/genes13050926.

Arkhangelskiy V.N., Shulgin S.G., Zinkina Yu.V. Reproductive behavior of Russian women as depending on their level of education. *Vestnik Rossiyskogo Universiteta Druzhby Narodov. Seriya: Sotsiologiya = RUDN Journal of Sociology*. 2020;20(3):546-559. DOI 10.22363/2313-2272-2020-20-3-546-559. (in Russian)

Arkhangelskiy V.N., Rostovskaya T.K., Vasilieva E.N. Influence of the standard of living on the reproductive behavior of Russians: gender aspect. *Zhenshchina v Rossiyskom Obshchestve = Woman in Russian Society*. 2021;Spec.iss.:3-24. DOI 10.21064/WinRS.2021.0.1. (in Russian)

Arner A.M., Grogan K.E., Grabowski M., Reyes-Centeno H., Perry G.H. Patterns of recent natural selection on genetic loci associated with sexually differentiated human body size and shape phenotypes. *PLoS Genet*. 2021;17(6):e1009562. DOI 10.1371/journal.pgen.1009562.

Balbo N., Billari F.C., Mills M. Fertility in advanced societies: a review of research: La fécondité dans les sociétés avancées: un examen des recherches. *Eur. J. Popul*. 2013;29(1):1-38. DOI 10.1007/s10680-012-9277-y.

Barban N., Jansen R., de Vlaming R., Vaez A., Mandemakers J.J., Tropf F.C., Shen X., Wilson J.F., Chasman D.I., Nolte I.M., … Lee J.J., Benjamin D.J., Cesarini D., Koellinger P.D., den Hoed M., Snieder H., Mills M.C. Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet*. 2016;48(12):1462-1472. DOI 10.1038/ng.3698.

Beauchamp J.P. Genetic evidence for natural selection in humans in the contemporary United States. *Proc. Natl. Acad. Sci. USA*. 2016;113(28):7774-7779. DOI 10.1073/pnas.1600398113.

Boutwell B., Hinds D.; 23andMe Research Team; Tielbeek J., Ong K.K., Day F.R., Perry J.R.B. Replication and characterization of *CADM2*

and *MSRA* genes on human behavior. *Heliyon*. 2017;3(7):e00349. DOI 10.1016/j.heliyon.2017.e00349.

Byars S.G., Ewbank D., Govindaraju D.R., Stearns S.C. Colloquium papers: natural selection in a contemporary human population. *Proc. Natl. Acad. Sci. USA*. 2010;107(Suppl.1):1787-1792. DOI 10.1073/pnas.0906199106.

Charpin D., Gouitaa M. Why is the prevalence of allergic diseases increasing? A critical assessment of some classical risk factors. *Mediators Inflamm*. 2001;10(6):292-294. DOI 10.1080/09629350152700920.

Cohen A., Dehejia R., Romanov D. Financial incentives and fertility. *Rev. Econ. Stat.* 2013;95(1):1-20. DOI 10.1162/REST_a_00342.

Demontis D., Walters R.K., Martin J., Mattheisen M., Als T.D., Agerbo E., Baldursson G., Belliveau R., Bybjerg-Grauholm J., Bækvad-Hansen M., … Werge T., Mors O., Mortensen P.B., Daly M.J., Faraone S.V., Børglum A.D., Neale B.M. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet*. 2019;51(1):63-75. DOI 10.1038/s41588-018-0269-7.

Fieder M., Huber S. Contemporary selection pressures in modern societies? Which factors best explain variance in human reproduction and mating? *Evol. Hum. Behav*. 2022;43(1):16-25. DOI 10.1016/j.evolhumbehav.2021.08.001.

Field Y., Boyle E.A., Telis N., Gao Z., Gaulton K.J., Golan D., Yengo L., Rocheleau G., Froguel P., McCarthy M.I., Pritchard J.K. Detection of human adaptation during the past 2000 years. *Science*. 2016;354(6313):760-764. DOI 10.1126/science.aag0776.

Grossman S.R., Andersen K.G., Shlyakhter I., Tabrizi S., Winnicki S., Yen A., Park D.J., Griesemer D., Karlsson E.K., Wong S.H., Cabili M., Adegbola R.A., Bamezai R.N., Hill A.V., Vannberg F.O., Rinn J.L.; 1000 Genomes Project; Lander E.S., Schaffner S.F., Sabeti P.C. Identifying recent adaptations in large-scale genomic data. *Cell.* 2013;152(4):703-713. DOI 10.1016/j.cell.2013.01.035.

Hugh-Jones D., Abdellaoui A. Human capital mediates natural selection in contemporary humans. *Behav. Genet*. 2022;52(4-5):205-234. DOI 10.1007/s10519-022-10107-w.

Kerner G., Laval G., Patin E., Boisson-Dupuis S., Abel L., Casanova J.L., Quintana-Murci L. Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2,000 years. *Am. J. Hum. Genet*. 2021;108(3):517-524. DOI 10.1016/j.ajhg.2021.02.009.

Klunk J., Vilgalys T.P., Demeure C.E., Cheng X., Shiratori M., Madej J., Beau R., Elli D., Patino M.I., Redfern R., DeWitte S.N., Gamble J.A., Boldsen J.L., Carmichael A., Varlik N., Eaton K., Grenier J.C., Golding G.B., Devault A., Rouillard J.M., Yotova V., Sindeaux R., Ye C.J., Bikaran M., Dumaine A., Brinkworth J.F., Missiakas D., Rouleau G.A., Steinrücken M., Pizarro-Cerdá J., Poinar H.N., Barreiro L.B. Evolution of immune genes is associated with the Black Death. *Nature*. 2022;611(7935):312-319. DOI 10.1038/s41586-022-05349-x.

Kong A., Frigge M.L., Thorleifsson G., Stefansson H., Young A.I., Zink F., Jonsdottir G.A., Okbay A., Sulem P., Masson G., Gudbjartsson D.F., Helgason A., Bjornsdottir G., Thorsteinsdottir U., Stefansson K. Selection against variants in the genome associated with educational attainment. *Proc. Natl. Acad. Sci. USA*. 2017;114(5):E727-E732. DOI 10.1073/pnas.1612113114.

Kuijpers Y., Domínguez-Andrés J., Bakker O.B., Gupta M.K., Grasshoff M., Xu C.J., Joosten L.A.B., Bertranpetit J., Netea M.G., Li Y. Evolutionary trajectories of complex traits in European populations of modern humans. *Front. Genet*. 2022;13:833190. DOI 10.3389/fgene.2022.833190.

Lewandowska M., Jędrychowska-Dańska K., Płoszaj T., Witas P., Zamerska A., Mańkowska-Pliszka H., Witas H.W. Searching for signals of recent natural selection in genes of the innate immune response – ancient DNA study. *Infect. Genet. Evol*. 2018;63:62-72. DOI 10.1016/j.meegid.2018.05.008.

Marnetto D., Pankratov V., Mondal M., Montinaro F., Pärna K., Vallini L., Molinaro L., Saag L., Loog L., Montagnese S., Costa R.; Estonian Biobank Research Team; Metspalu M., Eriksson A., Pagani L. Ancestral genomic contributions to complex traits in contemporary Europeans. *Curr. Biol.* 2022;32(6):1412-1419.e3. DOI 10.1016/j.cub.2022.01.046.

Mathieson I., Terhorst J. Direct detection of natural selection in Bronze Age Britain. *Genome Res*. 2022;32(11-12):2057-2067. DOI 10.1101/gr.276862.122.

Mathieson I., Lazaridis I., Rohland N., Mallick S., Patterson N., Roodenberg S.A., Harney E., Stewardson K., Fernandes D., Novak M., Sirak K., Gamba C., Jones E.R., Llamas B., Dryomov S., Pickrell J., Arsuaga J.L., de Castro J.M., Carbonell E., Gerritsen F., Khokhlov A., Kuznetsov P., Lozano M., Meller H., Mochalov O., Moiseyev V., Guerra M.A., Roodenberg J., Vergès J.M., Krause J., Cooper A., Alt K.W., Brown D., Anthony D., Lalueza-Fox C., Haak W., Pinhasi R., Reich D. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature.* 2015;528(7583):499-503. DOI 10.1038/nature16152.

Mathieson I., Day F.R., Barban N., Tropf F.C., Brazel D.M.; eQTLGen Consortium; BIOS Consortium; Vaez A., van Zuydam N., Bitarello B.D., … Zhao W., Zhao Y., Snieder H., den Hoed M., Ong K.K., Mills M.C., Perry J.R.B. Genome-wide analysis identifies genetic effects on reproductive success and ongoing natural selection at the FADS locus. *Nat. Hum. Behav*. 2023;7(5):790-801. DOI 10.1038/s41562-023-01528-6.

Mikhailova S.V., Ivanoshchuk D.E., Yushkevich E.A., Bairqdar A., Anisimenko M.S., Shcherbakova L.V., Denisova D.V., Orlov P.S. Prevalence of common alleles of some stress resilience genes among adolescents born in different periods relative to the socioeconomic crisis of the 1990s in Russia. *Curr. Issues Mol. Biol*. 2022;45(1):51-65. DOI 10.3390/cimb45010004.

Mills M.C., Mathieson I. The challenge of detecting recent natural selection in human populations. *Proc. Natl. Acad. Sci. USA*. 2022;119(15):e2203237119. DOI 10.1073/pnas.2203237119.

Mitteroecker P., Huttegger S.M., Fischer B., Pavlicev M. Cliff-edge model of obstetric selection in humans. *Proc. Natl. Acad. Sci. USA*. 2016;113(51):14680-14685. DOI 10.1073/pnas.1612410113.

Palamara P.F., Terhorst J., Song Y.S., Price A.L. High-throughput inference of pairwise coalescence times identifies signals of selection and enriched disease heritability. *Nat. Genet.* 2018;50(9):1311-1317. DOI 10.1038/s41588-018-0177-x.

Pankratov V., Yunusbaeva M., Ryakhovsky S., Zarodniuk M.; Estonian Biobank Research Team; Yunusbayev B. Prioritizing autoimmunity risk variants for functional analyses by fine-mapping mutations under natural selection. *Nat. Commun*. 2022;13(1):7069. DOI 10.1038/s41467-022-34461-9.

Raj T., Kuchroo M., Replogle J.M., Raychaudhuri S., Stranger B.E., De Jager P.L. Common risk alleles for inflammatory diseases are targets of recent positive selection. *Am. J. Hum. Genet*. 2013;92(4):517-529. DOI 10.1016/j.ajhg.2013.03.001.

Reeve C.L., Heeney M.D., Woodley of Menie M.A. A systematic review of the state of literature relating parental general cognitive ability and number of offspring. *Pers. Individ. Differ.* 2018;134:107-118. DOI 10.1016/j.paid.2018.05.036.

Saklayen M.G. The global epidemic of the metabolic syndrome. *Curr. Hypertens. Rep*. 2018;20(2):12. DOI 10.1007/s11906-018-0812-z.

Sanjak J.S., Sidorenko J., Robinson M.R., Thornton K.R., Visscher P.M. Evidence of directional and stabilizing selection in contemporary humans. *Proc. Natl. Acad. Sci. USA*. 2018;115(1):151-156. DOI 10.1073/pnas.1707227114.

Schork A.J., Peterson R.E., Dahl A., Cai N., Kendler K.S. Indirect paths from genetics to education. *Nat. Genet*. 2022;54(4):372-373. DOI 10.1038/s41588-021-00999-5.

Sohail M., Maier R.M., Ganna A., Bloemendal A., Martin A.R., Turchin M.C., Chiang C.W., Hirschhorn J., Daly M.J., Patterson N., Neale B., Mathieson I., Reich D., Sunyaev S.R. Polygenic adaptation on height is overestimated due to uncorrected stratification

in genome-wide association studies. *eLife.* 2019;8:e39702. DOI 10.7554/eLife.39702.

Song W., Shi Y., Wang W., Pan W., Qian W., Yu S., Zhao M., Lin G.N. A selection pressure landscape for 870 human polygenic traits. *Nat. Hum. Behav*. 2021;5(12):1731-1743. DOI 10.1038/s41562-021-01231-4.

Speidel L., Forest M., Shi S., Myers S.R. A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet*. 2019; 51(9):1321-1329. DOI 10.1038/s41588-019-0484-x.

Suvorov A. Population numbers and reproductive health. *Endocrinology*. 2021;162(11):bqab154. DOI 10.1210/endocr/bqab154.

Testa M.R. On the positive correlation between education and fertility intentions in Europe: individual- and country-level evidence. *Adv. Life Course Res*. 2014;21:28-42. DOI 10.1016/j.alcr.2014.01.005.

Tropf F.C., Stulp G., Barban N., Visscher P.M., Yang J., Snieder H., Mills M.C. Human fertility, molecular genetics, and natural selection in modern societies. *PLoS One*. 2015;10(6):e0126821. DOI 10.1371/journal.pone.0126821.

Turner N., Robbins K. Association between county-level natality and income in the US, 2000–2020. *JAMA Pediatr*. 2023;177(2):198-202. DOI 10.1001/jamapediatrics.2022.4814.

Wolf E., Sonenklar N., Schefft M., Haskell H., James J. Is there evidence of ADHD overdiagnosis in children? *Am. Fam. Physician*. 2023;107(3):292-296.

Wu Y., Furuya S., Wang Z., Nobles J.E., Fletcher J.M., Lu Q. GWAS on birth year infant mortality rates provides evidence of recent natural selection. *Proc. Natl. Acad. Sci. USA*. 2022;119(12):e2117312119. DOI 10.1073/pnas.2117312119.

Yasumizu Y., Sakaue S., Konuma T., Suzuki K., Matsuda K., Murakami Y., Kubo M., Palamara P.F., Kamatani Y., Okada Y. Genome-wide natural selection signatures are linked to genetic risk of modern phenotypes in the Japanese population. *Mol. Biol. Evol*. 2020; 37(5):1306-1316. DOI 10.1093/molbev/msaa005.

Zeidan J., Fombonne E., Scorah J., Ibrahim A., Durkin M.S., Saxena S., Yusuf A., Shih A., Elsabbagh M. Global prevalence of autism: a systematic review update. *Autism Res*. 2022;15(5):778-790. DOI 10.1002/aur.2696.

ORCID ID

S.V. Mikhailova orcid.org/0000-0002-0897-5473

Funding. This research was funded by Russian Science Foundation grant No. 22-28-00866.

Acknowledgements. The English language was corrected by shevchuk-editing.com

Conflict of interest. The author declares no conflict of interest.

Received April 29, 2023. Revised July 3, 2023. Accepted July 3, 2023.

ECOLOGICAL GENETICS
Original article

Вавиловский журнал генетики и селекции. 2023;27(6):694-702
DOI 10.18699/VJGB-23-80

Original Russian text https://vavilovj-icg.ru/

A study of macroinvertebrate communities in Bolshiye Koty Bay of Lake Baikal using DNA metabarcoding

L.S. Kravtsova[1] ✉, T.E. Peretolchina[1], T.I. Triboy[1], I.A. Nebesnykh[1], A.E. Tupikin[2], M.R. Kabilov[2]

[1] Limnological Institute of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia
[2] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ lk@lin.irk.ru

Abstract. The diversity of macroinvertebrates, the structure of their communities in Bolshiye Koty Bay (Lake Baikal) was studied by a DNA metabarcoding approach using an Illumina MiSeq system. Internal primer *mlCOIintF* in combination with *jgHCO2198* of the Folmer fragment of the *COI* gene were used for macroinvertebrate metabarcoding. A total of 118009 reads of the *COI* gene fragment (at least 313 bp in length) were obtained. The correlation of the Spearman coefficient (S = 0.6, $p < 0.05$) with the abundance of macroinvertebrates in the samples before DNA extraction showed that the number of reads can serve as an indirect characteristic of the abundance of a species (operational taxonomic unit, OTU). 115 OTUs belonging to the higher taxa of macroinvertebrates were identified: Porifera, 1; Platyhelminthes, 3; Annelida, 38; Arthropoda, 55; Mollusca, 18. At a high level of resolution (with homology with GenBank reference sequences ≥ 95 %, coverage ≥ 90 %), 46 taxa of macroinvertebrates comprising three communities were registered: one dominated by molluscs (*Choanomphalus* conf. *maacki*) and two dominated by chironomids (*Orthocladius gregarius* Linev., *Sergentia baicalensis* Tshern.). Communities are characterized by low species diversity according to Shannon (from 0.7 to 1.2 bits), high concentration of dominance according to Simpson (from 0.5 to 0.7) and low evenness according to Pielou (from 0.3 to 0.4). Dominants and subdominants in the communities account for 91 to 96 % of *COI* gene fragment reads. The spatial distribution of the dominant species identified in the communities is influenced by the geomorphological features of the bottom and the composition of sediments in the area studied. The approach proposed for studying the structure of macroinvertebrate communities based on DNA metabarcoding and next generation sequencing can be recommended for express assessment of the state of aquatic ecosystems in the monitoring.

Key words: communities of macroinvertebrates; diversity; DNA metabarcoding; *COI*; high-throughput sequencing technologies; Lake Baikal.

For citation: Kravtsova L.S., Peretolchina T.E., Triboy T.I., Nebesnykh I.A., Tupikin A.E., Kabilov M.R. A study of macroinvertebrate communities in Bolshiye Koty Bay of Lake Baikal using DNA metabarcoding. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):694-702. DOI 10.18699/VJGB-23-80

Исследование сообществ макробеспозвоночных животных в бухте Большие Коты озера Байкал с использованием ДНК метабаркодинга

Л.С. Кравцова[1] ✉, Т.Е. Перетолчина[1], Т.И. Трибой[1], И.А. Небесных[1], А.Е. Тупикин[2], М.Р. Кабилов[2]

[1] Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия
[2] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ lk@lin.irk.ru

Аннотация. Приводятся сведения о разнообразии макробеспозвоночных животных, структуре их сообществ в бухте Большие Коты оз. Байкал, полученные методом ДНК метабаркодинга на основе NGS-технологии (Illumina, MiSeq). Для ДНК метабаркодинга макробеспозвоночных был использован внутренний праймер *mlCOIintF* в комбинации с *jgHCO2198* для амплификации фолмеровского фрагмента гена *COI*. Всего получено 118009 прочтений фрагмента гена *COI* (длиной не менее 313 п.н.). Показано, что количество прочтений может служить опосредованной характеристикой обилия вида (операционной таксономической единицы – ОТЕ). Корреляция количества прочтений с численностью макробеспозвоночных в пробах до экстракции ДНК по коэффициенту Спирмена составляет 0.6 ($p < 0.05$). Выявлено 115 ОТЕ, принадлежащих высшим таксонам макробеспозвоночных животных: Porifera – 1, Platyhelminthes – 3, Annelida – 38, Arthropoda – 55, Mollusca – 18. На видовом уровне (при гомологии с референсными последовательностями GenBank ≥ 95 % и покрытии не менее 90 %) зарегистрировано 46 таксонов макробеспозвоночных, формирующих три сообщества: одно – с доминированием моллюсков *Choanomphalus* conf. *maacki* и два – с доминированием хирономид *Orthocladius gregarius* Linev., *Sergentia*

baicalensis Tshern. Сообщества характеризуются невысоким видовым разнообразием по Шеннону (от 0.7 до 1.2 бит), высокой концентрацией доминирования по Симпсону (от 0.5 до 0.7) и низкой выравненностью по Пиелу (от 0.3 до 0.4). На долю доминантов и субдоминантов в сообществах приходится от 91 до 96 % прочтений фрагмента гена *COI*. На пространственное распределение доминирующих видов сообществ влияют геоморфологические особенности дна в исследуемом районе и состав донных отложений. Предложенный подход для изучения структуры сообществ макробеспозвоночных на основе ДНК метабаркодинга может быть рекомендован для экспресс-оценки состояния водных экосистем при мониторинге.
Ключевые слова: сообщества макробеспозвоночных; разнообразие; ДНК метабаркодинг; *COI*; высокопроизводительное секвенирование; Байкал.

