2024 • 28 • 4

Online ISSN 2500-3259

Сетевое издание

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

VAVILOV JOURNAL OF GENETICS AND BREEDING

Основан в 1997 г.
Периодичность 8 выпусков в год
DOI 10.18699/vjgb-24-41

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

2024 • 28 • 4

Online ISSN 2500-3259

Online edition

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

VAVILOV JOURNAL OF GENETICS AND BREEDING

Founded in 1997
Published 8 times annually
DOI 10.18699/vjgb-24-41

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 • VAVILOV JOURNAL OF GENETICS AND BREEDING

СОДЕРЖАНИЕ • 2024 • 28 • 4

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII • VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2024 • 28 • 4

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):365-376

DOI 10.18699/vjgb-24-42

Disomic chromosome 3R(3B) substitution causes a complex of meiotic abnormalities in bread wheat *Triticum aestivum* L.

A.A. Zhuravleva, O.G. Silkova

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
silkova@bionet.nsc.ru

Abstract. *Triticum aestivum* L. lines introgressed with alien chromosomes create a new genetic background that changes the gene expression of both wheat and donor chromosomes. The genes involved in meiosis regulation are localized on wheat chromosome 3B. The purpose of the present study was to investigate the effect of wheat chromosome 3B substituted with homoeologous rye chromosome 3R on meiosis regulation in disomically substituted wheat line 3R(3B). Employing immunostaining with antibodies against microtubule protein, α-tubulin, and the centromere-specific histone (CENH3), as well as FISH, we analyzed microtubule cytoskeleton dynamics and wheat and rye 3R chromosomes behavior in 3R(3B) (*Triticum aestivum* L. variety Saratovskaya 29 × *Secale cereale* L. variety Onokhoiskaya) meiosis. The results revealed a set of abnormalities in the microtubule dynamics and chromosome behavior in both first and second divisions. A feature of metaphase I in 3R(3B) was a decrease in the chiasmata number compared with variety Saratovskaya 29, 34.9 ± 0.62 and 41.92 ± 0.38, respectively. Rye homologs 3R in 13.18 % of meiocytes did not form bivalents. Chromosomes were characterized by varying degrees of compaction; 53.33 ± 14.62 cells lacked a metaphase plate. Disturbances were found in microtubule nucleation at the bivalent kinetochores and in their convergence at the spindle division poles. An important feature of meiosis was the asynchronous chromosome behavior in the second division and dyads at the telophase II in 8–13 % of meiocytes, depending on the anther studied. Considering the 3R(3B) meiotic phenotype, chromosome 3B contains the genes involved in the regulation of meiotic division, and substituting 3B3B chromosomes with rye 3R3R does not compensate for their absence.

Key words: chromosome substitution; meiosis; FISH; immunostaining; rye *Secale cereale* L.; common wheat *Triticum aestivum* L.

For citation: Zhuravleva A.A., Silkova O.G. Disomic chromosome 3R(3B) substitution causes a complex of meiotic abnormalities in bread wheat *Triticum aestivum* L. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):365-376. DOI 10.18699/vjgb-24-42

Funding. The work was carried out with financial support from the budget project FWNR-2022-0017. The plants were grown at the Center for Common Use for Plant Reproduction of the Institute of Cytology and Genetics SB RAS with the financial support of the budget project FWNR-2022-0017. The analysis of preparations was carried out at the Center for Microscopic Analysis of Biological Objects of the SB RAS.

Дисомное замещение хромосом 3R(3B) приводит к комплексу аномалий в мейозе мягкой пшеницы *Triticum aestivum* L.