Introduction

The structure of aquatic organisms communities, in particular macroinvertebrates, is one of the indicators characterizing the state of water bodies. Research in this direction is relevant in connection with global climate change and increasing anthropogenic impact on aquatic ecosystems (O'Reilly et al., 2003; Bonada et al., 2007; Burgmer et al., 2007; Moss et al., 2011; Hampton et al., 2018).

Community of organisms is a set of populations of different species coexisting in space and time (Begon et al., 1986). Their structure is formed under the influence of both abiotic environmental factors (Brauns et al., 2007; McGoff et al., 2013; Rezende et al., 2014; Worrall et al., 2014) and biotic interactions (van den Berg et al., 1997; Arbačiauskas et al., 2008; Nalepa et al., 2009). As a rule, the species diversity and abundance of organisms are the basic characteristics of the structure of communities. The study of the diversity of organisms at a high-resolution species level requires the involvement of a large number of morphologists and is associated with a laborious process of taxa identification. Currently, molecular genetics methods such as DNA metabarcoding using high-throughput sequencing technologies are used as an alternative to the classical methods of the taxonomic diversity studying of aquatic ecosystems. This method is widely used to study the diversity of both marine and freshwater fauna (Porazinska et al., 2009; Hajibabaei et al., 2011; Aylagas et al., 2014; Elbrecht et al., 2017; Haenel et al., 2017; Kuntke et al., 2020).

In the early 2000s, the mitochondrial *COI* gene was adopted as a standard for DNA barcoding of animal taxa (Hebert et al., 2003). Despite the universality of primers (*LCO1490* and *HCO2198*) for the "Folmer fragment" of the mitochondrial *COI* gene amplification (Folmer et al., 1994), many researchers began to use different, more variable regions of it to obtain a clear phylogenetic signal. For example, for marine nematodes, region *I3–M11* was used (Derycke et al., 2010). The mini-barcode of the Folmer *COI* fragment using a combination of primers *mlCOIintF* and *jgHCO2198* has become popular to assess the diversity of Metazoa (Meusnier et al., 2008; Leray et al., 2013). The mini-barcode also proved to be effective in assessing the diversity of benthic invertebrates in the Listvennichny Bay of Lake Baikal (Kravtsova et al., 2021).

In the last decade, the application of environmental DNA (eDNA) approaches for rapid assessment of the biodiversity from water and sediment samples has been increased (Yu et al., 2012; Lacoursière-Roussel et al., 2018). The advantage of this method is a quick result, since no preliminary isolation of any organisms from the samples is required. However, it was less effective for studying the diversity of Metazoa in water bodies. It was shown that DNA metabarcoding of invertebrate tissues gives a more accurate estimation of the diversity of multicellular organisms (99 % of reads) than DNA from the environment (only 12 % of reads) (Gleason et al., 2021). In this study, we tried to find out the acceptability of the method of DNA metabarcoding from the tissues of organisms for assessing not only the diversity of macroinvertebrates, but also for the quantitative ratio of species that form communities.

The aim of this work is to study the features of the structural organization of macroinvertebrate communities distributed in the coastal zone of open Baikal using DNA metabarcoding.

Materials and methods

Quantitative samples of zoobenthos were collected in July 2019 in Bolshiye Koty Bay of Lake Baikal along the coastline over a 1 km long (Fig. 1).

Macroinvertebrates were collected from different types of bottom sediments at three stations (No. 1–3) (see Fig. 1). The first type of bottom sediment included large and small pebbles with individual boulders located in the subaqueous part of the beach (depths 0.3–0.4 m). The second type of bottom sediment was represented by boulders, unrounded rock fragments with crushed stone, and the third type was represented by silted sand. The last two types of sediments were found mainly on a shallow terrace at depths of 2–5 m (see Fig. 1). On the underwater part of the beach and on a shallow terrace samples were taken manually and with the help of divers, respectively. On each type of bottom sediment, five quantitative samples of zoobenthos were collected using a 0.1 m^2 counting frame. Invertebrates from the surface of stones were brushed off into a cuvette with water, and those from silted sand were removed by flotation in a saturated sugar solution with a specific gravity of 1.12 g/L. Samples were washed through a mill sieve No. 23 and fixed with 96 % ethanol. In total, 15 quantitative samples of zoobenthos were collected and sorted under laboratory conditions using an MBS-10 microscope (at 20× magnification).

According to Elbrecht et al. (2017), preliminary sorting of organisms by size significantly improves the result of sequencing of all taxa, regardless of the biomass of the organism. Since the zoobenthos in Baikal is represented by different size groups: mega-, macro-, meioorganisms, in this work we limited ourselves to only one size group – macroinvertebrates. Their sizes in samples varied from 2 to 50 mm. All macroinvertebrates found in the sample were used for DNA extraction. First, invertebrates were soaked in distilled water (1 hour), small organisms (2–3 mm) were taken as a whole, and tissue pieces of 2–3 mm were taken from large individuals (more than 5 mm) in order to level the scatter of biomass

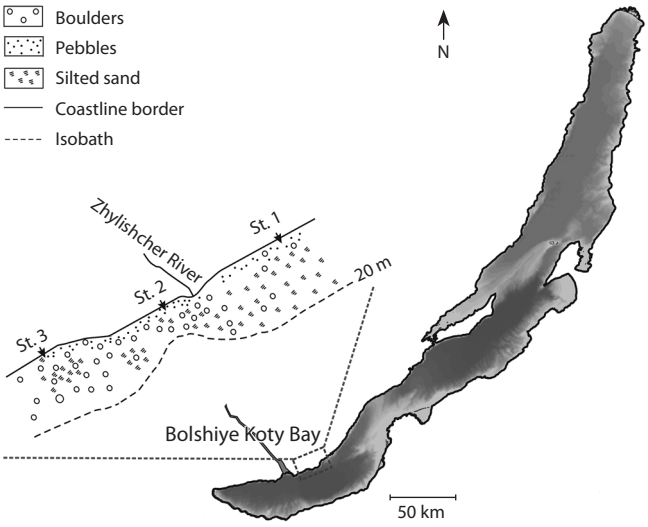

Fig. 1. Map-scheme of sampling localities of macrozoobenthos in Bolshiye Koty Bay, Lake Baikal (July, 2019).

in size. Then, pieces of tissue and small organisms collected from the same type of bottom sediment were combined into one sample, placed in a porcelain cup with 2 % CTAB solution, and ground with a pestle. DNA was extracted according to a modified (chloroform was used instead of a phenol-chloroform-isoamyl mixture) method of Doyle and Dickson (Doyle, Dickson, 1987). In total, three samples of genomic DNA (at least 20 ng each) of invertebrate animals were prepared for metabarcoding.

Primers *mlCOIintF*: GGWACWGGWTGAACWGTW TAYCCYCC (Leray et al., 2013) and *jgHCO2198*: TAIA CYTCIGGRTGICCRAARAAYCA (Geller et al., 2013), where "I" is inosine, were used to obtain *COI* gene amplicons. Amplification was performed in a volume of 20 µl containing 0.2 mM of each dNTP, 0.5 µM of each primer, 2 mM $MgCl_2$, 5 µM SYTO9, 10 ng of DNA, 25 U/ml of Maxima Hot Start Taq DNA Polymerase (Thermo Scientific, Lithuania). Real-time PCR was performed on a CFX96 Touch Real-Time PCR system (Bio-Rad, USA) according to the program: initial denaturation at 95 °C, 4 min; 32 cycles at 95 °C – 30 s, 48 °C – 30 s and 72 °C – 30 s. The primer annealing temperature was selected using gradient PCR. PCR products were analyzed on an MCE-202 MultiNA Microchip Electrophoresis System using a DNA 12000 Reagent Kit (Shimadzu, Japan). The resulting amplicons were quantified on a Qubit 2.0 fluorimeter with a Qubit DNA High Sensitivity Assay Kits (Invitrogen, USA). NEBNext Ultra II DNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1) (NEB, UK) were used to obtain DNA libraries from amplicons. The resulting libraries were quantified using the Kapa SYBR Fast Universal qPCR Kit (KapaBiosystems, USA). Library sequencing was performed using the MiSeq Reagent Standard Kit v3 PE300 on Miseq (Illumina, USA) at the Central Collective Use Center "Genomics" (ICBFM SB RAS).

The resulting paired reads were analyzed with UPARSE scripts (Edgar, 2013) using Usearch v11.0.667 (Edgar, 2010). Bioinformatics analysis included overlapping paired reads, filtering by quality and length, accounting for identical sequences, discarding singletons, removing chimeras, and obtaining an OTU (operational taxonomic unit). The taxonomic identification of the OTU sequences was determined using the SINTAX algorithms (Edgar, 2016) and the reference database MIDORI_UNIQUE_20180221_COI_SINTAX (Machida et al., 2017), as well as BLAST. The representativeness of OTU samples was analyzed using iNEXT 2.0.15 (Hsieh et al., 2016). The nucleotide sequences of the OTUs identified were tested for the presence of stop codons using the SeqKit v0.16.1 (https://doi.org/10.1371/journal.pone.0163962).

The Spearman correlation coefficient (S) was calculated to assess the relationship between the initial abundance of organisms of higher taxa in samples (before DNA extraction) and the number of reads of the *COI* gene fragment.

Communities of macroinvertebrates were identified by clustering the OTUs of the species rank and their *COI* reads by the Ward method; the Euclidean metric was used as a distance measure. Dominants and subdominants in the communities were determined by descending ranking of the modified density index (occurrence of OTUs in samples and their relative abundance = reads, in %) (Brotskaya, Zenkevich, 1939; Konstantinov, 1986). The communities of macroinvertebrates were named according to the dominant species with the highest density index. Species with a density index of ≥10 % were classified as subdominants. Species with a density index ≤10 % were considered minor. To characterize the degree of complexity of the community structure, the Shannon diversity (H, bit), Simpson dominance (D), and Pielou evenness (e) indices were used (Odum, 1983).

The spatial distribution of dominants and subdominants of communities in Bolshiye Koty Bay was analyzed using the principal component method, the relative abundance of OTUs and the composition of bottom sediments were used as variables. The data for calculations using multivariate statistics were previously transformed by the log(X+1) function. Calculations were carried out using the statistical programming environment R 3.0.0, package vegan 2.0-7.

Results

A total of 118,009 reads of the *COI* gene fragment (at least 313 bp in length) were obtained using high-throughput sequencing. Bioinformatics analysis showed that the proportion of unclassified *COI* sequences was less than 1 % of the total (1,157). 116,852 reads account for 115 OTUs belonging to the higher taxa of macroinvertebrates: Porifera – 1, Platyhelminthes – 3, Annelida – 38, Arthropoda – 55, Mollusca – 18.

The Spearman correlation coefficient (S) between the abundance of macroinvertebrates of higher taxa (Platyhelminthes, Hirudinea, Polychaeta, Oligochaeta, Isopoda, Amphipoda, Trichoptera, Chironomidae, Bivalvia, Gastropoda) found in samples before DNA extraction with the number of reads of the *COI* gene fragment is 0.6 ($p < 0.05$), which allows us to take the indices of the latter as equivalent to the relative abundance of organisms.

Forty-six taxa of species and genus ranks with ≥98 % and ≥95 % homology with GenBank sequences, respectively, were identified with coverage of at least 90 % (Table 1). They account for 88 % of the *COI* gene fragment reads.

Table 1. Composition of OTUs with ≥ 95% homology to GenBank reference sequences

Taxon	Species	GenBank numbers of reference sequences	Homology to the GenBank reference sequences, %
Porifera	*Baikalospongia fungiformis* Mak.	MH985288	99.7
Polychaeta/Fabriciidae	*Manayunkia baicalensis* (Nusb.)	MK393734	97.6
	Manayunkia godlewskii (Nusb.)	MK393737	99.3
	Manayunkia zenkewitschii Sit., Shcherb. et Kharch.	KF289863	99.2
Oligochaeta	*Rhynchelmis alyonae* Mart., Ferrag. et Kayg.	GU328670	98.3
	Rhynchelmis sp.	AJ577632	95.0
Amphipoda	*Baikalogammarus pullus* (Dyb.)	FJ756303	99.7
	Eulimnogammarus cyaneus (Dyb.)	MK887720	99.7
	Eulimnogammarus verrucosus (Gerstf.)	MK887569	100.0
	Eulimnogammarus vittatus (Dyb.)	MK887750	99.7
	Linevichella vortex (Dyb.)	MN148355	99.7
	Micruropus whalii (Dyb.)	MN148354	99.7
Diptera/Chironomidae	*Microtendipes* sp.	LC329125	95.0
	Paratanytarsus baicalensis (Tshern.)	MT020734	99.3
	Sergentia baicalensis Tshern.	AF116586	99.0
	Sergentia conf. *affinis*	AF116588	96.5
	Cricotopus fuscus (Kieff.)	MN673037	98.6
	Cricotopus conf. *intersectus*	MN683031	96.9
	Cricotopus sylvestris (Fabr.)	KC250789	100.0
	Cricotopus triannulatus (Macq.)	KJ439943	99.6
	Diplocladius cultriger Kieff.	HQ941599	98.6
	Orthocladius gregarius Linev.	KC879234	100.0
	Orthocladius nitidoscutellatus Lundstr.	MT048130	98.6
	Orthocladius sp.	KT248920	100.0
	Orthocladius sp. 2	KT248920	100.0
	Paratrichocladius rufiventris (Meig.)	HQ941597	99.3
Trichoptera	*Apatania majuscula* MacLach.	KX103052	99.3
	Baicalina bellicosa Mart.	KR153132	99.7
	Baicalina thamastoides Mart.	KR153144	99.6
	Baicalinella foliata (Mart.)	KR153101	99.7
	Protobaicalina spinosa (Mart.)	KR153124	98.9
Mollusca/Bivalvia	*Pisidium casertanum* (Poli)	KF483386	99.6
	Pisidium henslowanum (Shepp.)	KF483398	98.9
	Pisidium sp.	KF000182	99.3
	Pisidium sp. 1	KF483372	95.0
Mollusca/Gastropoda	*Benedictia fragilis* W. Dyb.	KX241839	99.3
	Choanomphalus amauronius Bourg.	Y14721	98.0
	Choanomphalus conf. *anomphalus*	Y14714	95.4
	Choanomphalus conf. *maacki*	LC429414	95.3
	Choanomphalus conf. *schrencki*	Y14713	96.1
	Gerstfeldtiancylus conf. *roepstorfi*	KR822550	96.2
	Maackia herderiana (Lindh.)	KY697388	99.7
	Megalovalvata baicalensis (Gerstf.)	LC377798	99.7
	Pseudancylastrum sibiricum (Gerstf.)	KR822557	100.0
	Pseudobaikalia contabulata (W. Dyb.)	Z92987	99.7
	Radix auricularia (Linn.)	MH190039	100.0

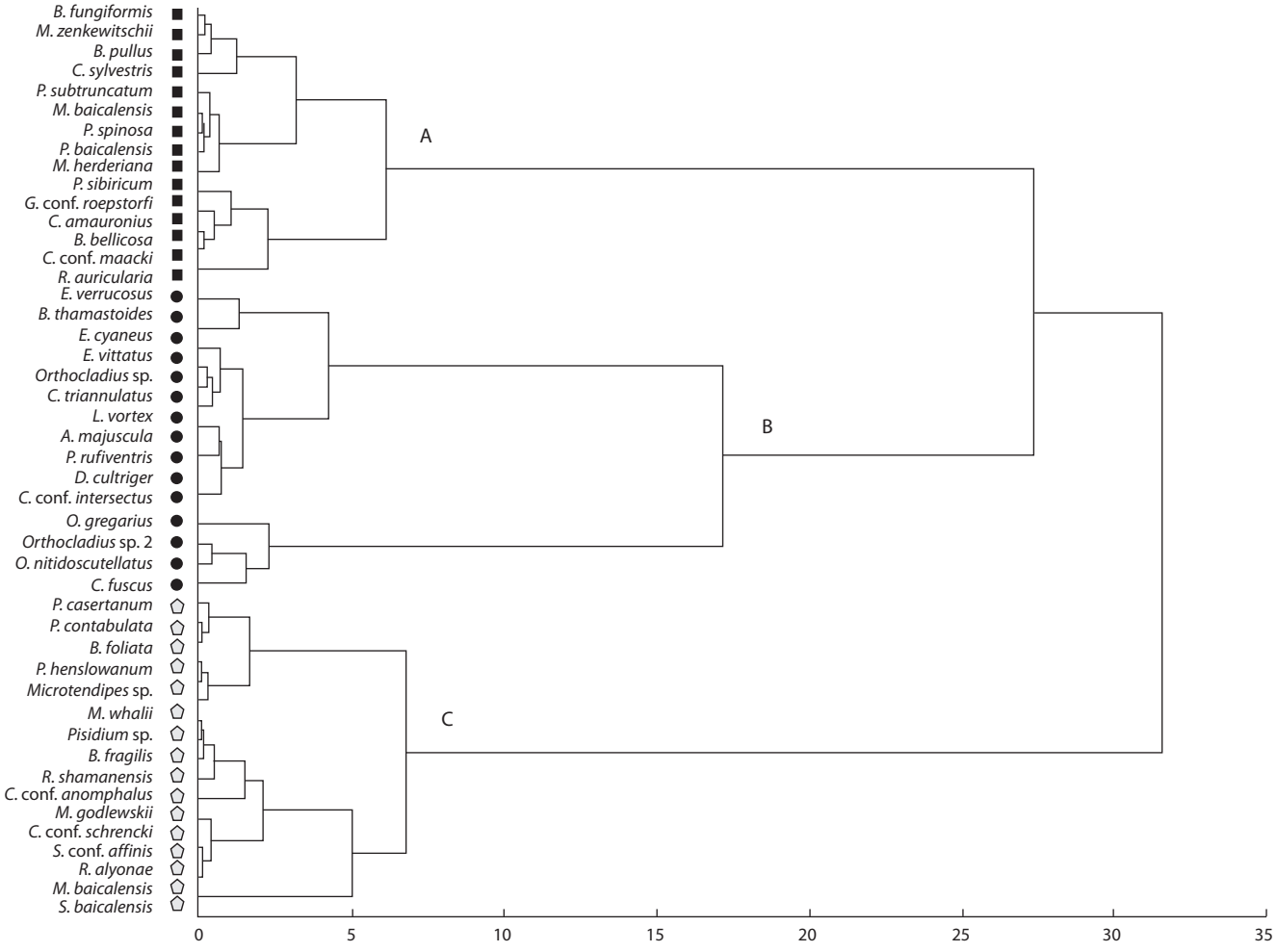

Fig. 2. Dendrogram of OTUs of macroinvertebrates constructed by the Ward clustering method using the Euclidean metric as a measure of distance. Clusters A, B, C on the dendrogram are macroinvertebrate communities identified based on reads of the *COI* gene fragment.

Table 2. Structural parameters of macroinvertebrate communities found in Bolshiye Koty Bay, Lake Baikal (July, 2019)

Community	Species number	Percentage of reads of dominants	H, bit	D	e	Number of reads per community
Orthocladius gregarius	15	96	1.20	0.45	0.44	39,941
Choanomphalus conf. *maacki*	15	93	0.74	0.70	0.28	32,573
Sergentia baicalensis	16	91	0.70	0.74	0.25	30,118

Note. H – species diversity according to Shannon, D – Simpson dominance index, e – Pielou's evenness index.

Three communities of macroinvertebrates were identified in the region studied, one of which was dominated by the molluscs *Choanomphalus* conf. *maacki* (A) and two were chironomid-dominated by *Orthocladius gregarius* Linev. (B), *Sergentia baicalensis* Tshern. (C) (Fig. 2).