А.А. Журавлева, О.Г. Силкова

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
silkova@bionet.nsc.ru

Аннотация. У линий мягкой пшеницы с интрогрессией чужеродных хромосом создается новый генетический фон, который изменяет экспрессию генов как пшеницы, так и хромосом-доноров родственных видов. На хромосоме 3B пшеницы локализованы гены, участвующие в регуляции мейоза. Целью работы было изучить влияние замещения хромосомы пшеницы 3B гомеологичной хромосомой ржи 3R на регуляцию мейоза у дисомно замещенной линии пшеницы 3R(3B). С помощью иммуноокрашивания с антителами к белку микротрубочек, α-тубулину и центромероспецифичному гистону H3 (CENH3), а также с использованием флуоресцентной *in situ* гибридизации проведен анализ динамики микротрубочкового цитоскелета и поведения хромосом пшеницы и ржи 3R в мейозе линии 3R(3B) (*Triticum aestivum* L. сорт Саратовская 29 × *Secale cereale* L. сорт Онохойская). В результате работы обнаружен комплекс аномалий в динамике микротрубочек и поведении хромосом как в первом, так и во втором делениях. Особенностью метафазы I у линии 3R(3B) являлось уменьшение числа хиазм в сравнении с сортом Саратовская 29 – 34.9 ± 0.62 и 41.92 ± 0.38 соответственно. Гомологи хромосомы ржи 3R в 13.18 % мейоцитов не формировали биваленты. Хромосомы характеризовались различной степенью компактизации,

в 53.33 ± 14.62 клетки отсутствовала метафазная пластинка. Установлены нарушения в нуклеации микротрубочек на кинетохорах отдельных бивалентов и в их конвергенции на полюсах деления веретена. Важной особенностью мейоза было асинхронное поведение хромосом во втором делении и наличие диад на стадии телофазы II в 8–13 % мейоцитов в зависимости от изученного пыльника. Таким образом, согласно мейотическому фенотипу линии 3R(3B), на хромосоме 3B сорта Саратовская 29 находятся гены, участвующие в регуляции комплекса мейотических процессов, а замещение хромосомами ржи 3R3R хромосом 3B3B не компенсирует их отсутствия.
Ключевые слова: замещение хромосом; мейоз; FISH; иммуноокрашивание; рожь *Secale cereale* L.; мягкая пшеница *Triticum aestivum* L.

Introduction

Bread wheat *Triticum aestivum* L. is characterized by tolerance to genomic introgressions of genetic material from wild and cultivated relatives. Alien chromosomes or their fragments may yield valuable traits, such as resistance to biotic and abiotic stresses, which is widely used in breeding programs (Mohammed et al., 2014; Yudina et al., 2014; Kroupin et al., 2019). However, introduction of alien chromosomes may also affect the regulation of basic biological processes, such as meiotic division, as early as in first-generation hybrids (Loginova et al., 2020).

Meiotic regulation in wheat has its peculiarities. While it is a hetero-hexaploid species (2*n* = 42, AABBDD genome), the meiotic behavior of its chromosomes matches that of a diploid organism. Chromosome pairing is controlled by *Ph* (*Pairing homoeologous*) genes. The *Ph1* gene suppressing meiotic homoeologous pairing is localized on 5BL chromosome (Sears, 1977; Giorgi, 1978), and the *Ph2* gene with the same albeit weaker effect, on 3DS chromosome (Mell-Sampayo, 1971). The *Ph1* locus sized 2.5 MB contains subtelomeric heterochromatin inserted within a cluster of *CDK2*-like genes (Griffiths et al., 2006; Al-Kaff et al., 2008; Martín et al., 2017). The gene initially referred to as "*hypothetical 3*" (*Hyp3*) (Griffiths et al., 2006; Al-Kaff et al., 2008) and later reannotated as *ZIP4* (*TaZIP4-B2*) (UniProtKB–Q2L3T5), based on meiotic phenotype of the *ph1b* common wheat mutants, was incorporated into a heterochromatin segment during wheat polyploidization (Martín et al., 2017). Here, *TaZIP4-B2* was responsible for progression of homologous and inhibition of homoeologous crossover, including by being involved in synaptonemal complex formation (Martín et al., 2017, 2018).