Communities are characterized by a simple structure; 15–16 species were noted in their composition (Table 2). The Shannon index is low, ranging from 0.7 to 1.2 bits. At the same time, in the communities, there is a high concentration of dominance of one species (D varies from 0.5 to 0.7) and low evenness (e varies from 0.3 to 0.4).

In the community *C.* conf. *maacki* among the subdominants there are molluscs *Pseudancylastrum sibiricum* (Gerstf.), *Gerstfeldtiancylus* conf. *roepstorfi*. The subdominants of the *O. gregarius* community include the caddisflies *Baicalina thamastoides* Mart., chironomids *Orthocladius* sp. 2, *O. nitidoscutellatus* Lundstr., and *Cricotopus fuscus* (Kieff.). The community dominated by *S. baicalensis* contains numerous polychaetes *Manayunkia godlewskii* (Nusb.), molluscs *Choanomphalus* conf. *anomphalus*. Dominants and subdominants in communities account for 91 to 96 % of *COI* gene fragment reads.

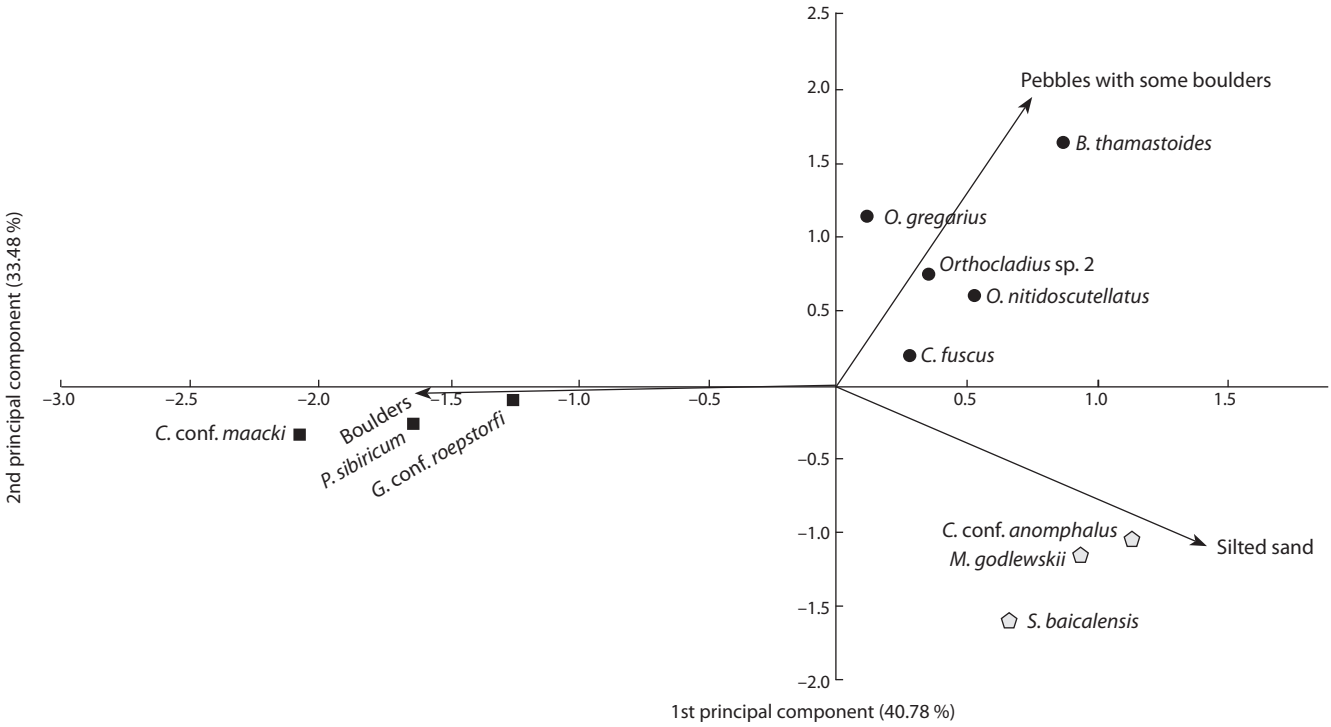

Fig. 3. Distribution of dominant species (OTUs) of macroinvertebrate communities (A–C) in the space of the first two principal components, taking into account 74 % of the variability of the relative abundance data set.

The spatial distribution of the dominant species of macroinvertebrate communities depending on environmental factors is shown in Figure 3. The first principal component characterizes the distribution of communities depending on the composition of bottom sediments. The dominant species of macroinvertebrate communities found on fine clastic material (pebbles, silted sand) have positive loads on the first principal component, and negative loads on coarse clastic material (boulders, rock fragments). The second principal component characterizes the distribution of dominants of the studied macroinvertebrate communities depending on the bottom geomorphology. The species found in the beach area have positive loads, while those found on the shallow-water terrace have negative loads.

Discussion

The studies of Bailey et al. (2001) indicate the effectiveness of monitoring observations of the state of water bodies using taxonomy at the genus or species level. Metabarcoding, as a modern genetic tool, is widely used for express assessment of biodiversity in ecosystems (Elbrecht, Leese, 2015). The authors of the work recommend using the presence/absence metric to characterize diversity, since the high resolution of primers makes it possible to take into account mass and minor species, but at the same time it does not make it possible to measure the absolute values of their abundance or biomass in samples. However, to study the structure of communities, indicators characterizing not only the composition of taxa, but also their quantitative ratio are needed. The number of reads, apparently, can be attributed to an indirect characteristic of species abundance (OTU). This confirms the presence of a positive correlation between the abundance of macroinvertebrates before DNA extraction and the abundance of OTUs from Lake Baikal not only in the samples collected in Bolshiye Koty Bay ($S = 0.6$, $p < 0.05$), but also in Listvennichny Bay ($S = 0.5$) (Kravtsova et al., 2021). In addition, a positive correlation was noted between the number of reads and the biomass of organisms collected in ponds in Germany (Elbrecht, Leese, 2015). Therefore, when studying the structure of macroinvertebrate communities, emphasis was placed on the number of reads per OTU, as well as the occurrence of OTUs in samples. To understand the extent to which the molecular genetics method using high-throughput sequencing technologies is applicable to the study of macroinvertebrate communities, the same approach was used for comparison as in ecological studies of water bodies based on classical hydrobiological methods (Brotskaya, Zenkevich, 1939; Begon et al., 1986; Konstantinov, 1986).

The fauna of macroinvertebrates on the shallow-water terrace and in the underwater part of the beach of the studied area of Lake Baikal in Bolshiye Koty Bay (excluding the population of the underwater slope and canyon) is quite diverse; in 1988, at least 177 species were recorded in its composition (Kravtsova et al., 2003). Most of the species found using DNA metabarcoding in 2019 were present here earlier, but in general, the diversity of benthic fauna (see Table 1) identified using molecular genetics methods is lower. This can be explained, on the one hand, by the smaller volume of collected quantitative samples of macrozoobenthos, and, on the other hand, by the absence in GenBank of *COI* sequences belonging to Baikal rich in species groups, for example, oligochaetes. In Baikal, 202 species of oligochaetes have been recorded, among them 165 are endemic (Semernoy, 2004). It is possible that the low representation of *COI* sequences of this group

of invertebrates in GenBank can be connected with the poor knowledge of the fauna due to its high endemism. In total, Annelida accounts for about 30 % of the total number of OTUs identified by DNA metabarcoding both from Bolshiye Koty Bay and Listvennichny Bay of Lake Baikal (Kravtsova et al., 2021).

Polychaeta, which are not rich in species, are also included in Annelida, but they were not previously included in the list of taxa from Bolshiye Koty Bay (Kravtsova et al., 2003), since their morphological identification had not been carried out. However, using DNA metabarcoding, in 2019, all three polychaete species found in Baikal were recorded in the fauna with high homology (98–99 %) with sequences from GenBank (see Table 1) (Pudovkina et al., 2016): *Manayunkia baicalensis* (Nusb.), *M. godlewskii*, *M. zenkewitschii* Sit., Shcherb. et Kharch.

Chironomids (Diptera) are among the objects, the identification of species of which by the morphological characters of the larvae is extremely difficult and they are often determined to a group of species (gr.) or species (sp.). It is possible that *O. gregarius*, the dominant of one of the three communities identified in 2019, was previously listed as *O.* gr. *thienemanni*, and the subdominant *O. nitidoscutellatus* was in *O.* gr. *olivaceus* (Kravtsova et al., 2003). Despite these species having been present in Bolshiye Koty Bay in 1988, they did not play a community-forming role among other macroinvertebrates (Kravtsova et al., 2004). The species *O. gregarius* and *O. nitidoscutellatus*, as shown by further molecular genetic studies (Kravtsova et al., 2014), have a rather long evolutionary history and their existence in Lake Baikal is beyond doubt (Makarchenko E.A., Makarchenko M.A., 2008).

The low diversity of species rank OTUs indicates that the macroinvertebrate communities dominated by *C*. conf. *maacki*, *O. gregarius*, *S. baicalensis* in Bolshiye Koty Bay are characterized by a simple structure (see Table 2) due to the above reasons.

It is known that abiotic environmental factors play an important role in the distribution and formation of the diversity of macroinvertebrate communities (Rezende et al., 2014). The spatial distribution of dominants and subdominants of macroinvertebrate communities from Bolshiye Koty Bay is consistent with the distribution of these species in the coastal zone of Lake Baikal. So, a community dominated by *C*. conf. *maacki* is confined, as before, to the rocky bottom sediments of the shallow terrace. The community of *O. gregarius* is distributed on pebbles with individual boulders in the beach zone, and with the dominance of *S. baicalensis*, on the silty sands of a shallow-water terrace, a typical biotope for this species. In Lake Baikal, representatives of the genus *Orthocladius* prefer to settle in a hydrodynamically active zone of wave mixing and water flow, while *Sergentia* prefer to settle in conditions where sedimentation processes prevail over erosion and transfer of terrigenous material, organic matter.

The study of the structure of macroinvertebrate communities in aquatic ecosystems using DNA metabarcoding has its own characteristics, in contrast to the rapid assessment of fauna diversity (based on the presence/absence metric). First of all, it is necessary to pay attention to the degree of study of the diversity of the bottom fauna of the reservoir, the size groups of its representatives. Of no small importance for the assessment of diversity is the presence of sequences of the studied gene fragment in the GenBank database. To obtain an adequate characterization of the relative abundance (reads of the *COI* gene fragment) of organisms in the community, quantitative sampling of macrozoobenthos should be carried out from a certain area, taking into account the biotopic heterogeneity of the bottom. In order to avoid the influence of body size of organisms on the number of reads per OTU, when preparing samples for DNA extraction, it is necessary to select tissue pieces of the same size from all individuals found in quantitative samples. This makes it possible to obtain an integral characteristic of the relative abundance of an organism (number of reads), taking into account its biomass, leveled by scatter (due to body size), as well as abundance. Since macroinvertebrates make up one third of the representatives of the unique Baikal fauna (2,565 species and subspecies (Timoshkin, 1997)), a more complete database of *COI* reference sequences is required for an objective assessment of α-diversity.

Conclusion

DNA metabarcoding using a primer (*mlCOIintF*) in combination with *jgHCO2198* to amplify the Folmer *COI* gene fragment showed its effectiveness in studies of the diversity and structure of Baikal macroinvertebrate communities. The fauna of Bolshiye Koty Bay contains typical representatives of most groups of macroinvertebrates inhabiting the coastal zone of Lake Baikal. It is shown that the number of reads, as a characteristic of the relative abundance of taxa, can be used to study the features of the structural organization of macroinvertebrate communities. In general, the proposed approach is quite acceptable for assessing the stability of macroinvertebrate communities in the temporal aspect in the monitoring of aquatic ecosystems.

References

Arbačiauskas K., Semenchenko V., Grabovski M., Leuven R., Paunović M., Son M., Csanyi B., Gumuliauskaitė S., Konopacka A., Nehring S., van der Velde G., Vezhnovetz V., Panov V. Assessment of biocontamination of benthic macroinvertebrate communities in European inland waterways. *Aquat. Invasions*. 2008;3(2):211-230. DOI 10.3391/ai.2008.3.2.12.

Aylagas E., Borja A., Rodríguez-Ezpeleta N. Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). *PLoS One*. 2014;9(3):e90529. DOI 10.1371/journal.pone.0090529.

Bailey R.C., Norris R.H., Reynoldson T.B. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. *J. North Am. Benthol. Soc.* 2001;20(2):280-286. DOI 10.2307/1468322.

Begon M., Harper J., Townsend K. Ecology: From Individuals to Ecosystems. Malden, MA: Blackwell Publishing, 1986. (Russ. ed.: Begon M., Harper J., Taunsend K. Ekologiya. Moscow: Mir Publ., 1989)

Bonada N., Dolédec S., Statzner B. Taxonomic and biological trait differences of stream macroinvertebrate communities between mediterranean and temperate regions: implications for future climatic scenarios. *Glob. Chang. Biol.* 2007;13(8):1658-1671. DOI 10.1111/j.1365-2486.2007.01375.x.

Brauns M., Garcia X.-F., Pusch M.T., Walz N. Eulittoral macroinvertebrate communities of lowland lakes: discrimination among trophic states. *Freshw. Biol.* 2007;52(6):1022-1032. DOI 10.1111/j.1365-2427.2007.01.

Brotskaya V.A., Zenkevich L.A. Quantitative accounting of the benthic fauna of the Barents Sea. *Proceedings of Russian Federal Research Institute of Fisheries and Oceanography*. 1939;4:5-98. (in Russian)

Burgmer T., Hillebrand H., Pfenninger M. Effects of climate-driven temperature changes on the diversity of freshwater macroinvertebrates. *Oecologia*. 2007;151(1):93-103. DOI 10.1007/s00442-006-0542-9.

Derycke S., Vanaverbeke J., Rigaux A., Backeljau T., Moens T. Exploring the use of cytochrome oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. *PLoS One*. 2010;5(10): e13716. DOI 10.1371/journal.pone.0013716.

Doyle J.J., Dickson E.E. Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*. 1987;36(4):715-722. DOI 10.2307/1221122.

Edgar R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26(19):2460-2461. DOI 10.1093/bioinformatics/btq461.

Edgar R.C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods*. 2013;10(10):996-998. DOI 10.1038/nmeth.2604.

Edgar R.C. SINTAX, a Simple Non-Bayesian Taxonomy Classifier for 16S and ITS sequences. *bioRxiv.* 2016. DOI 10.1101/074161.

Elbrecht V., Leese F. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass – sequence relationships with an innovative metabarcoding protocol. *PLoS One*. 2015;10(7):e0130324. DOI 10.1371/journal.pone.0130324.

Elbrecht V., Vamos E.E., Meissner K., Aroviita J., Leese F. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods Ecol. Evol.* 2017;8(10):1265-1275. DOI 10.1111/2041-210X.12789.

Folmer O., Hoeh W.R., Black M.B., Vrijenhoek R.C. Conserved primers for PCR amplification of mitochondrial DNA from different invertebrate phyla. *Mol. Mar. Biol. Biotechnol.* 1994;3(5):294-299.

Geller J., Meyer C., Parker M., Hawk H. Redesign of PCR primers for mitochondrial cytochrome *c* oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* 2013;13(5):851-861. DOI 10.1111/1755-0998.12138.

Gleason J.E., Elbrecht V., Braukmann T.W.A., Hanner R.H., Cottenie K. Assessment of stream macroinvertebrate communities with eDNA is not congruent with tissue-based metabarcoding. *Mol. Ecol.* 2021;30(13):3239-3251. DOI 10.1111/mec.15597.

Haenel Q., Holovachov O., Jondelius U., Sundberg P., Bourlat S.J. NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö island, Smögen, and soft mud from Gullmarn Fjord, Sweden. *Biodivers. Data J.* 2017;5:e12731. DOI 10.3897/BDJ.5.e12731.

Hajibabaei M., Shokralla S., Zhou X., Singer G.A.C., Baird D.J. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*. 2011; 6(4):e17497. DOI 10.1371/journal.pone.0017497.

Hampton S.E., McGowan S., Ozersky T., Virdis S.G.P., Vu T.T., Spanbauer T.L., Kraemer B.M., Swann G., Mackay A.W., Powers S.M., Meyer M.F., Labou S.G., O'Reilly C.M., DiCarlo M., Galloway A.W.E., Fritz S.C. Recent ecological change in ancient lakes. *Limnol. Oceanogr.* 2018;63(5):2277-2304. DOI 10.1002/lno.10938.

Hebert P.D.N., Cywinska A., Ball S.L., DeWaard J.R. Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 2003;270(1512): 313-321. DOI 10.1098/rspb.2002.2218.

Hsieh T.C., Ma K.H., Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* 2016;7(12):1451-1456. DOI 10.1111/2041-210X.12613.

Konstantinov A.S. General Hydrobiology. Moscow, 1986. (in Russian)

Kravtsova L.S., Karabanov E.B., Kamaltynov R.M., Mekhanikova I.V., Sitnikova T.Ya., Rozhkova N.A., Slugina Z.V., Izhboldina L.A., Weinberg I.V., Akinshina T.V., Krivonogov S.K., Shcherbakov D.Yu. Macrozoobenthos of subaqueous landscapes in the shoal of southern Lake Baikal: 1. Local diversity of bottom populations and features of their spatial distribution. *Zoologicheskiy Zhurnal = Zoological Journal*. 2003;82(3):307-317. (in Russian)

Kravtsova L.S., Kamaltynov R.M., Karabanov E.B., Mekhanikova I.V., Sitnikova T.Ya., Rozhkova N.F., Slugina Z.V., Izhboldina L.A., Weinberg I.V., Akinshina T.V., Sherbakov D.Yu. Macrozoobenthic communities of underwater landscapes in the shallow-water zone of southern Lake Baikal. *Hydrobiol*. 2004;522:193-205. DOI 10.1023/B:HYDR.0000029979.68265.3e.

Kravtsova L.S., Peretolchina T.E., Triboy T.I., Sherbakov D.Y. The evolutionary history of two species of Orthocladiinae (Diptera: Chironomidae) from Lake Baikal (Eastern Siberia). *Aquat. Insects*. 2014;36(3-4):171-185. DOI 10.1080/01650424.2015.1062111.

Kravtsova L.S., Peretolchina T.E., Triboy T.I., Nebesnykh I.A., Kupchinskiy A.B., Tupikin A.E., Kabilov M.R. The study of the diversity of hydrobionts from Listvennichny Bay of Lake Baikal by DNA metabarcoding. *Russ. J. Genet*. 2021;57(4):460-467. DOI 10.1134/S1022795421040050.

Kuntke F., de Jonge N., Hesselsøe M., Nielsen J.L. Stream water quality assessment by metabarcoding of invertebrates. *Ecol. Indic.* 2020; 111:105982. DOI 10.1016/j.ecolind.2019.105982.

Lacoursière-Roussel A., Howland K., Normandeau E., Grey E.K., Archambault P., Deiner K., Lodge D.M., Hernandez C., Leduc N., Bernatchez L. eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecol. Evol.* 2018;8(16):7763-7777. DOI 10.1002/ece3.4213.

Leray M., Yang J.Y., Meyer C.P., Mills S.C., Agudelo N., Ranwez V., Boehm J.T., Machida R.J. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 2013;10(1):34. DOI 10.1186/1742-9994-10-34.

Machida R.J., Leray M., Ho S.-L., Knowlton N. Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Sci. Data*. 2017;4(1):170027. DOI 10.1038/sdata.2017.27.

Makarchenko E.A., Makarchenko M.A. New findings of chironomids (Diptera, Chironomidae, Orthocladiinae) in Far East and bordering territories. III. *Orthocladius* van der Wulp. *Evraziatskii Entomologicheskii Zhurnal = Euroasian Entomological Journal*. 2008; 7(2):243-262. (in Russian)

McGoff E., Aroviita J., Pilotto F., Miler O., Solimini A.G., Porst G., Jurca T., Donohue L., Sandin L. Assessing the relationship between the Lake Habitat Survey and littoral macroinvertebrate communities in European lakes. *Ecol. Indic.* 2013;25:205-214. DOI 10.1016/j.ecolind.2012.09.018.

Meusnier I., Singer G.A.C., Landry J.-F., Hickey D.A., Hebert P.D.N., Hajibabaei M. A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*. 2008;9(1):214. DOI 10.1186/1471-2164-9-214.

Moss B., Kosten S., Meerhoff M., Battarbee R.W., Jeppesen E., Mazzeo N., Havens K., Lacerot G., Liu Z., de Meester L., Paerl H., Scheffer M. Allied attack: climate change and eutrophication. *Inland Waters*. 2011;1(2):101-105. DOI 10.5268/IW-1.2.359.

Nalepa T.F., Fanslow D.L., Lang G.A. Transformation of the offshore benthic community in Lake Michigan: recent shift from the native amphipod *Diporeia* spp. to the invasive mussel *Dreissena rostriformis bugensis*. *Freshw. Biol.* 2009;54(3):466-479. DOI 10.1111/j.1365-2427.2008.02123.x.

Odum Yu. Basic Ecology. Philadelphia–New York–Chicago–San Francisco–Montreal–Toronto–London–Sydney–Tokyo–Mexico City–Rio de Janeiro–Madrid: Saunders College Publ., 1983. (Russ. ed.: Odum Yu. Ekologiya. Moscow, 1986)

O'Reilly C.M., Alin S.R., Plisnier P.-D., Cohen A.S., McKee B.A. Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa. *Nature*. 2003;424(6950):766-768. DOI 10.1038/nature01833.

Porazinska D.L., Giblin-Davis R.M., Faller L., Farmerie W., Kanzaki N., Morris K., Powers T.O., Tucker A.E., Sung W.A.Y., Thomas W.K. Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Mol. Ecol. Resour.* 2009;9(6):1439-1450. DOI 10.1111/j.1755-0998.2009.02611.x.

Pudovkina T.A., Sitnikova T.Y., Matveyev A.N., Shcherbakov D.Y. Kindred relations of Baikal polychaete of the *Manayunkia* genus [Polychaeta: Sedentaria: Sabellidae] according to CO1 and settlement history analysis. *Russ. J. Genet. Appl. Res*. 2016;6(2):129-137. DOI 10.1134/S207905971602009X.

Rezende R.S., Santos A.M., Henke-Oliveira C., Gonçalves J.F., Jr. Effects of spatial and environmental factors on benthic a macroinvertebrate community. *Zoologia.* 2014;31(5):426-434. DOI 10.1590/S1984-46702014005000001.

Semernoy V.P. Oligochaeta of Lake Baikal. Novosibirsk: Nauka Publ., 2004. (in Russian)

Timoshkin O.A. Biodiversity of Baikal fauna: state-of-the-art (Preliminary analysis). In: New Scope on Boreal Ecosystems in East Siberia. Proc. of the Intern. Workshop, Kyoto, Japan, 23–25 Nov. 1994. Novosibirsk: Russ. Acad. Sci. Publ. Siberian Branch, 1997;35-76.

van den Berg M.S., Coops H., Noordhuis R., van Schie J., Simons J. Macroinvertebrate communities in relation to submerged vegetation in two Chara-dominated lakes. *Hydrobiologia*. 1997;342:143-150. DOI 10.1023/A:1017094013491.

Worrall T.P., Dunbar M.J., Extence C.A., Laize C.L.R., Monk W.A., Wood P.J. The identification of hydrological indices for the characterization of macroinvertebrate community response to flow regime variability. *Hydrol. Sci. J.* 2014;59(3-4):645-658. DOI 10.1080/02626667.2013.825722.

Yu D.W., Ji Y., Emerson B.C., Wang X., Ye C., Yang C., Ding Z. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol. Evol.* 2012;3(4):613-623. DOI 10.1111/j.2041-210X.2012.00198.x.

ORCID ID

L.S. Kravtsova orcid.org/0000-0003-0862-4726
T.E. Peretolchina orcid.org/0000-0002-2950-9762
T.I. Triboy orcid.org/0000-0002-4830-0155
I.A. Nebesnykh orcid.org/0000-0001-5961-0514
A.E. Tupikin orcid.org/0000-0002-8194-0322
M.R. Kabilov orcid.org/0000-0003-2777-0833

Acknowledgements. The work was supported by the project of the Russian Foundation for Basic Research No. 19-05-00398_a. Collection and analysis of samples was partially carried out within the state-funded project conducted by LIN SB RAS No. 121032300196-8; bioinformatics analysis was partially supported by the state-funded project conducted by LIN SB RAS No. 121031300042-1.

Conflict of interest. The authors declare no conflict of interest.

Received February 20, 2023. Revised May 16, 2023. Accepted May 16, 2023.

ECOLOGICAL GENETICS
Original article

Вавиловский журнал генетики и селекции. 2023;27(6):703-711
DOI 10.18699/VJGB-23-81

Soil Alveolata diversity in the undisturbed steppe and wheat agrocenoses under different tillage

N.B. Naumova[1] ✉, P.A. Barsukov[1], O.A. Baturina[2], O.A. Rusalimova[1], M.R. Kabilov[2]

[1] Institute of Soil Science and Agrochemistry of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ naumova@issa-siberia.ru

Abstract. Microeukaryotes are vital for maintaining soil quality and ecosystem functioning, however, their communities are less studied than bacterial and fungal ones, especially by high throughput sequencing techniques. Alveolates are important members of soil microbial communities, being consumers and/or prey for other microorganisms. We studied alveolate diversity in soil under the undisturbed steppe (US) and cropped for wheat using two tillage practices (conventional, CT, and no-till, NT) by amplifying the ITS2 marker with ITS3_KYO2/ITS4 primers and sequencing amplicons using Illumina MiSeq. A total of 198 Alveolata OTUs were identified, with 158 OTUs attributed to the Ciliophora phylum, containing five classes: Litostomatea, Spirotrichea and Oligohymenophorea, Nassophorea and Phyllopharyngea. Litostomatea and Phyllopharyngea were more abundant in US as compared with CT and NT. The observed OTU richness was higher in US than in CT and NT. The β-biodiversity of soil ciliates also very distinctly differentiated the US field from CT and NT. In the US, Nassophorea and Spirotrichea correlated positively with sand and negatively with clay, silt and SOM contents. This is the first report about soil ciliates diversity in Siberia as assessed by metabarcoding technique. The revealed clear effect of land use on the relative abundance of some taxa and a lack of tillage effect suggest the importance of the quantity and quality of plant material input for shaping the prey for ciliates. The ITS-metabarcoding technique was used for the first time in the research of ciliates diversity; further studies, embracing diverse aspects of soil ciliates by combining -omics methodology with the traditional one, are needed to get a better insight on the ecological roles of the main ciliate taxa in the complex soil system.
Key words: ITS region; ciliates; Chernozem; conventional tillage; no tillage.

For citation: Naumova N.B., Barsukov P.A., Baturina O.A., Rusalimova O.A., Kabilov M.R. Soil Alveolata diversity in the undisturbed steppe and wheat agrocenoses under different tillage. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):703-711. DOI 10.18699/VJGB-23-81

Разнообразие почвенных Alveolata в ненарушенной степи и агроценозах пшеницы при разной обработке почвы

Н.Б. Наумова[1] ✉, П.А. Барсуков[1], О.А. Батурина[2], О.А. Русалимова[1], М.Р. Кабилов[2]

[1] Институт почвоведения и агрохимии Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ naumova@issa-siberia.ru

Аннотация. Микроскопические эукариоты крайне важны для обеспечения качества почвы и функционирования экосистем. Однако сообщества почвенных микроэукариот менее изучены по сравнению с сообществами бактерий и грибов, особенно с применением методов высокопроизводительного секвенирования. Значимыми компонентами почвенных микробных сообществ являются альвеоляты, участвующие в ключевых процессах почвенных экосистем (разложение органического вещества, трансформация питательных элементов и др.). Цель работы заключалась в изучении разнообразия альвеолят в почве под ненарушенной степной растительностью и при возделывании пшеницы с помощью двух методов обработки почвы (традиционная вспашка и нулевая обработка) путем амплификации маркера ITS2 с праймерами ITS3_KYO2/ITS4 и последующего секвенирования ампликонов (Illumina MySeq). Всего идентифицировано 198 OTE альвеолят, из которых 158 относились к типу Ciliophora и пяти его классам: Litostomatea, Spirotrichea, Oligohymenophorea, Nassophorea и Phyllopharyngea. Litostomatea и Phyllopharyngea оказались более обильны в почве под ненарушенной степной растительностью по сравненению с почвой под пшеницей обоих вариантов обработки. Богатство OTE в верхнем слое ненарушенной почвы под степной растительностью было также заметно выше, чем в обоих вариатах возделываемых полей, которые четко отличались от степи и по β-биоразнообразию. Nassophorea и Spirotrichea положительно коррелировали с содержанием песка в ненарушенной и пахотной почве; в последней к ним присоединились Litostomatea. Данная работа представляет собой первое исследование разнообразия почвенных альвеолят с применением метода метабаркодирования. Выявленное воздействие землепользования на относительное обилие некоторых таксонов наряду с отсутствием влияния

обработки почвы свидетельствует о важном значении количества и качества поступающего в почву растительного материала для формирования сообществ микроорганизмов, поедаемых инфузориями. Дальнейшие исследования с применением методолгии -омик и традиционных способов необходимы для лучшего понимания экологической роли инфузорий в частности и альвеолят в целом в сложной почвенной системе.
Ключевые слова: ITS район; инфузории; чернозем; традиционная вспашка; нулевая обработка.

Introduction

As it is currently argued, "bridging the gaps between biodiversity science and agricultural practices is crucial to meet food security in the Anthropocene" (Cappelli et al., 2022. P. 674). Therefore, belowground biodiversity is an ultimately important part/mediator of such efforts. Eukaryotic microorganisms, such as fungi, alveolates, metazoans, algae and other organisms with ≤ 5000 μm^3 of body volume (Coleman, 1985) are important players in biotic interactions in soil, and, as such, involved in key ecosystem processes: organic matter transformation, nutrient cycling, etc. (Bardgett, Putten, 2014). Protozoa were shown to benefit plant growth (Bonkowski, 2004), for instance, by improving N mineralization from soil organic matter via stimulating bacterial biomass turnover (Kuikman et al., 1990). Thus, microeukaryotes presence is vital for maintaining soil quality and ecosystem functioning and sustainability. However, soil microbial eukaryotic communities are much less studied as compared with bacterial and fungal ones, and especially by the high throughput sequencing techniques. Alveolates are important members of soil microbial assemblages, where they serve as consumers or prey for other microorganisms. The abundance and taxonomic diversity of alveolates used to be studied by culturing (by the so-called most probable numbers technique) and microscopy. Currently microscopy is the main methodology for the enumeration of alveolates (Adl et al., 2008), but species identification, requiring a complicated staining protocol (Acosta-Mercado, Lynn, 2003), is rather laborious and sometimes not definitive. Therefore, metagenomic approach and state-of-the-art high throughput sequencing has greatly extended the methodology for assessing the biodiversity of alveolates in soil.

In agricultural ecosystems soil and its residential biota is strongly affected by all aspects of production technologies, such as tillage, fertilization, crops, pesticides and others. Over the last decades the possibility to reduce damage to soil by minimizing tillage has gained much attention from both researchers and practitioners. Although there are many reports about bacterial and fungal biodiversity estimated metagenomically under minimal and/or no tillage, alveolates have remained relatively understudied (Ritter et al., 2021). While assessing the ITS2 region DNA sequence reads diversity, using the ITS3_KYO2/ITS4 primer set (Liu K. et al., 2012), under different tillage practices in the chernozem in the south of West Siberia (Naumova et al., 2022), we found that those fungal primers also amplified DNA belonging to other domains, specifically Alveolata, Amoebozoa, Heterolobosea, Metazoa, Rhizaria and Eukaryota kingdoms of uncertain taxonomic attribution. All those reads were discarded for the mycobiome analysis (Naumova et al., 2022); however, a substantial number of alveolate operational taxonomic units (OTUs), with their rarefaction curves reaching plateau with increasing number of sequence reads, convinced us to proceed with analyzing alveolate ITS-based diversity.

Materials and methods

Experimental site and conditions. The field trial was described earlier (Naumova et al., 2022) (https://www.mdpi.com/2075-1729/12/8/1169). Briefly, the study area is the forest-steppe zone (54°4′6″ N, 79°36′3″ E) with a sharply continental climate[1], the mean monthly temperature in the area of experimental site location in October is 3.5 °C with Luvic Endocalcic Chernozem (Siltic) (World Reference Base for Soil Resources…, 2015) as the widely spread and agriculturally significant soil of the region.

Experimental setup. The field trail was described earlier as well (Naumova et al., 2022). Briefly, it was started in 2009 on the area of 40 ha when a portion of the conventionally tilled soil (CT, mouldboard ploughing in the fall and disking in the spring) was subjected to the no-till technology (NT); both plots were getting the same rates of herbicides and fertilizers simultaneously.

The wheat grain yield, harvested at the beginning of September 2021, reached 4.8 t ha^{-1} in the NT field and 4.1 t ha^{-1} in the CT field. An undisturbed site (Un), located near the experimental field and covered by a true bunchgrass steppe (with *Stipa capillata*, *Festuca valesiaca*, some *Poa* spp. and *Puccinellia* sp.), was used to get the data about the zonal soil bacteriobiome as a reference.

Soil sampling and chemical analyses. Soil was sampled in October 2021 from the 0–5 and 5–15 cm layers in five individual replicates from each layer. In total, 30 soil samples were collected and chemically analyzed as described before (Naumova et al., 2022): soil pH ranged 6.3–6.8, total soil carbon content ranged 3.6–4.2 %, and total soil nitrogen content was 0.29–0.37 %.

DNA extraction, amplification and sequencing. Total DNA was extracted with the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The bead-beating was performed using TissueLyser II (Qiagen, Hilden, Germany) for 10 min at 30 Hz. Agarose gel electrophoresis was used to assess the quality of the extracted DNA; additional DNA purification was not necessary.

The ITS2 gene marker was amplified with the primer pairs ITS3_KYO2/ITS4, combined with Illumina adapter sequences (Fadrosh et al., 2014). PCR amplification was performed as described earlier (Kryukov et al., 2020). A total of 200 ng PCR product from each sample was pooled together and purified using the MinElute Gel Extraction Kit (Qiagen, Hilden, Germany). The obtained amplicon libraries were sequenced with 2x300 bp paired-ends reagents on MiSeq (Illumina, CA, USA) in the SB RAS Genomics Core Facility (ICBFM SB RAS, Novosibirsk, Russia). The read data reported in this study were submitted to the NCBI Short Read Archive under bioproject accession number PRJNA845814.

[1] Hydrometcenter of Russia. Available online at: https://meteoinfo.ru/en/climate/monthly-climate-means-for-towns-of-russia-temperature-and-precipitation (accessed on March 27, 2022).

Bioinformatic analysis. Raw sequences were analyzed with the UPARSE pipeline (Edgar, 2013) using Usearch v.11.0.667. The UPARSE pipeline included merging paired reads; read quality filtering (-fastq_maxee_rate 0.005); length trimming (remove less 100 nt); merging identical reads (dereplication); discarding singleton reads; removing chimeras and OTU clustering using the UPARSE-OTU algorithm. The OTU sequences were assigned a taxonomy using SINTAX (Edgar, 2013) and ITS UNITE USEARCH/UTAX v.8.3 (Abarenkov et al., 2021) as a reference. Taxonomic structure of sequences thus obtained was estimated by the ratio of the number of taxon-specific sequence reads (with non-fungal removed from the data matrix) to the total number of sequence reads, i. e. by the relative abundance of taxa, expressed as percentage.

The OTUs datasets were analyzed by individual rarefaction with the help of the PAST software (Hammer et al., 2001): the number of alveolate OTUs detected, reaching plateau with increasing number of sequences, showed that the sampling effort was close to saturation for all samples, thus being enough for comparing biodiversity (Hughes, Hellmann, 2005).

Statistical analyses. Statistical analyses (descriptive statistics, ANOVA and correlation analyses) were performed by using Statistica v.13.3 (TIBCO Software Inc., Palo Alto, CA, USA). Rarefaction curves, OTUs-based α-diversity indices were calculated and principal coordinates analysis was performed using PAST software. Factor effects and mean differences in *post-hoc* comparisons by Fisher's LSD test were considered statistically significant at the $p \leq 0.05$ level.

Results

Alveolata taxonomic diversity. After quality filtering, chimera and other domains' sequences removal, a total of 198 different Alveolata OTUs were identified at 97 % sequence identity level, with 158 OTUs attributed to the Ciliophora phylum, the rest remaining unclassified below the domain level. The rarefaction curves showed that the sampling effort was enough to compare diversity (Hughes, Hellmann, 2005), as the number of OTUs dependent on the total number of sequence reads reached plateau (Fig. 1).

Five Ciliophora classes were identified: Litostomatea with 32 OTUs being the most OTU-rich one, followed by Spirotrichea and Oligohymenophorea with 25 OTUs each, Nassophorea and Phyllopharyngea being represented by just three OTUs each. Of the total number of Ciliophora OTUs, many (70, or 44 %) remained unassigned to the lower taxonomic levels.

Taxonomic composition and structure in different fields. The relative abundance of the Ciliophora phylum did not differ between the fields and the layers (Table 1), whereas at the class level there were some differences: Litostomatea was much more abundant in the undisturbed soil as compared to both cropped ones, and Phyllopharyngea, albeit being a minor member of the ciliate assemblage, was also markedly increased in the undisturbed soil. At the order level, Sporadotrichida was almost seven times more abundant in the 0–5 cm layer of the undisturbed soil than in the no-till one. Oligohymenophorea_is, an order-level cluster, had almost five times higher abundance in the 5–15 cm layer of the no-till soil as compared with the undisturbed one (Table 1). Haptorida (Litostomatea), Hymenostomatida (Oligohymenophorea) and Cyrtophorida (Phyllopharyngea) displayed much higher abundance in the undisturbed soil.

The number of dominant OTUs slightly exceeded 20 in each field and both layers (Fig. 2). The sets, bulked over both soil layers, comprised 30–43 dominant OTUs, the number of the dominant OTUs being maximal in the undisturbed soil. Three OTUs were common for all fields, with two OTUs not

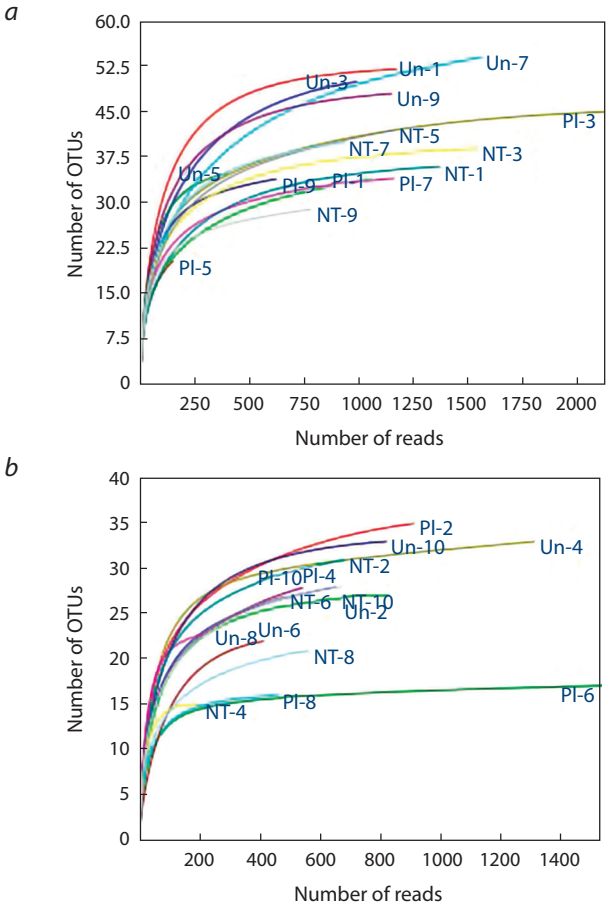

Fig. 1. The rarefaction curves for alveolate OTUs in the 0–5 (*a*) and 5–15 (*b*) cm soil.