Bread wheat genome sequencing revealed the phylogenomic origin of *ZIP4* (Appels et al., 2018). It was demonstrated that *ZIP4* was a transduplication of a 3B chromosome locus having inparalogs on chromosomes 3A and 3D. In other words, hexaploid wheat carries four *ZIP4* copies, i. e. one copy on each chromosome of group 3 (3A, 3B, 3D) and a duplicated copy on 5B chromosome. Earlier, while establishing aneuploid lines of the Chinese Spring common wheat variety, it was shown that the absence of 3B chromosome resulted in meiotic asynapsis and reduced plant fertility (Sears, 1954). The latter findings were confirmed later, and the gene was localized on the long arm of 3BL (Bassi et al., 2013). It was shown that the loss of 3B chromosome resulted in pairing inhibition and reduced chiasmata count in meiosis. Notably, the effect of short arm (3BS) deletion was less significant than that of long arm deletion (Darrier et al., 2022). The desynapsis gene had no official designation in wheat (McIntosh et al., 2013), so, given its possible synthetic relationship to *des2* on chromosome 3H in barley (Ramage, Hernandez-Soriano, 1972), the designation *Tdes2* was suggested, with "des" standing for desynaptic and "T" for Triticum (Bassi et al., 2013). In addition, *QTug.sau-3B*, a QTL responsible for unreduced gamete production in interspecific hybrids, was identified on 3B chromosome (Hao et al., 2014). A total of 16 meiotic genes were localized on 3B chromosome in the Chinese spring reference variety (Darrier et al., 2022). It was also shown that orthologs of wheat meiotic genes interacted with *TaZIP4* of group 3 chromosomes in various meiotic processes (Alabdullah et al., 2019).

In addition to meiotic genes, there are also genes responsible for agriculturally valuable traits, such as yield, kernel weight, shape, and color, seed dormancy period, resistance to *Stagonospora nodorum*, *Puccinia graminis* f. sp. *tritici*, *P. recondita*, as well as synthesis of certain isozymes, localized on homoeologous group 3 wheat chromosomes (Munkvold et al., 2004). *Qss.msub-3BL*, a QTL for stem solidness, controlling sawfly resistance in bread and durum wheats, was also localized on 3BL (Cook et al., 2004). Overall, a total of 6,000 genes were localized on chromosome 3B (Paux et al., 2006). Another noteworthy discovery was the evolutionary recent (100 ka) amplification burst of LTR retrotransposons (Ling et al., 2018) capable of affecting gene structure and expression (Bariach et al., 2020). Thus, chromosome 3B substitutions or its absence become relevant in terms of hybrid genotype development. It was also shown that gene expression changes occurred in both wheat and alien chromosomes in wheat-alien addition and substitution lines (Rey et al., 2018; Dong et al., 2020).

Therefore, studying the effect of substituting wheat chromosome 3B with rye chromosome 3R on meiosis regulation in wheat-rye disomic chromosome 3R(3B) substitution line (*T. aestivum* L. Saratovskaya 29 variety – *Secale cereale* L. Onokhoiskaya variety) is a relevant research issue (Silkova et al., 2006); in the present study, we analyzed microtubule cytoskeleton dynamics and investigated meiotic cycle progression as well as the behavior of wheat chromosomes and rye chromosome 3R.

Materials and methods

Plant material. The study employed the Saratovskaya 29 (S29) variety of *Triticum aestivum* L. bread wheat and wheat-rye disomic chromosome 3R(3B) substitution line (*T. aestivum* L. Saratovskaya 29 variety × *Secale cereale* L. Onokhoiskaya variety), where chromosome 3B of wheat was substituted with chromosome 3R of rye (Silkova et al., 2006) (Table 1). The plants were grown in a hydroponic greenhouse at the Institute of Cytology and Genetics, SB RAS at a 24/18 °C day/night temperature and photoperiod of LD 16:8.

Table 1. Methods used for cytogenetic analysis of the 3R(3B) line and Saratovskaya 29 plants

Method	3R(3B) line		Saratovskaya 29 variety	
	Total		Total	
	spikes	meiocytes	spikes	meiocytes
3 % acetocarmine staining, fixation in Navashin's fluid	11	982	6	470
3 % acetocarmine staining, fixation in 96 % ethanol : glacial acetic acid (3:1)	5	534	5	456
Immunostaining	5	648	5	573
Genomic *in situ* hybridization (GISH)	5	431	–	–

Cytogenetic analysis

Acetocarmine staining. Routine study of microtubule (MT) cytoskeleton dynamics in meiosis in the S29 variety and the 3R(3B) line was performed using the technique described earlier (Loginova et al., 2020). Modified Navashin's fluid was used as a fixative for immature spikes (Wada, Kusunoki, 1964), and meiocytes were analyzed at all stages of the first and second divisions of microsporogenesis (Table 1).