Un – undisturbed soil, PI – ploughed and NT – no-till soil; the numbers indicate individual soil replicates from a tillage treatment.

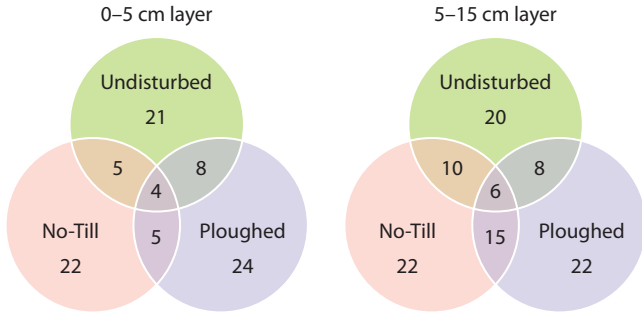

Fig. 2. Venn's diagram of the number of the dominant Alveolata OTUs in soil under different tillage treatment. OTUs were considered dominant if they accounted for ≥ 1 % of the total number of sequence reads.

Table 1. Relative abundance (%, mean) of the dominant Alveolata taxa in Chernozem 0–5 and 5–15 cm layers in the experimental fields in the south of West Siberia

Taxon	Undisturbed		Ploughed		No-till	
	0–5 cm	5–15 cm	0–5 cm	5–15 cm	0–5 cm	5–15 cm
Phylum level						
Ciliophora	86.1	87.9	88.0	88.4	86.8	91.6
un. Alveolata	13.9	12.1	12.0	11.6	13.2	8.4
Class level						
un. Ciliophora	36.3	34.8	33.1	41.0	39.7	50.7
Spirotrichea	25.2	40.1	35.0	34.1	23.7	20.5
Oligohymenophorea	15.9^{ab}	6.1^{a}	17.4^{ab}	12.5^{ab}	19.4^{b}	18.9^{b}
Litostomatea	5.5^{b}	6.8^{b}	1.0^{a}	0.5^{a}	0.8^{a}	1.6^{a}
Nassophorea	2.8^{b}	0.0^{a}	1.3^{ab}	0.2^{ab}	3.1^{b}	0.0^{a}
Phyllopharyngea	0.4^{b}	0.1^{a}	0.1^{a}	0.0^{a}	0.1^{a}	0.0^{a}
Order level						
un. Spirotrichea	12.3	31.1	31.2	27.9	21.8	14.7
Sporadotrichida	12.9^{b}	8.9^{ab}	3.9^{ab}	6.2^{ab}	1.9^{a}	5.8^{ab}
Oligohymenophorea_is	13.8^{ab}	3.9^{a}	17.4^{b}	11.7^{ab}	19.4^{b}	18.9^{b}
Haptorida	4.9^{b}	6.8^{b}	1.0^{a}	0.5^{a}	0.7^{a}	1.5^{a}
Hymenostomatida	2.1^{b}	2.2^{b}	0.0^{a}	0.4^{a}	0.0^{a}	0.0^{a}
Nassulida	2.8^{ab}	0.0^{a}	1.3^{ab}	0.2^{ab}	3.1^{b}	0.0^{a}
Philasterida	0.0	0.0	0.0	0.4	0.00	0.00
un. Litostomatea	0.6^{b}	0.00^{a}	0.00^{a}	0.00^{a}	0.1^{ab}	0.0^{a}
Cyrtophorida	0.30^{b}	0.07^{a}	0.05^{a}	0.00^{a}	0.06^{a}	0.00^{a}
Exogenida	0.12	0.00	0.00	0.00	0.00	0.00

Note. "un." stands for unclassified. Letters in rows indicate that the values are different ($p \leq 0.05$, Fisher's LSD test); the absence of letters after the values in a row indicates that there was no difference.

Table 2. Alpha-biodiversity indices (calculated on the OTU's basis) of Alveolata OTUs assemblages in the Chernozem in the experimental fields in the south of West Siberia

Index	Undisturbed		Ploughed		No-till	
	0–5 cm	5–15 cm	0–5 cm	5–15 cm	0–5 cm	5–15 cm
OTU richness	48^{d}	28^{ab}	34^{bc}	25^{a}	37^{c}	24^{a}
Chao1	50^{d}	29^{ab}	35^{bc}	26^{ab}	41^{cd}	25^{a}
Simpson (S)	0.91	0.81	0.90	0.84	0.91	0.87
Shannon's	3.0^{b}	2.4^{a}	2.7^{ab}	2.3^{a}	2.8^{ab}	2.4^{a}
Evenness	0.42	0.44	0.47	0.44	0.43	0.49
Equitability	0.77	0.72	0.78	0.74	0.76	0.77
Berger–Parker	0.21	0.30	0.19	0.30	0.18	0.25
Dominance (1-S)	0.09	0.19	0.10	0.16	0.09	0.13

Note. Letters in rows indicate that the values are different ($p \leq 0.05$, Fisher's LSD test); the absence of letters after the values in a row indicates that there was no difference.

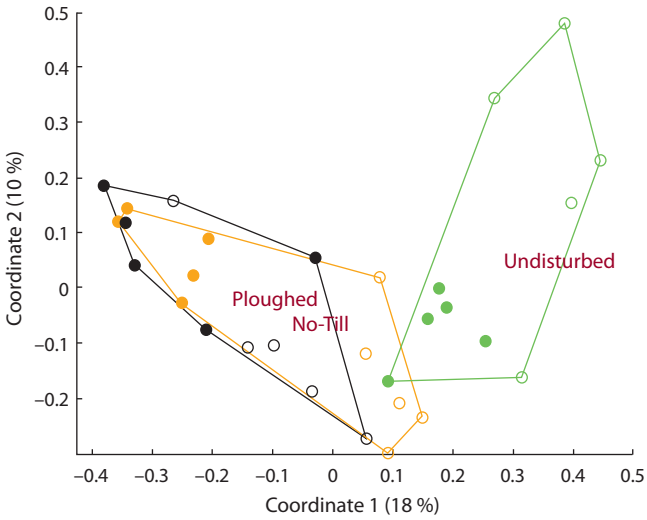

Fig. 3. Principal coordinates analysis of the soil alveolate assemblage composition (OTU level, Bray–Curtis dissimilarity distance) under different soil tillage in the forest-steppe zone in West Siberia: location of samples in the plane of the first two coordinates.

Solid circles indicate samples from the 0–5 cm layer, and open circles indicate samples from the 5–15 cm layer.

identified below the class level (Spirotrichea) and one OTU, below the phylum level (Ciliophora). As for the 0–5 cm layer, the fourth dominant OTU belonged to the Nassulida order of the Nassophorea class; in the 5–15 cm layer, in addition to the three OTU-level clusters common for all samples, there were two unidentified ones below the phylum level and one was identified to the Spirotrichida order of the Spirotrichea class.

Alpha- and beta-biodiversity in different fields. The observed OTU richness was notably higher in the 0–5 cm layer of the undisturbed soil (Table 2) as compared with both cropped ones; the cropped soils did not differ from each other in this index. As for the potential OTU richness, though its estimator (Chao1) in the 0–5 cm layer of the undisturbed soil was markedly higher than in the respective layer of the CT soil, the same was true, albeit to a lesser extent, for the respective layer of the NT soil. Shannon index was much increased in the 0–5 cm layer of the undisturbed soil as compared with its 5–15 cm layer, displaying no difference between the fields.

As for β-diversity, it clearly separated the undisturbed field from the cropped ones, and the latter were not separated from each other (Fig. 3).

Ciliate assemblage composition and soil properties. Correlation of the ciliate classes abundance with soil properties in the 0–5 cm layer showed specific spectra for each field (Fig. 4), albeit with few correlation coefficients being statistically significant.

In the undisturbed soil, Nassophorea and Spirotrichea correlated positively with sand and negatively with clay, silt and soil organic matter contents, being also sensitive to pH. In the same soil, Litostomatea and Oligohymenophorea showed preference for soil organic matter with a wide C/N ratio. The pattern was different in the CT soil, where all major classes, except for Spirotrichea, correlated positively with sand and negatively with clay, silt and soil organic matter ones. A different correlation pattern was revealed in the NT soil, where

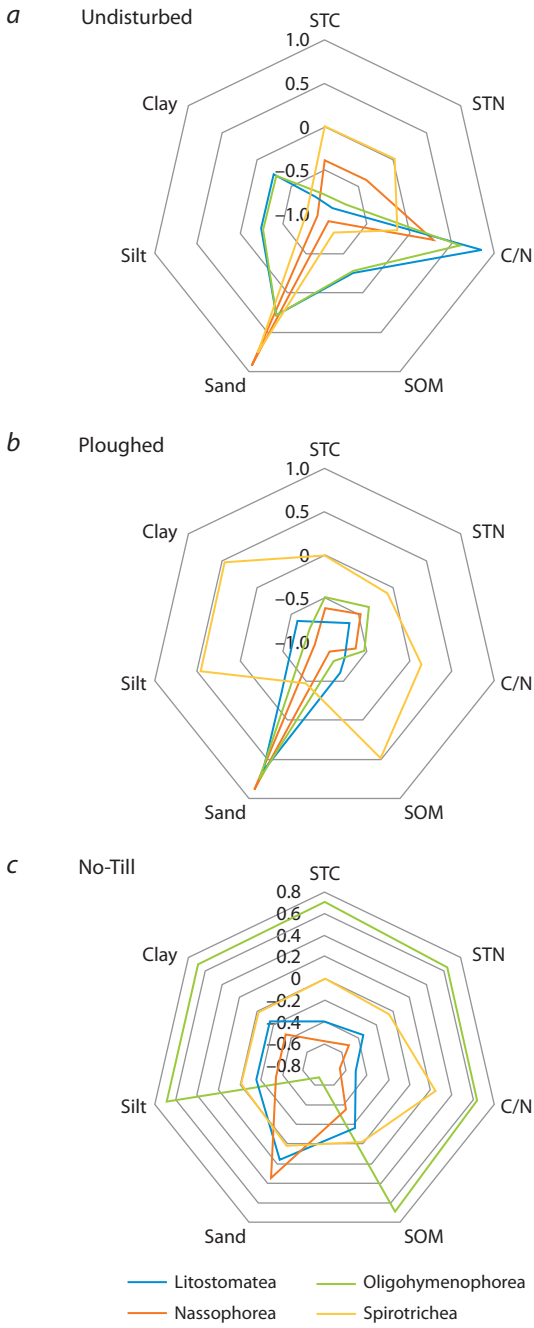

Fig. 4. Correlation coefficients (Pearson) between the relative abundance of the dominant ciliate classes in 0–5 cm layer and soil properties: STC, STN – soil total carbon and nitrogen, SOM – soil organic matter, C/N – the STC/STN ratio in soil; undisturbed soil (*a*), ploughed soil (*b*) and no-tillage soil (*c*).

The coefficients $|r| \geq 0.88$ were statistically significant at $p \leq 0.05$ level.

Oligohymenophorea correlated positively with silt, clay, soil organic matter, soil total carbon and nitrogen and negatively with sand.

Discussion

To assess the diversity of microscopic eukaryotes, most studies used primers to 18S rRNA genes (Ritter et al., 2021), either taxonomically broad (Chaib De Mares et al., 2017) or specific for soil ciliates (Lara et al., 2007; Ting et al., 2015). However,

even with specific primers, non-target taxa sequences are commonly amplified (Pastorelli et al., 2022). The ITS primers we used in this study were designed for fungi (Liu K. et al., 2012), but often these primers did not fail to amplify a plethora of other domains, such as Alveolata, Amoebozoa, Heterolobosea, Metazoa, Rhizaria and Eukaryota kingdoms of uncertain taxonomic attribution. That was precisely what happened in our research: we analysed and reported the mycobiome data (Naumova et al., 2022), but, besides fungal sequences, obtained many sequence clusters belonging to other domains, including Alveolata. It seemed a waste not to attempt an analysis of such sequences; the more so as the number of OTUs reached plateau with the increasing number of sequence reads, tempting us to compare diversity of amplicon sequences. The primers we used were not specifically designed for alveolates, and we are far from claiming that we examined alveolate assemblages in their entirety (though even with specific primers, such claims would be unjustified). However, the fact that diversity estimates of Alveolata OTUs showed agronomically and ecologically meaningful differences, together with our humble hope that our study would "clearly benefit from incorporating more protistology alongside the study of bacteria, fungi and animals" (Geisen et al., 2018), strongly encouraged us to discuss how ciliate diversity relates to the context of this study.

Averaged over all samples in the study, Ciliophora accounted for 88 % of the Alveolata sequence reads. As Ciliophora were shown to prefer arid and semi-arid soil environments (Bates et al., 2013), the phylum ultimate dominance in Alveolata assemblage complies with the climate of the study region, characterized as semi-arid. Less Ciliophora presence, i. e. 45 % of the total number of sequence reads, was reported for the meadow soils in the Alps (Seppey et al., 2020). However, ciliates are relatively more studied and better represented in various databases: in particular, they are highly overrepresented in molecular surveys because of their shorter SSU rRNA sequences that ease amplification, and the presence of extremely high SSU rRNA gene copy numbers (Gong et al., 2013), ranging from 10^3 to 10^6 (Wang et al., 2019). All these might have also been a factor contributing to higher Ciliophora presence in our study.

This study explicitly identified five class-level sequence reads clusters, i. e. Spirotrichea, Oligohymenophorea, Litostomatea, Nassophorea and Phyllopharyngea. The first four were the dominant ones, being commonly found in other studies (employing the same or different methodology) as the main members of soil ciliate assemblages (and even estuarine ones (Jiang et al., 2021)). In the meadow soils in the Alps, the Ciliophora phylum was mostly represented by the Spirotrichea, Oligohymenophorea, Litostomatea and Colpodea classes (Seppey et al., 2020). In soil ciliate community at the Baiyun Mountain in China, the most species-rich classes were Spirotrichea, Colpodea, Litostomatea, Oligohymenophorea, Nassophorea, Armophorea and Phyllopharyngea, as determined by microscopy (Li et al., 2010); exactly the same class-level composition (by DGGE + sequencing) was reported for the oil palm plantation in Malaysia (Ting et al., 2012). Another study reported the ultimate (55 %) dominance of colpodids (by classical methodology) (Bamforth, 2001) in a range of different soils; a notable presence of the group (found by metatransciptomics) was also reported (Geisen et al., 2015). Another very recent study of ciliates in Castanozems in the north-west of China identified nine classes of soil ciliates (Liu H. et al., 2022), with Spirotrichea and Litostomatea being the most species-rich and Colpodea ranking third. Unlike all those studies, here, we did not find any Colpodea. This notable and surprising discrepancy concerning the Colpodea presence may be explained by differences in methodology, primers used, ecosystems, soil, as well as weather conditions preceding soil sampling. According to BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), the primers used for this study had poor homology for Colpodea. As for the weather conditions, they could have contributed somewhat as well: we collected soil samples at the end of the growing season with air temperatures dropping to negatives at night, but in the Alps soil samples were collected in July, i. e. at the height of the growing season, whereas at the Baiyun Mountain they were collected seasonally throughout a year. Also, as Colpodea are well known as *r*-strategists (Lüftenegger et al., 1985), e. g. thriving in unstable environments, they could hardly display a notable presence after gradual change of environmental conditions at the very end of the growing season, more than a month after harvest in the cropped fields and no disturbance in all three fields. Although there exists a discrepancy between the morphological and molecular databases used in the molecular barcoding of protists (Venter et al., 2018), the factor could have hardly contributed to the absence of such conspicuous and well-studied taxon as Colpodea in our study. Therefore, the primers are the primary culprits, and this difference between the ciliate classes might be worth looking into.

Our finding that the Litostomatea class was markedly increased in the undisturbed soil, as compared with the both cropped ones, suggests that in the undisturbed soil a) there was much more prey available for these free-living predators of other protists or microscopic animals (Vd'ačný et al., 2012), or b) these ciliates or their prey were susceptible to the agronomic practices in the cropped fields, as protists were shown to be the most susceptible soil microbiome component to the application of nitrogen fertilizers (Zhao Z. et al., 2019). Anyway, the undisturbed soil environment in both layers apparently benefited Litostomatea, perhaps indirectly, via different chemical composition/stoichiometry of the versatile plant litter: nutrient characteristics of the latter were shown to be important ecological factors that affect protozoan community diversity (Jia et al., 2021). And our finding of positive Litostomatea correlation with the wider C/N ratio of soil indirectly corroborates this.

The Spirotrichea representatives are commonly found in various soil environments, ranging from the high Arctic deserts (Choe et al., 2021) to meadows in the Alps (Seppey et al., 2020). Our results showing that it was the most abundant class in all soil samples from undisturbed and cropped fields (averaging 30 %) imply increased availability of their various prey, from bacteria to other ciliates (Subphylum 2. INTRAMACRONUCLEATA…, 2010), stimulated by dead phytomass abundance at the end of the growing season. The Spirotrichea diversity was reported to be positively correlated with physicochemical parameters such as interstitial water, total organic carbon, nitrogen and phosphorous content (Abraham et al., 2019). However, within the context of our study,

i. e. small gradients of physicochemical properties, we did not reveal such a correlation at the wheat-cropped sites, but at the undisturbed site, the relative abundance of the Spirotrichea-specific sequences was found to correlate positively with sand content in soil. Such correlation suggests sensitivity of these ciliates to soil pore space and aeration, or both.

The fact that Oligohymenophorea had a substantial presence in all soil samples in our study (15 % on average) agrees with the notion that the class is a common member of soil ciliate assemblages (Zhao F. et al., 2013; Tribun et al., 2022). The finding that the class did not show any tillage-related differences (in the top layer) may suggest a broad spectrum of the class niches within the context of the study. As for the 5–15 cm layer, the Oligohymenophorea relative abundance increased from the undisturbed site to the ploughed and no-till ones: the results suggest the beneficial effect of some environmental property (unaccounted in the study) in the no-till soil.

As for the Nassophorea representatives, in this study, the class presence seemed to benefit from the unploughed soil environment similarly in both the undisturbed and no-tillage 0–5 cm layers, which indicates mostly the effect of soil properties, rather than plant species and phytomass. Although Phyllopharyngea demonstrated layer-related differential abundance, as the class was the rare member of the soil ciliate community at all sites, we would not attempt to draw any ecological inference from the finding.

It is noteworthy that in both soil layers only few OTUs were common for all three fields, belonging to the Spirotrichea class and not being explicitly attributed to any of its orders.

Our finding that ciliate relative abundance was in most cases positively correlated with sand content in the soil indicates the importance of relatively bigger pore space needed for these microscopic eukaryotes as typically they are longer than 50 μm in body length (Lynn, 2017).