To study meiotic chromosome pairing in the S29 variety and the 3R(3B) line, we used acetic acid : 96 % ethanol mixture (1:3 volumetric ratio) as a fixative for immature spikes. All meiocytes at metaphase I and anaphase I qualified for evaluation were studied in all anthers (Table 1).

The specimens were studied using a Leica DM 2000 microscope (Leica Microsystems), and the images were recorded using a DFC 295 camera (Leica Microsystems).

Fluorescent in situ hybridization (FISH) and indirect immunostaining. Specimen preparation and FISH were performed using the technique described earlier (Loginova et al., 2020). Meiocytes at metaphase I and telophase II were analyzed. For the purposes of this study, we employed centromere-specific probe pAet6-09 for rice, wheat, rye, and barley chromosomes (Zhang et al., 2004), as well as genomic rye DNA. The DNA repeat sample of pAet6-09 was the courtesy of Dr. A. Lukaszewcki (University of California, Riverside, United States). Probe pAet6-09 was labeled with digoxigenin-11-dUTP via polymerase chain reaction (PCR). The total DNA of rye was labeled with Nick-translation (Invitrogen, Carlsbad, California, United States, cat. no. 18160-010) with biotin-16-dUTP. Probes were combined in various ratios and mixed with blocking wheat DNA. To reduce fluorescence fading, Vectashield antifade solution (Vector Laboratories No. X1215) containing 1μg/ml DAPI (4′,6-diamidino-2-phenylindol, Sigma-Aldrich, No. D9542, United States) for chromatin staining was used.

Specimen preparation and indirect immunostaining were performed using the technique described earlier (Loginova et al., 2020). The primary antibodies were anti-α-tubulin ones (Monoclonal Anti-α-Tubulin antibody produced in mouse, Sigma-Aldrich, No. T5168) (1:2,000 solution) and antibodies specific to kinetochore protein CENH3, i.e. a centromeric histone H3 variant for cereals (courtesy of Dr. A. Houben, IPK Gatersleben, Germany), 1:850 solution in 1xPBS buffer with 1 % BSA. The secondary anti-CENH3 antibodies were Rhodamine (TRITC)-conjugated AffiniPure Goat Anti-Rabbit IgG (H+L) (Jackson ImmunoResearch, No. 111-025-003) (1:100 solution); the secondary anti-α-tubulin antibodies were FITC-conjugated anti-mouse IgG (Sigma, 1:100 solution). To reduce fluorescence fading, we applied Vectashield antifade solution (Vector Laboratories No. X1215) containing 1 μg/ml DAPI (4′,6-diamidino-2-phenylindol, Sigma-Aldrich, No. D9542, United States) for chromatin staining.

The specimens were studied using an Axio Imager M1 microscope (Carl Zeiss AG, Germany) with ProgRes MF camera (Meta Systems, Jenoptic, Germany), Isis imaging software (Meta Systems, Jenoptic, Germany) as well as a LSM 780 NLO laser scanning microscope (Zeiss) with an AxioCam MRm camera (Zeiss) and ZEN imaging software (Zeiss). The images obtained were processed in Adobe Photoshop CS2.

Results

Chromatin and microtubule cytoskeleton dynamics at prophase of the first meiotic division in the S29 variety and the 3R(3B) line

Comparative analysis of prophase progression in the S29 variety and the 3R(3B) line did not show any differences before zygotene (Supplementary Materials 1 and 2)[1]. Meiocytes in the S29 variety and the 3R(3B) line changed their shape from rectangular (Supplementary Materials 1*a* and 2*a*, *d*) and triangular (Supplementary Materials 1*b* and 2*c*) to rounded (Supplementary Materials 1*d* and 2*b*) starting with early leptotene.