Temperature is an important factor controlling the ciliate community (Oshima et al., 2020) and affects the structure and functions of the soil microbial food web. For instance, ciliates can be destroyed by freezing temperatures, especially at soil moisture content exceeding 30 % (Müller et al., 2010). Some researchers suggested that "internally governed encystment may be an essential adaptation to an unpredictable environment in which individual protozoa cannot sense when the soil will dry out and will survive desiccation only if they have encysted in time" (Ekelund et al., 2002. P. 1096), or, extending the statement, when the soil will freeze or experience other adverse condition. In general, soil helps to preserve ciliate cysts in a viable state. Since the soil for our study was sampled at the end of October when freezing temperatures, at least at nights, are common, it is most likely that the diversity profile reflects the diversity of viable but non-active organisms.

It should be noted that relationships between ribotypic and phenotypic traits of protists across their life cycle stages remain largely unknown: recently, encystment and temperature were shown to influence intraindividual sequence polymorphisms of rDNA and rRNA (Zou et al., 2021). Thus, the rDNA copy number may affect the composition and structure of soil ciliate assemblages.

Nowadays, it is commonplace to reiterate that "conventional agricultural production systems… reduce soil biodiversity" (Harkes et al., 2019); and our finding of reduced ciliate OTUs' richness, both observed and potential, in the top 5-cm layer of conventionally ploughed soil as compared with the undisturbed one, agrees with the statement. However, the same was true for the no-till soil as far as the observed OTUs' richness is concerned. But the potential richness, i. e. the Chao1 index, albeit being somewhat lower still, did not differ statistically from the value in the undisturbed soil, very likely indicating the ongoing, albeit slowly, process of ciliate assemblage diversification due to the no-till treatment. As for the Shannon index, it showed no difference between the soil tillage managements. The same pattern, i. e. reduced OTUs' richness under conventional tillage as compared with the undisturbed soil, and no difference in the Shannon index, was found by us in the mycobiome of the same soil samples (Naumova et al., 2022). This finding implies that the Shannon index, calculated on the basis of sequence reads, in the case of eukaryotic microorganisms, at least such as fungi and alveolates, cannot adequately reflect biodiversity changes.

With ITS primers, this study recorded 158 Ciliophora OTUs in the soil samples collected from adjacent fields. Recently, a comprehensive study of ciliated protozoans in soils and fresh water bodies of the Russian Far East, performed by employing traditional microscopic techniques to detect and identify ciliates, found 307 species (Tribun et al., 2022), which, bearing in mind the number, biotope diversity and area surveyed (in total, about 900,000 km^2) did not strike us as seriously exceeding the species richness in our study. The core of the ciliate communities in both studies belonged to the classes Oligohymenophorea, Spirotrichea and Litostomatea, together accounting for 65 % of species richness in the Far East study and 52 % of species richness and 41–50 % of the total number of sequence reads in our study. Another very recent study in the north-west of China, also employing traditional methodology to identify and enumerate ciliates, found 114 species of ciliates among four sampling sites, varying in vegetation and land use (Liu H. et al., 2022). Thus, we can safely conclude that soil ciliates diversity data, obtained here by sequencing amplicons of ITS2 region of rRNA genes, encompassed a significant portion of true ciliate diversity in soil, providing ecologically relevant and meaningful assemblage profiles in the context of our study.

Conclusion

This is the first report about soil ciliates diversity, as assessed by metagenomic technique, in Siberia, and specifically in Chernozem under different land use and tillage practices (undisturbed steppe vs. cropped for wheat by conventional or no tillage). We found a clear effect of land use on the relative abundance of some taxa at the order level, but did not find any effect of the tillage treatments: this strongly suggests the importance of primary producers, i. e. the quantity and quality of plant material input in soil, in shaping the prey available for ciliates. Soil ciliate β-diversity differentiated the undisturbed field from the cropped ones very distinctly as well. Further multifaceted studies, focusing on many aspects of soil ciliates by combining -omics methodology with the traditional one, are needed to get a better insight on the ecological roles of the main ciliate taxa in the complex soil system.

References

Abarenkov K., Zirk A., Piirmann T., Pöhönen R., Ivanov F., Nilsson R.H., Kõljalg U. UNITE USEARCH/UTAX release for Fungi. UNITE Community, 2021. DOI 10.15156/BIO/1280276.

Abraham J.S., Sripoorna S., Dagar J., Jangra S., Kumar A., Yadav K., Singh S., Goyal A., Maurya S., Gambhir G., Toteja R., Gupta R., Singh D.K., El-Serehy H.A., Al-Misned F.A., Al-Farraj S.A., Al-Rasheid K.A., Maodaa S.A., Makhija S. Soil ciliates of the Indian Delhi Region: Their community characteristics with emphasis on their ecological implications as sensitive bio-indicators for soil quality. *Saudi J. Biol. Sci.* 2019;26(6):1305-1313. DOI 10.1016/j.sjbs.2019.04.013.

Acosta-Mercado D., Lynn D.H. The edaphic quantitative protargol stain: A sampling protocol for assessing soil ciliate abundance and diversity. *J. Microbiol. Meth*. 2003;53(3):365-375. DOI 10.1016/S0167-7012(03)00042-3.

Adl S.M., Acosta-Mercado D., Lynn D.H. Protozoa. In: Carter M.R., Gregorich E.G. (Eds.). Soil sampling and methods of analysis. 2nd ed. Boca Raton, USA: CRC Press, 2008;77-91.

Bamforth S. Proportions of active ciliate taxa in soils. *Biol. Fertil. Soils.* 2001;33:197-203. DOI 10.1007/s003740000308.

Bardgett R.D., van der Putten W.H. Belowground biodiversity and ecosystem functioning. *Nature.* 2014;515(7528):505-511. DOI 10.1038/nature13855.

Bates S.T., Clemente J.C., Flores G.E., Walters W.A., Parfrey L.W., Knight R., Fierer N. Global biogeography of highly diverse protistan communities in soil. *ISME J.* 2013;7(3):652-659. DOI 10.1038/ismej.2012.147.

Bonkowski M. Protozoa and plant growth: the microbial loop in soil revisited. *New Phytol*. 2004;162(3):617-631. DOI 10.1111/j.1469-8137.2004.01066.x.

Cappelli S.L., Domeignoz-Horta L.A., Loaiza V., Laine A.L. Plant biodiversity promotes sustainable agriculture directly and via belowground effects. *Trends Plant Sci.* 2022;27(7):674-687. DOI 10.1016/j.tplants.2022.02.003.

Chaib De Mares M., Sipkema D., Huang S., Bunk B., Overmann J., van Elsas J.D. Host specificity for bacterial archaeal and fungal communities determined for high- and low-microbial abundance sponge species in two genera. *Front. Microbiol*. 2017;8:2560. DOI 10.3389/fmicb.2017.02560.

Choe Y-H., Kim M., Lee Y.K. Distinct microbial communities in adjacent rock and soil substrates on a high arctic polar desert. *Front. Microbiol*. 2021;11:607396. DOI 10.3389/fmicb.2020.607396.

Coleman D. Through a ped darkly – an ecological assessment of root soil-microbial-faunal interactions. In: Fitter A.H., Atkinson D., Read D.J., Usher M.B. (Eds). Ecological interactions in the soil: plants microbes and animals. Oxford: Blackwell Science Publication, 1985;1-21.

Edgar R.C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods.* 2013;10(10):996-998. DOI 10.1038/nmeth.2604.

Ekelund F., Frederiksen H.B., Rønn R. Population dynamics of active and total ciliate populations in arable soil amended with wheat. *Appl. Environ. Microbiol*. 2002;68(3):1096-1101. DOI 10.1128/AEM.68.3.1096-1101.2002.

Fadrosh D.W., Ma B., Gajer P., Sengamalay N., Ott S., Brotman R.M., Ravel J. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome.* 2014;2:6. Available at: https://microbiomejournal.biomedcentral.com/articles/10.1186/2049-2618-2-6

Geisen S., Tveit A.T., Clark I.M., Richter A., Svenning M.M., Bonkowski M., Urich T. Metatranscriptomic census of active protists in soils. *ISME J.* 2015;9:2178-2190. DOI 10.1038/ismej.2015.30.

Geisen S., Mitchell E.A.D., Adl S., Bonkowski M., Dunthorn M., Ekelund F., Fernández L.D., Jousset A., Krashevska V., Singer D., Spiegel F.W., Walochnik J., Lara E. Soil protists: a fertile frontier in soil biology research. *FEMS Microbiol. Rev.* 2018;42(3):293-323. DOI 10.1093/femsre/fuy006.

Gong J., Dong J., Liu X., Massana R. Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist.* 2013;164(3):369-379. DOI 10.1016/j.protis.2012.11.006.

Hammer O., Harper D.A.T., Ryan P.D. PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica.* 2001;4:9. Available at: https://palaeo-electronica.org/2001_1/past/past.pdf

Harkes P., Suleiman A.K.A., van den Elsen S.J.J., de Haan J.J., Holterman M., Kuramae E.E., Helder J. Conventional and organic soil management as divergent drivers of resident and active fractions of major soil food web constituents. *Sci. Rep*. 2019;9(1):13521. DOI 10.1038/s41598-019-49854-y.

Hughes J.B., Hellmann J.J. The Application of Rarefaction Techniques to Molecular Inventories of Microbial Diversity. *Methods Enzymol.* 2005;397:292-308. DOI 10.1016/S0076-6879(05)97017-1.

Jia T., Liang X., Guo T., Chai B. Impact of nutrients on protozoa community diversity and structure in litter of two natural grass species in a copper tailings dam China. *Microorganisms.* 2021;9(11):2250. DOI 10.3390/microorganisms9112250.

Jiang C., Liu B., Zhang J., Gu S., Liu Z.,Wang X., Chen K., Xiong J., Lu Y., Miao W. Diversity and seasonality dynamics of ciliate communities in four estuaries of Shenzhen China (South China Sea). *J. Mar. Sci. Eng*. 2021;9(3):260. DOI 10.3390/jmse9030260.

Kryukov V.Y., Kosman E., Tomilova O., Polenogova O., Rotskaya U., Tyurin M., Alikina T., Yaroslavtseva O., Kabilov M., Glupov V. Interplay between fungal infection and bacterial associates in the wax moth *Galleria mellonella* under different temperature conditions. *J. Fungi* (*Basel*). 2020;6(3):170. DOI 10.3390/jof6030170.

Kuikman P.J., Jansen A.G., Veen J.A., Zehnder A.J.B. Protozoan predation and the turnover of soil organic carbon and nitrogen in the presence of plants. *Biol. Fertil Soils*. 1990;10:22-28. DOI 10.1007/bf00336120.

Lara E., Berney C., Harms H., Chatzinotas A. Cultivation-independent analysis reveals a shift in ciliate 18S rRNA gene diversity in a polycyclic aromatic hydrocarbon-polluted soil. *FEMS Microbiol. Ecol*. 2007;62(3):365-373. DOI 10.1111/j.1574-6941.2007.00387.x.

Li J., Li M.-G., Yang J., Ai Y., Xu R.-L. Community characteristics of soil ciliates at Baiyun Mountain, Guangzhou, China. *Zool. Stud*. 2010;49(6):713-723. Available at: https://zoolstud.sinica.edu.tw/Journals/49.6/713.pdf

Liu K.L., Porras-Alfaro A., Kuske C.R., Eichorst S.A., Xie G. Accurate rapid taxonomic classification of fungal large-subunit rRNA genes. *Appl. Environ. Microbiol*. 2012;78(5):1523-1533. DOI 10.1128/AEM.06826-11.

Liu H., Ning Y., Yang Y., Yang H., Wang L., Chen L., Wanma Y., Shen X. Use of ciliate communities for monitoring ecological restoration of grain for the green in north-western China. *Soil Ecol. Lett.* 2022;4:264-275. DOI 10.1007/s42832-021-0105-3.

Lüftenegger G., Foissner W., Adam H. r- and K-selection in soil ciliates: a field and experimental approach. *Oecologia.* 1985;66(4):574-579. DOI 10.1007/BF00379352.

Lynn D.H. Ciliophora. In: Archibald J., Simpson A., Slamovits C. (Eds.). Handbook of the Protists. USA: Springer: Cham, 2017;679-730. DOI 10.1007/978-3-319-28149-0_23.

Müller H., Achilles-Day U.E., Day J.G. Tolerance of the resting cysts of *Colpoda inflata* (*Ciliophora Colpodea*) and *Meseres corlissi* (*Ciliophora Spirotrichea*) to desiccation and freezing. *Eur. J. Protistol*. 2010;46(2):133-142. DOI 10.1016/j.ejop.2009.12.004.

Naumova N., Barsukov P., Baturina O., Rusalimova O., Kabilov M. Soil Mycobiome Diversity under Different Tillage Practices in the South of West Siberia. *Life* (*Basel*). 2022;12(8):1169. DOI 10.3390/life12081169.

Oshima T., Shinohara Y., Asakawa S., Murase J. Susceptibility and resilience of the soil ciliate community to high temperatures. *Soil Sci. Plant Nutr*. 2020;66(6):870-877. DOI 10.1080/00380768.2020.1819148.

Pastorelli R., Cucu M.A., Lagomarsino A., Paletto A., De MeoI. Analysis of Ciliate Community Diversity in Decaying Pinus nigra Logs. *Forests.* 2022;13(5):642. DOI 10.3390/f13050642.

Ritter C.D., Machado A.F., Ribeiro K.F., Dunthorn M. Metabarcoding advances for ecology and biogeography of Neotropical protists: what do we know, where do we go? *Biota Neotropica.* 2021; 21(4):e20211214. DOI 10.1590/1676-0611-BN-2021-1214.

Seppey C.V.W., Broennimann O., Buri A., Yashiro Pinto-Figueroa E., Singer D., Blandenier Q., Mitchell E.A.D., Niculita-Hirzel H., Guisan A., Lara E. Soil protist diversity in the Swiss western Alps is better predicted by topo-climatic than by edaphic variables. *J. Biogeogr*. 2020;47(4):866-878. DOI 10.1111/jbi.13755.

Subphylum 2. INTRAMACRONUCLEATA: Class 1. SPIROTRICHEA – Ubiquitous and Morphologically Complex. In: Lynn D.H. (Ed.) The Ciliated Protozoa. Dordrecht: Springer, 2010;141-173. DOI 10.1007/978-1-4020-8239-9_7.

Ting L.T., King W.S., Hong L.W., Ali S.R.A. Diversity of soil protozoa (ciliates) in oil palm plantation at Sungai Asap Sarawak. In: Proceedings of the Third International Plantation Industry Conference and Exhibition. Kota Kinabalu, Sabah, Malaysia, 2012;5.

Ting L.T., King W.S., Hong L.W., Ali S.R.A. New combination of primer pairs for PCR-DGGE detection of soil ciliates. *Malays. Appl. Biol.* 2015;44:67-72.

Tribun M., Panov A., Nikitina L. Fauna of ciliates (*Alveolata Ciliophora*) of the southern part of the Russian Far East. *Protistology.* 2022; 16(2):109-121. DOI 10.21685/1680-0826-2022-16-2-5.

Venter P.C., Nitsche F., Scherwass A., Arndt H. Discrepancies between molecular and morphological databases of soil ciliates studied for temperate grasslands of Central Europe. *Protist.* 2018;169(4):521-538. DOI 10.1016/j.protis.2018.04.001.

Vd'ačný P., Bourland W.A., Orsi W., Epstein S..S., Foissner W. Genealogical analyses of multiple loci of litostomatean ciliates (*Protista Ciliophora Litostomatea*). *Mol. Phyl. Evol*. 2012;65(2):397-411. DOI 10.1016/j.ympev.2012.06.024.

Wang Y., Wang C., Jiang Y., Katz L.A., Gao F., Yan Y. Further analyses of variation of ribosome DNA copy number and polymorphism in ciliates provide insights relevant to studies of both molecular ecology and phylogeny. *Sci. China Life Sci*. 2019;62(2):203-214. DOI 10.1007/s11427-018-9422-5.

World Reference Base for Soil Resources 2014, update 2015. International soil classification system for naming soils and creating legends for soil maps. IUSS Working Group WRB. Rome: FAO, 2015.

Zhao F., Xu K., Zhang D. Spatio-temporal variations in the molecular diversity of microeukaryotes in particular ciliates in soil of the Yellow River delta China. *J. Eukaryotic Microbiol.* 2013;60(3): 282-290. DOI 10.1111/jeu.12035.

Zhao Z.B., He J.Z., Geisen S., Han L.L., Wang J.T., Shen J.P., Wei W.X., Fang Y.T., Li P.P., Zhang L.M. Protist communities are more sensitive to nitrogen fertilization than other microorganisms in diverse agricultural soils. *Microbiome*. 2019;7(1):33. DOI 10.1186/s40168-019-0647-0.

Zou S., Fu R., Deng H., Zhang Q., Gentekaki E., Gong J. Coupling between ribotypic and phenotypic traits of protists across life cycle stages and temperatures. *Microbiol. Spectr*. 2021;9(3):e0173821. DOI 10.1128/Spectrum.01738-21.

ORCID ID

N.B. Naumova orcid.org/0000-0003-2354-5065
M.R. Kabilov orcid.org/0000-0003-2777-0833

Acknowledgements. This research was funded by the Ministry of Science and Higher Education of the Russian Federation (grant No. 075-15-2021-1085). The authors are very thankful to the private farmer Mr. A.E. Weiss for his permission to sample soils from the experimental fields and to Mrs. Galina A. Bugrovskaya for carrying out soil chemical analyses.

Conflict of interest. The authors declare no conflict of interest.

Data availability statement. The read data reported in this study were submitted to the GenBank under the study accession PRJNA845814.

Received March 29, 2023. Revised August 3, 2023. Accepted August 7, 2023.

Вавиловский журнал генетики и селекции. 2023;27(6):712-718
DOI 10.18699/VJGB-23-82

Original Russian text https://vavilovj-icg.ru/

Smallpox vaccination in a mouse model

S.N. Shchelkunov[1, 2] ✉, A.A. Sergeev[1], S.A. Pyankov[1], K.A. Titova[1], S.N. Yakubitskiy[1]

[1] State Research Center of Virology and Biotechnology "Vector", Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ snshchel@rambler.ru

Abstract. The monkeypox epidemic, which became unusually widespread among humans in 2022, has brought awareness about the necessity of smallpox vaccination of patients in the risk groups. The modern smallpox vaccine variants are introduced either intramuscularly or by skin scarification. Intramuscular vaccination cannot elicit an active immune response, since tissues at the vaccination site are immunologically poor. Skin has evolved into an immunologically important organ in mammals; therefore, intradermal delivery of a vaccine can ensure reliable protective immunity. Historically, vaccine inoculation into scarified skin (the s.s. route) was the first immunization method. However, it does not allow accurate vaccine dosing, and high-dose vaccines need to be used to successfully complete this procedure. Intradermal (i.d.) vaccine injection, especially low-dose one, can be an alternative to the s.s. route. This study aimed to compare the s.s. and i.d. smallpox immunization routes in a mouse model when using prototypic second- and fourth-generation low-dose vaccines (10^4 pfu). Experiments were conducted using BALB/c mice; the LIVP or LIVP-GFP strains of the vaccinia virus (VACV) were administered into the tail skin via the s.s. or i.d. routes. After vaccination (7, 14, 21, 28, 42, and 56 days post inoculation (dpi)), blood samples were collected from the retro-orbital venous sinus; titers of VACV-specific IgM and IgG in the resulting sera were determined by ELISA. Both VACV strains caused more profound antibody production when injected via the i.d. route compared to s.s. inoculation. In order to assess the level of the elicited protective immunity, mice were intranasally infected with a highly lethal dose of the cowpox virus on 62 dpi. The results demonstrated that i.d. injection ensures a stronger protective immunity in mice compared to s.s. inoculation for both VACV variants.
Key words: smallpox; monkeypox; vaccinia virus; vaccination; intradermal injection; skin scarification.

For citation: Shchelkunov S.N., Sergeev A.A., Pyankov S.A., Titova K.A., Yakubitskiy S.N. Smallpox vaccination in a mouse model. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):712-718. DOI 10.18699/VJGB-23-82

Противооспенная вакцинация на модели мышей

С.Н. Щелкунов[1, 2] ✉, А.А. Сергеев[1], С.А. Пьянков[1], К.А. Титова[1], С.Н. Якубицкий[1]

[1] Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ snshchel@rambler.ru

Аннотация. Необычно широко распространившаяся в 2022 г. эпидемия оспы обезьян среди людей привела к заключению о необходимости противооспенной вакцинации пациентов из групп риска. При этом современные варианты противооспенной вакцины вводят либо внутримышечно, либо скарификацией кожи. Внутримышечное введение не обеспечивает активного иммунного ответа, так как ткани, в которые при этом вводится вакцина, являются иммунологически бедными. Кожа эволюционно развилась в иммунологически важный орган млекопитающих, поэтому введение вакцины в дерму кожи может обеспечивать надежный протективный иммунный ответ. Исторически первым способом иммунизации стал метод инокуляции вакцины в скарифицированную кожу (с/к). Однако этот метод не обеспечивает точного дозирования вакцины, для успешного выполнения процедуры нужно использовать вакцину в высокой концентрации. Альтернативой методу с/к может служить процедура внутрикожной (в/к) инъекции вакцины, особенно при использовании ее в низкой концентрации. Целью настоящей работы было сравнение способов внутрикожной противооспенной иммунизации на модели мышей с применением прототипных вакцин второго и четвертого поколений в низкой дозе 10^4 БОЕ. Эксперименты выполняли на мышах линии BALB/c, штаммы LIVP или LIVP-GFP вируса осповакцины (VACV) вводили в кожу хвоста с/к или в/к способами. Через 7, 14, 21, 28, 42 и 56 дней после вакцинации (дпв) у мышей проводили забор проб крови из ретроорбитального венозного синуса и получали сыворотки, в которых методом ИФА определяли титры VACV-специфичных IgM и IgG. Оба штамма VACV обусловливали более выраженную продукцию антител при в/к инъекции по сравнению со с/к инокуляцией. Для проверки уровня развившегося протективного иммунитета на 62-й дпв мышей интраназально инфицировали высоколетальной дозой вируса оспы коров. Полученные результаты показали, что в/к инъекция обеспечивает развитие протективного иммунитета у мышей в значительно большей степени, по сравнению с с/к инокуляцией обоих вариантов VACV.
Ключевые слова: оспа; оспа обезьян; вирус осповакцины; вакцинация; внутрикожная инъекция; скарификация кожи.

Introduction

Smallpox (*lat*. variola) is an especially dangerous infectious disease that has claimed lives of many hundreds of millions of people over the past centuries. During smallpox epidemics, the death toll among the infected people could be as high as 30–40 %. The variola virus (VARV) is an infectious agent of this disease (Fenner et al., 1998).

The VARV was mostly transmitted amongst humans via the airborne or aerosol route during close personal contacts. The incubation period lasting one or two weeks was followed by abrupt onset of fever, headache, and sacral pain. Several days later, rash lesions appeared on the tongue as well as oral and oropharyngeal mucosa; maculopapular rash then developed on the face and hands, subsequently spreading over the entire body and progressing to pustules. By day 10–13 of illness, the pustules reached their maximum size, and then gradually flattened, dried, and evolved into scabs. By day 30–40 of the disease, the scabs fell off to leave reddish spots. These scabs subsequently left typical deep scars known as pockmarks in some body areas, mostly on the face (the pock-pitted face). Hence, smallpox survivors could be easily phenotypically differentiated from people who had not had this disease (Shchelkunov et al., 2005).

It turned out that smallpox survivors were not susceptible to it during the later epidemics. Many centuries ago, this fact apparently gave an idea to Indian and Chinese doctors to develop a procedure that subsequently became known as variolation (variola inoculation). According to this method, the infectious material obtained by rubbing scabs taken from epidemic patients was placed (inoculated) into skin incisions. People infected intradermally typically had a milder form of smallpox compared to the naturally occurring smallpox. After the infectious process, a characteristic scar was formed at the site of VARV inoculation into the skin. This procedure made people resistant to smallpox. However, 0.5–2.0 % of variolated individuals died, so this smallpox protection method has not become common (Fenner et al., 1988).

In the XVIII century, a smallpox-like disease in cattle and horses, which became known as cowpox, was reported in England. This disease was clinically characterized by development of skin rashes on animal bodies, most frequently on the udder and teats. The skin elements underwent typical evolutionary transformation stages (papules to vesicles to pustules); scabs and ulcers were subsequently formed. This infection was easily transmitted to people who had contacted the infected animals. In most cases, cowpox in humans had a mild course and was characterized by isolated topical lesions, mostly on hands and forearms, at skin microtrauma sites. After infection resolution, cicatrices resembling variolation scars were formed at former skin lesion sites. Furthermore, people who recovered from cowpox did not get infected during smallpox epidemics.

Having gained this knowledge, an English physician Ed. Jenner inferred that people can be protected against smallpox by being preliminarily infected with cowpox. Starting with 1796, he conducted several experimental inoculations of the infectious material collected from pustules of cowpox-infected humans into skin incisions (the skin scarification route) in people and, after some time, infected them with smallpox using the variolation procedure. In all the cases, people infected with the cowpox were resistant to smallpox infection. Ed. Jenner called the developed smallpox protection procedure "vaccination" (or vaccine inoculation, the term derived from *Latin* vacca – cow) (Fenner et al., 1988; Esparza et al., 2017).

It is noteworthy that it was not until one century after the invention of the smallpox vaccination method that the kingdom of viruses was discovered. However, it has only recently been found that different vaccinia virus (VACV) strains that have been used for immunization for a long time are closest to the horsepox virus rather than the cowpox virus in terms of their genomic organization (Tulman et al., 2006; Esparza et al., 2017).

Smallpox was completely eradicated by 1977 using mass smallpox vaccination and strict epidemiologic surveillance under the World Health Organization's Global Smallpox Eradication Program (Fenner et al., 1988).

In the overwhelming majority of cases, the VACV was inoculated for smallpox vaccination by the skin scarification (s.s.) route. This procedure is relatively easy to perform but does not allow accurate dosing of the vaccine preparation; therefore, high-dose viral preparation needs to be used to ensure reliable immunization (Fenner et al., 1988; Jacobs et al., 2009; Sanchez-Sampedro et al., 2015).

Intradermal (i.d.) injection of the vaccine preparation can be a modern alternative to the s.s. route. This approach allows accurate vaccine dosing and ensures higher immunization reliability, so the dose of the administered vaccine can be reduced, which is especially important in the case of mass vaccination.

This study aimed to compare the effectiveness of i.d. and s.s. smallpox vaccination with low-dose VACV in a model of BALB/c mice. For correct comparison, the VACV was introduced by both routes within the same region of mouse tail skin for both procedures. The clonal variant of the LIVP strain and the constructed recombinant LIVP-GFP (mutant with respect to viral thymidine kinase), which can be regarded as prototypic second- and fourth-generation smallpox vaccines, respectively, were used as study objects.

Materials and methods

Viruses and cell culture. Clonal variant 14 of the VACV LIVP strain produced by limiting dilution and triple plaque purification using agarose overlay (Yakubitskiy et al., 2015), the mutant LIVP-GFP, with inactivated virus thymidine kinase gene, generated based on it (Petrov et al., 2013), and the cowpox virus (CPXV) strain GRI-90 (Shchelkunov et al., 1998) were used in this study. The viruses were grown and titrated using the African green monkey kidney cells line CV-1 from the collection of the State Research Center of Virology and Biotechnology "Vector", Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing.

Animals. BALB/c mice were procured from the Laboratory Animals Farm of the SRC VB Vector. The experimental animals were fed a standard diet with a sufficient amount of water, in compliance with the veterinary laws and the requirements for the humane care and use of laboratory animals. Animal manipulations were approved by the Bioethics Committee of SRC VB Vector (Protocol No. 02-06.2022 dated June 28, 2022).

Pathogenicity assessment of VACV strains. Three-week-old BALB/c mice weighing 10–12 g (10 animals per group)

were used in the studies to assess the pathogenicity of the VACV LIVP and LIVP-GFP strains upon intranasal (i.n.) infection. Mice preliminarily subjected to inhaled anesthesia with diethyl ester received 50 μL of virus-containing fluid at a dose of 10^7 plaque-forming units (pfu) or normal saline inoculated into the nasal cavity. The animals were followed up for 14 days; clinical signs of infection and animal deaths were documented.

The score grading system for assessing the revealed symptoms was as follows: 0 – no signs of the disease; 1 – slightly ruffled hair coat; 2 – significantly ruffled hair coat; 3 – significantly ruffled hair coat and hunched posture or conjunctivitis; 4 – labored breathing or remaining immobile; and 5 – death.

Each mouse was weighed every two days. The arithmetic mean mouse body weight in each group at each time point was calculated and expressed as a percentage from the baseline value. Data diffusion with respect to the mean value was presented as standard deviation of the mean and also expressed as a percentage.

Immunization of mice. Female BALB/c mice starting with age of 6–7 weeks (body weight, 16–19 g) were immunized by intradermal (i.d.) injection or skin scarification (s.s.) using VACV LIVP or LIVP-GFP at a dose of 10^4 pfu.

When performing an i.d. injection or s.s. inoculation, the inoculation site (the dorsal side of the tail, approximately 1 cm from its base) was pretreated with 70 % ethanol. In the case of i.d. injection, 20 μL of viral material (10^4 pfu) or normal saline (control group) was inoculated according to the procedure described earlier (Shchelkunov et al., 2022a). For s.s. immunization, 10 skin incisions were made using a 26G needle (0.45 × 16 mm) within the uppermost layer of epidermis. Viral material (10^4 pfu) or normal saline (control group) (5 μL) was immediately applied onto the damaged skin and allowed to be absorbed into the skin.

On days 7, 14, 21, 28, 42, and 56 post inoculation (dpi) with the LIVP or LIVP-GFP viruses, blood samples were collected from the retro-orbital venous sinus in mice (six animals from each group) according to the procedure described previously (Shchelkunov et al., 2022a).

Serum preparations were obtained from individual blood samples of mice by centrifuging blood cells. Serum samples obtained from mouse blood were stored at –20 °C.

Enzyme-linked immunosorbent assay of serum samples. Enzyme-linked immunosorbent assay of individual serum samples of mice was carried out according to the procedure described previously (Shchelkunov et al., 2020). Purified VACV LIVP preparation was used as an antigen. The geometric means of log reciprocal titer of VACV-specific IgM and IgG were determined for the study groups, and the confidence intervals were calculated for the 95 % matching between each sample and the total population.

Assessment of protectivity in immunized mice. On 62 dpi, the groups of animals immunized with the LIVP or LIVP-GFP viruses and control animals were i.n. inoculated with CPXV GRI-90 at a dose of 46 LD_{50} (9.4×10^5 pfu/mouse). The animals were followed up for 14 days, and their deaths were documented.

The data were obtained for groups consisting of six animals immunized, either i.d. or by s.s., with VACV LIVP or LIVP-GFP, as well as groups of non-immunized mice and non-infected animals (the negative control) or animals infected with CPXV GRI-90 (the positive control).

Statistical analysis. Statistical analysis and comparison of the results was carried out with standard methods using the Statistica 13.0 software package (StatSoft Inc., 1984–2001). The 50 % lethal dose (LD_{50}) was calculated using the Spearman–Karber method according to the number of animals that had died (Sachs, 1972). The p-value < 0.05 was considered statistically significant.

Results

Comparison of the pathogenic properties of LIVP and LIVP-GFP strains intranasally inoculated to mice

Three-week-old BALB/c mice were used to assess the pathogenicity of the VACV LIVP and LIVP-GFP strains in this study. The mice were i.n. inoculated with the viruses at a dose of 10^7 pfu. For VACV LIVP, pronounced clinical manifestations of the infection were observed starting with 4 dpi; their maximum intensity was detected on 8 dpi; and the animals recovered after 10 dpi (Fig. 1, *b*). The disease was accompanied by significant body weight loss of mice (see Fig. 1, *a*).

Under the same conditions, the LIVP-GFP virus with the mutant thymidine kinase gene caused minimal clinical manifestations of infection 6–8 dpi (see Fig. 1, *b*) and an insignificant body weight loss in infected animals compared to those in the control group (see Fig. 1, *a*). I.n. inoculation of mice with the LIVP strain resulted in the death of 50 % animals, whereas all the animals survived the inoculation with the LIVP-GFP strain (Fig. 2).

The results demonstrate that VACV LIVP was significantly attenuated in the case of inactivation of the thymidine kinase gene that had occurred when producing the recombinant LIVP-GFP strain.

Comparison of changes in the development of humoral immune response to vaccination of mice with the LIVP and LIVP-GFP viruses over time

Adult BALB/c mice, starting with the age of 6–7 weeks, were vaccinated by i.d. injection or s.s. inoculation with low-dose VACV LIVP or LIVP-GFP (10^4 pfu).

Upon i.d. injection of the LIVP virus, significant production of VACV-specific IgM was observed as early as on 7 dpi; its maximum level was reached by 21 dpi, while the IgM titer dropped to the level observed for the negative control group by 28 dpi and later. Therefore, the results of testing IgM in mouse serum samples are shown in Fig. 3 only for the time points of 7, 14, 21, and 28 dpi. Immunization of mice with the LIVP virus by s.s. inoculation resulted in later and less marked IgM production (see Fig. 3, *a*).

Both in the case of i.d. and s.s. vaccination of mice with the LIVP-GFP virus (10^4 pfu), IgM production was minimal; there were no significant differences compared to the IgM level in the control serum samples of non-immunized animals (see Fig. 3, *b*).

Much higher titers of VACV-specific immunoglobulins IgG were produced compared to those of IgM (see Figs. 3 and 4). After i.d. vaccination with the LIVP virus, significant IgG production was observed as early as on 7 dpi, reaching its

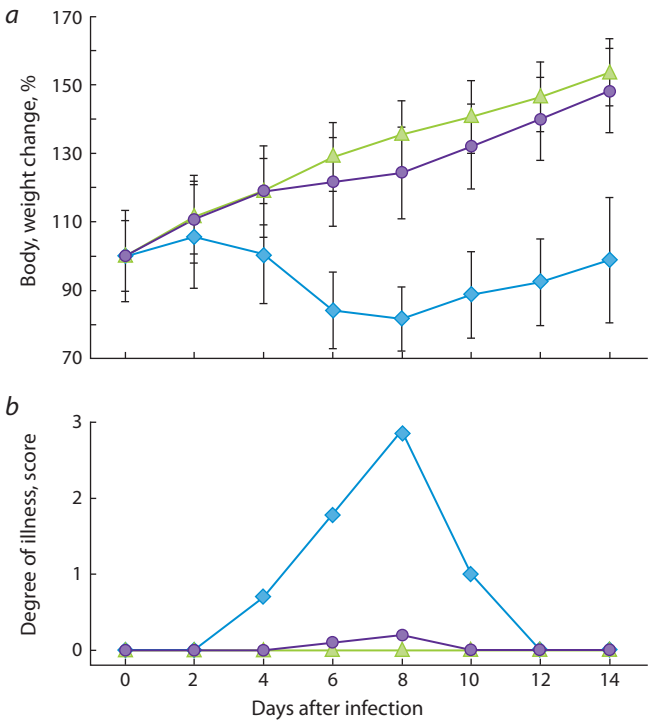

Fig. 1. Changes in mouse body weight (*a*) and clinical manifestations of infection (*b*) after intranasal inoculation of the LIVP (shown in blue) or LIVP-GFP (shown in violet) viruses at a dose of 10^7 pfu.

The mean data for groups consisting of 10 animals infected with the respective viruses and the control group (shown in green) are presented.

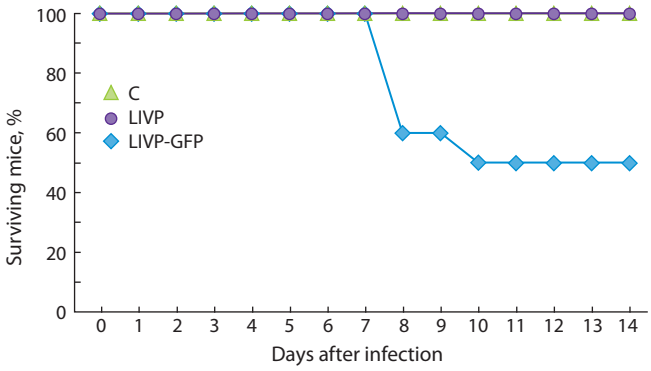

Fig. 2. Deaths of mice intranasally inoculated with the LIVP (shown in blue) or LIVP-GFP (shown in violet) viruses at a dose of 10^7 pfu. The control group consisted of non-infected animals (shown in green).

maximum on 21 dpi and remaining at virtually the same level until 28 dpi. The titer of VACV-specific IgG then gradually decreased by 42 and 56 dpi (see Fig. 4, *a*). For s.s. inoculation of the LIVP virus to mice, synthesis of specific IgG was delayed and had lower intensity compared to i.d. immunization (see Fig. 4, *a*).

When mice were i.d. inoculated with the LIVP-GFP virus, the IgG production level was considerably lower compared to that observed after vaccination with LIVP; the production of these antibodies was maximal on 28 dpi (see Fig. 4, *b*). S.s. inoculation of the LIVP-GFP virus resulted in lower-intensity production of analyzed IgG (see Fig. 4).

Assessment of protection against the lethal orthopoxvirus infection in immunized mice

In order to assess how the VACV strains under study, as well as the i.d. and s.s. vaccine administration routes, affect the development of protective immunity against orthopoxvirus reinfection in mice, groups of mice immunized with the LIVP or LIVP-GFP, as well as control (non-immunized) animals, were i.n. inoculated with CPXV GRI-90 at a dose of 46 LD_{50} on 62 dpi. The results of these experiments (Fig. 5) demonstrate that only the group of mice i.d. immunized with the LIVP virus were fully protected. In the group of mice vaccinated with the same virus through the s.s. route, 83 % of animals died after being infected with CPXV-GRI (see Fig. 5, *a*).

I.d. injection of the LIVP-GFP virus protected 80 % of mice against reinfection with CPXV-GRI under the same conditions, while all the mice s.s. inoculated with LIVP-GFP died (see Fig. 5, *a*). The level of protection against the lethal CPXV infection in mice correlated with the intensity of clinical manifestations of this infection (see Fig. 5, *b*).

Hence, i.d. low-dose immunization with VACV (10^4 pfu) used in this study for mice is obviously more effective compared to s.s. inoculation in the development of protective immunity against heterologous orthopoxvirus infection (the cowpox virus).

Discussion

The large-scale epidemic of monkeypox among humans that spread to all continents in 2022 (Harapan et al., 2022; Shchelkunova, Shchelkunov, 2023) has put the question about mass vaccination against this infection in the risk groups on the agenda. Important issues were the need to properly choose the type of vaccine and the optimal route of smallpox vaccine administration.

The first-generation live smallpox vaccine is a VACV preparation produced by viral replication in skin of calves or other animals. Recent studies have shown that these vaccines consist of a mixture of different VACV variants (Osborne et al., 2007; Qin et al., 2011).

In present-day conditions, the VACV vaccine strains obtained by isolating clonal variants from first-generation vaccines are produced on mammalian cell cultures, and these preparations are known to be second-generation smallpox vaccines (Sanchez-Sampedro et al., 2015). Application of first- and second-generation smallpox vaccines for mass vaccination is currently limited because of the relatively high risk of severe complications (Fenner et al., 1988; Sanchez-Sampedro et al., 2015), since the number of compromised people, including those infected with HIV, has recently increased.

Third-generation attenuated smallpox viruses (having reduced pathogenicity) are produced by multiple passages of a certain VACV strain in the cell culture of a heterologous host. This process is accompanied by emergence of VACV variants carrying spontaneous deletions and mutations in the viral genome (Jacobs et al., 2009; Olson, Shchelkunov, 2017; Albarnaz et al., 2018).

The novel approach to producing fourth-generation smallpox vaccines consists in introducing targeted deletions/insertions that disrupt selected viral genes and lead to VACV

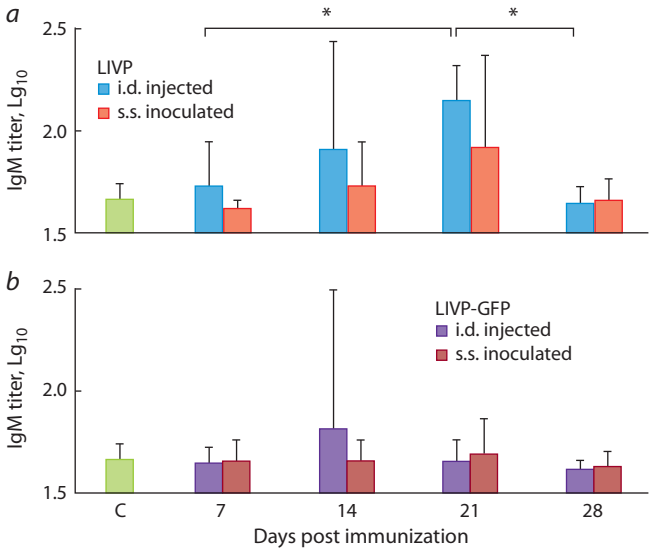

Fig. 3. Titers of VACV-specific IgM in serum samples from mice immunized with the LIVP (*a*) or LIVP-GFP viruses (*b*). C – serum samples of mice that received normal saline (control group).

* Statistically significant differences with $p < 0.05$.

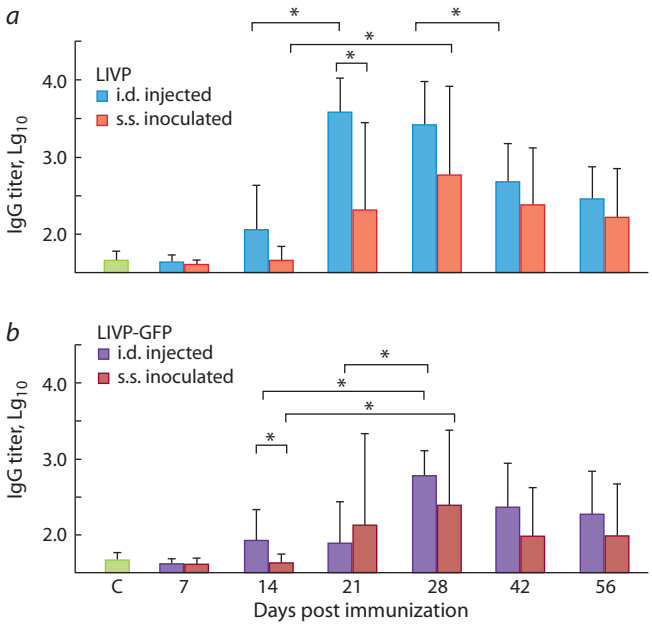

Fig. 4. Titers of VACV-specific IgG in serum samples of mice immunized with the LIVP (*a*) or LIVP-GFP (*b*) viruses. C – serum samples of mice that received normal saline (control group).

* Statistically significant differences with $p < 0.05$.

attenuation by genetic engineering (Yakubitskiy et al., 2015; Li et al., 2017; Shchelkunov et al., 2022b).

First-generation live smallpox vaccine based on the VACV LIVP strain is used for smallpox immunization in Russia. The LIVP strain was produced by epicutaneous passaging of the Lister vaccine strain provided by the Lister Institute (Elstree, UK) in rabbits and calves. A preparation of this vaccine is the virus grown in scarified calf skin (Perekrest et al., 2013). The patients receive this vaccine via s.s. inoculation.

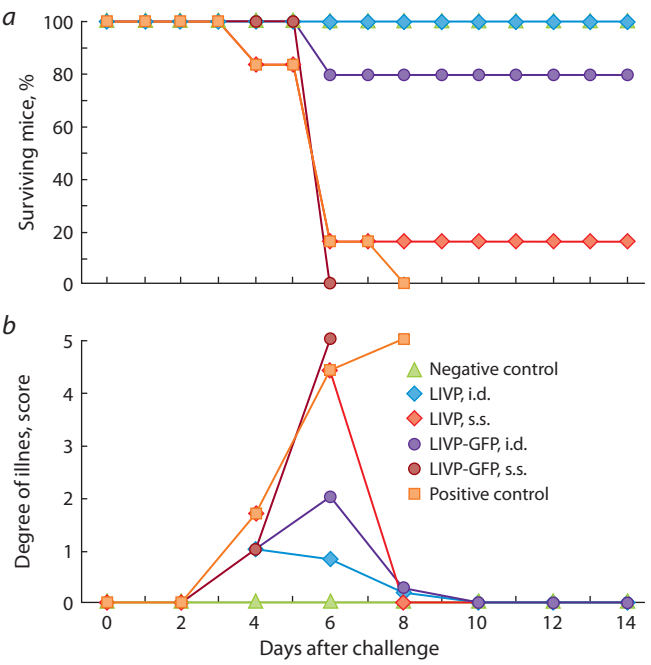

Fig. 5. Deaths (*a*) and clinical manifestations of infection in mice (*b*) immunized with the LIVP or LIVP-GFP viruses (10^4 pfu) after being intranasally infected with CPXV GRI-90 at a dose of 46 LD_{50} on 62 dpi.

The mean data for the groups consisting of six animals immunized with the respective viruses, as well as the non-immunized and non-infected mice (the negative control group) or mice infected with CPXV GRI-90 (the positive control group) are presented.

We used the LIVP strain to produce and characterize the clonal variant of LIVP (Yakubitskiy et al., 2015) that can be viewed as a prototype of second-generation smallpox vaccine. The recombinant LIVP-GFP strain with inactivated thymidine kinase gene generated based on it (Petrov et al., 2013) is a prototype variant of fourth-generation smallpox vaccine.

At the first stage of this study, we compared the pathogenicities of the LIVP and LIVP-GFP strains. The sensitivity of mice to orthopoxviruses significantly depends on their age (Shchelkunov et al., 2005); therefore, young (3-week-old) BALB/c mice were used in the experiments. The animals were i.n. inoculated with the viruses, since this route imitates the natural route of infection and ensures the highest sensitivity of mice to this infection (Hughes et al., 2020; Shchelkunov et al., 2021).

It turned out that after i.n. inoculation of young mice with the LIVP strain (10^7 pfu), it induced clinically apparent infection (see Fig. 1) resulting in death of 50 % of animals (see Fig. 2). Meanwhile, the LIVP-GFP strain led only to mild signs of the disease in mice (see Fig. 1) and complete recovery (see Fig. 2). Therefore, inactivation of the thymidine kinase gene in LIVP-GFP resulted in its substantial attenuation compared to the parental LIVP strain, which is consistent with the results obtained for other VACV strains (Taylor et al., 1991; Jacobs et al., 2009).

Numerous studies have previously demonstrated that s.s. immunization with second- and fourth-generation VACV-based vaccines at doses of at least 10^5–10^6 pfu fully protected mice against repeated lethal orthopoxvirus infection (Melamed et al., 2007; Jacobs et al., 2009; Shchelkunov et al., 2022a).

In this work, we studied the feasibility of reducing the dose of prototypic smallpox vaccines to 10^4 pfu when performing s.s. inoculation or i.d. injection to mice. For correct comparison, the VACV was introduced by the s.s. and i.d. routes within the same region of mouse tail skin.

Adult mice (aged 6–7 weeks) with a mature immune system were used for studying the immunogenicity of VACV LIVP and LIVP-GFP. The antibody response is known to make the most significant contribution to the development of adaptive immune response to VACV vaccination (Belyakov et al., 2003; Moss, 2011). Therefore, we studied changes in the synthesis of VACV-specific IgM and IgG after i.d. or s.s. vaccination of mice with the LIVP or LIVP-GFP strains. The results of these experiments demonstrated (see Figs. 3 and 4) that both VACV strains ensured more profound antibody production upon i.d. injection compared to s.s. inoculation. Meanwhile, statistically significant differences in the results between the compared groups were revealed only for IgG values on 21 dpi for LIVP (see Fig. 4, *a*) and 14 dpi for LIVP-GFP (see Fig. 4, *b*). No statistically significant differences in the results were observed for IgM (see Fig. 3).

In order to assess the level of protective immunity that developed in mice in response to s.s. or i.d. immunization with the LIVP or LIVP-GFP viruses, these animals were subjected to i.n. infection with a highly lethal dose of CPXV. It was considered to be the most adequate approach to assessing the effectiveness of VACV vaccination on the mouse model (Ferrier-Rembert et al., 2007; Melamed et al., 2007). The results (see Fig. 5) demonstrated that i.d. injection ensured a much stronger protective immunity compared to s.s. inoculation of the VACV. Only i.d. low-dose immunization with the LIVP strain fully protected mice against the lethal CPXV infection. The attenuated LIVP-GFP strain did not form a sufficiently strong protective immunity under the same conditions. S.s. inoculation with VACV LIVP or LIVP-GFP at the selected low dose did not protect animals against reinfection with CPXV (see Fig. 5).

Conclusion

These findings give grounds for inferring that i.d. injection of both studied VACV variants induces a much stronger protective immunity in mice compared to s.s. inoculation of these viruses at the same dose. In addition to more accurate vaccine dosing for i.d. immunization compared to the s.s. route, the former one is associated with less significant skin damage, thus substantially reducing the intensity of inflammation reaction that impedes efficient VACV replication and lowering the risk of bacterial infection at the vaccination site (Shmeleva et al., 2022). When using an attenuated fourth-generation vaccine with reduced specific immunogenicity for smallpox immunization, a higher dose of VACV needs to be used as compared to that of the second-generation vaccine.

References

Albarnaz J.D., Torres A.A., Smith G.L. Modulating vaccinia virus immunomodulators to improve immunological memory. *Viruses.* 2018; 10(3):101. DOI 10.3390/v10030101.

Belyakov I.M., Earl P., Dzutsev A., Kuznetsov V.A., Lemon M., Wyatt L.S., Snyder J.T., Ahlers J.D., Franchini G., Moss B., Berzofsky J.A. Shared models of protection against poxvirus infection by attenuated and conventional smallpox vaccine viruses. *Proc. Natl. Acad. Sci. USA.* 2003;100(16):9458-9463. DOI 10.1073/pnas.1233578100.

Esparza J., Schrick L., Damaso C.R., Nitsche A. Equination (inoculation of horsepox): An early alternative to vaccination (inoculation of cowpox) and the potential role of horsepox virus in the origin of the smallpox vaccine. *Vaccine*. 2017;35(52):7222-7230. DOI 10.1016/j.vaccine.2017.11.003.

Fenner F., Henderson D.A., Arita I., Jezek Z., Ladnyi I.D. Smallpox and its Eradication. Geneva: World Health Organization, 1988.

Ferrier-Rembert A., Drillien R., Tournier J.-N., Garin D., Crance J.-M. Intranasal cowpox virus infection of the mouse as a model for preclinical evaluation of smallpox vaccines. *Vaccine*. 2007;25(25): 4809-4817. DOI 10.1016/j.vaccine.2007.04.011.

Harapan H., Ophinni Y., Megawati D., Frediansyah A., Mamada S.S., Salampe M., Bin Emran T., Winardi W., Fathima R., Sirinam S., Sittikul P., Stoian A.M., Nainu F., Sallam M. Monkeypox: A comprehensive review. *Viruses.* 2022;14(10):2155. DOI 10.3390/v14102155.

Hughes L.J., Townsend M.B., Gallardo-Romero N., Hutson C.L., Patel N., Dotty J.B., Salzer J.S., Damon I.K., Carroll D.S., Satheshkumar P.S., Karem K.L. Magnitude and diversity of immune response to vaccinia virus is dependent on route of administration. *Virology.* 2020;544:55-63. DOI 10.1016/j.virol.2020.02.001.

Jacobs B.L., Langland J.O., Kibler K.V., Denzler K.L., White S.D., Holechek S.A., Wong S., Huynh T., Baskin C.R. Vaccinia virus vaccines: past, present and future. *Antiviral Res*. 2009;84(1):1-13. DOI 10.1016/j.antiviral.2009.06.006.

Li Y., Zhu Y., Chen S., Li W., Yin X., Li S., Xiao P., Han J., Li X., Sun L., Jin N. Generation of an attenuated Tiantan vaccinia virus strain by deletion of multiple genes. *Front. Cell. Infect. Microbiol.* 2017;7:462. DOI 10.3389/fcimb.2017.00462.

Melamed S., Paran N., Katz L., Ben-Nathan D., Israely T., Schneider P., Levin R., Lustig S. Tail scarification with vaccinia virus Lister as a model for evaluation of smallpox vaccine potency in mice. *Vaccine*. 2007;25(45):7743-7753. DOI 10.1016/j.vaccine.2007.09.023.

Moss B. Smallpox vaccines: targets of protective immunity. *Immunol. Rev.* 2011;239(1):8-26. DOI 10.1111/j.1600-065X.2010.00975.x.

Olson V.A., Shchelkunov S.N. Are we prepared in case of a possible smallpox-like disease emergence? *Viruses.* 2017;9(9):242. DOI 10.3390/v9090242.

Osborne J.D., da Silva M., Frace A.M., Sammons S.A., Olsen-Rasmussen M., Upton C., Buller R.M., Chen N., Feng Z., Roper R.L., Liu J., Pougatcheva S., Chen W., Wohlhueter R.M., Esposito J.J. Genomic differences of *Vaccinia virus* clones from Dryvax smallpox vaccine: The Dryvax-like ACAM2000 and the mouse neurovirulent Clone-3. *Vaccine*. 2007;25(52):8807-8832. DOI 10.1016/j.vaccine.2007.10.040.

Perekrest V.V., Movsesyants A.A., Mukhacheva A.V., Shevtsov V.A., Shvedov D.V., Borisevich I.V. Preparations for the specific prophylaxis of smallpox approved in the Russian Federation. *Boipreparaty = Biopreparation (Biopharmaceuticals)*. 2013;2:4-13. (in Russian)

Petrov I.S., Goncharova E.P., Pozdnyakov S.G., Shchelkunov S.N., Zenkova M.A., Vlasov V.V., Kolosova I.V. Antitumor effect of the LIVP-GFP recombinant vaccinia virus. *Dokl. Biol. Sci.* 2013; 451(1):248-252. DOI 10.1134/S0012496613040133.

Qin L., Upton C., Hazes B., Evans D.H. Genomic analysis of the *Vaccinia virus* strain variants found in Dryvax vaccine. *J. Virol*. 2011; 85(24):13049-13060. DOI 10.1128/JVI.05779-11.

Sachs L. Statistische Auswertungsmethoden. Heidelberg (Germany): Springer, 1972.

Sanchez-Sampedro L., Perdiguero B., Mejias-Perez E., Garcia-Arriaza J., Di Pilato M., Esteban M. The evolution of poxvirus vaccines. *Viruses.* 2015;7(4):1726-1803. DOI 10.3390/v7041726.

Shchelkunov S.N., Safronov P.F., Totmenin A.V., Petrov N.A., Ryazankina O.I., Gutorov V.V., Kotwal G.J. The genomic sequence analysis of the left and right species-specific terminal region of a cowpox virus strain reveals unique sequences and a cluster of intact ORFs for immunomodulatory and host range proteins. *Virology.* 1998;243(2):432-460. DOI 10.1006/viro.1998.9039.

Shchelkunov S.N., Marennikova S.S., Moyer R.W. Orthopoxviruses Pathogenic for Humans. New York: Springer, 2005.

Shchelkunov S.N., Yakubitskiy S.N., Sergeev A.A., Kabanov A.S., Bauer T.V., Bulichev L.E., Pyankov S.A. Effect of the route of administration of the vaccinia virus strain LIVP to mice on its virulence and immunogenicity. *Viruses*. 2020;12(8):795. DOI 10.3390/v12080795.

Shchelkunov S.N., Sergeev A.A., Kabanov A.S., Yakubitskiy S.N., Bauer T.V., Pyankov S.A. Route-coupled pathogenicity and immunogenicity of vaccinia virus variant inoculated mice. *Infektsiya i Immunitet = Russian Journal of Infection and Immunity*. 2021;11(2):357-364. DOI 10.15789/2220-7619-PAI-1375. (in Russian)

Shchelkunov S.N., Sergeev A.A., Titova K.A., Pyankov S.A., Starostina E.V., Borgoyakova M.B., Kisakova L.A., Kisakov D.N., Karpenko L.I., Yakubitskiy S.N. Comparison of the effectiveness of transepidemal and intradermal immunization of mice with the vaccinia virus. *Acta Naturae*. 2022a;14(4):111-118. DOI 10.32607/actanaturae.11857.

Shchelkunov S.N., Yakubitskiy S.N., Nesterov A.E., Kolosova I.V., Sergeev A.A., Zaykovskaya A.V., Kabanov A.S., Nechaeva E.A., Bogryantseva M.P., Usova S.V., Kutserubova N.S., Tregubchak T.V., Gavrilova E.V., Maksyutov R.A. Preclinical studies of the specific activity of the live culture vaccine VACΔ6 against smallpox and other orthopoxvirus infections. *Epidemiologiya i Vaktsinoprofilaktika = Epidemiology and Vaccinal Prevention*. 2022b;21(6):34-47. DOI 10.31631/2073-346-2022-21-6-34-47. (in Russian)

Shchelkunova G.A., Shchelkunov S.N. Smallpox, monkeypox and other human orthopoxvirus infections. *Viruses.* 2023;15(1):103. DOI 10.3390/v15010103.

Shmeleva E.V., Gomez de Aguero M., Wagner J., Enright A.J., Macpherson A.J., Ferguson B.J., Smith G.L. Smallpox vaccination induces a substantial increase in commensal skin bacteria that promote pathology and influence the host response. *PLoS Pathog*. 2022;18(4):e1009854. DOI 10.1371/journal.ppat.1009854.

Taylor G., Stott E.J., Wertz G., Ball A. Comparison of the virulence of wild-type thymidine kinase (tk)-deficient and tk^+ phenotypes of vaccinia virus recombinants after intranasal inoculation of mice. *J. Gen. Virol*. 1991;72(Pt. 1):125-130. DOI 10.1099/0022-1317-72-1-125.

Tulman E.R., Delhon G., Afonso C.L., Lu Z., Zsak L., Sandybaev N.T., Kerembekova U.Z., Zaitsev V.L., Kutish G.F., Rock D.L. Genome of horsepox virus. *J. Virol*. 2006;80(18):9244-9258. DOI 10.1128/JVI.00945-06.

Yakubitskiy S.N., Kolosova I.V., Maksyutov R.A., Shchelkunov S.N. Attenuation of vaccinia virus. *Acta Naturae.* 2015;7(4):113-121. DOI 10.32607/20758251-2015-7-4-113-121.

ORCID ID

S.N. Shchelkunov orcid.org/0000-0002-6255-9745
A.A. Sergeev orcid.org/0000-0001-8355-5551
S.A. Pyankov orcid.org/0000-0002-6593-6614
S.N. Yakubitskiy orcid.org/0000-0002-0496-390X

Acknowledgements. This work was supported by the Russian Science Foundation (grant No. 19-14-00006-P).

Conflict of interest. The authors declare no conflict of interest.

Received January 30, 2023. Revised May 27, 2023. Accepted May 29, 2023.