Three to four nucleoli were present in early leptotene (Supplementary Materials 1*a*, *b* and 2*b*) to later fuse into one (Supplementary Materials 1*c*, *d* and 2*e*). In leptotene-zygotene, thin chromatin threads formed a dense ball containing a single nucleolus shifted toward the nuclear envelope (Supplementary Materials 1*d* and 2*e*; Fig. 1*a*′, *b*′). Meiocyte maturation was accompanied by chromatin condensation. In zygotene, chromatin fiber thickening was observed (Supplementary Materials 1*d* and 2*e–g*).

In zygotene and pachytene, the S29 variety and the 3R(3B) line showed different chromatin distribution across the nucleus. Compared to the S29 variety, the 3R(3B) line was characterized by radial and irregular chromatin looping (Supplementary Material 2*i*, *j*). In pachytene, asymmetric chromosome grouping on one side of the nucleus was observed in both wheat and wheat-rye substitution line (Supplementary Mate-

[1] Supplementary Materials 1–7 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Zhuravleva_Engl_28_4.pdf

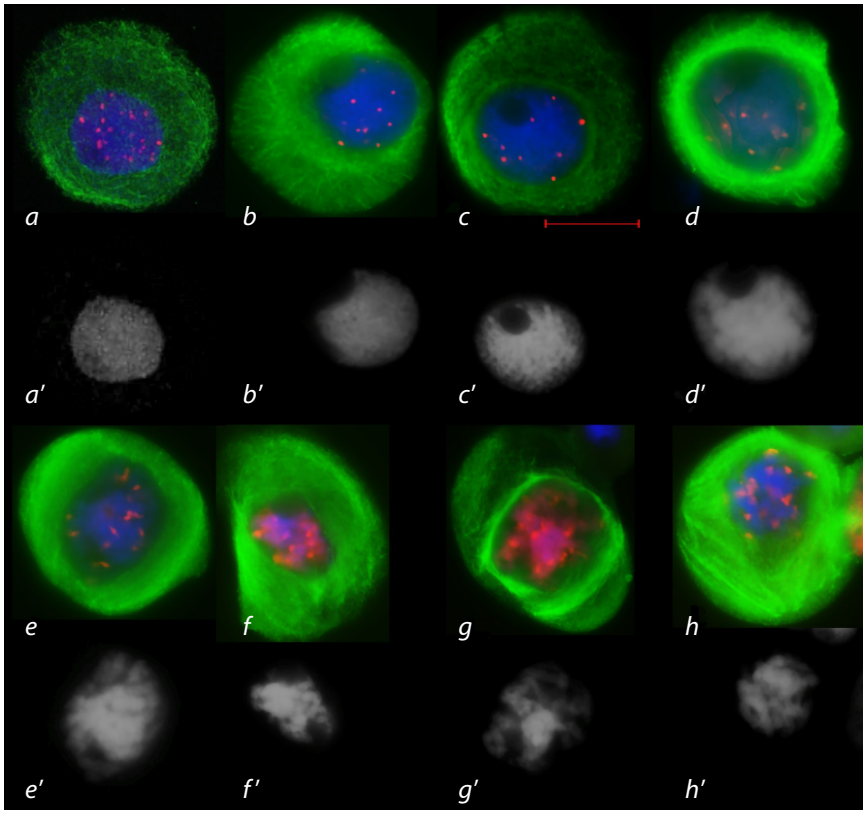


Fig. 1. MT cytoskeleton reorganization at prophase of the first meiotic division in the 3R(3B) line. *a* – leptotene-zygotene, prophase reticular cytoskeleton; *b* – zygotene, MTs move toward the nucleus to form a perinuclear ring; *c* – pachytene, dense MT ring; *d* – diplotene, MTs form a dense ring around the nucleus; *e*–*h* – diakinesis, consecutive stages of ring disintegration and MT reorientation; *l* – pro-spindle formation.

Immunostaining. DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm. *a*′–*h*′ – DAPI staining.

rials 1*e* and 2*h*). In diplotene, chromatin threads were shortened even more, while still in contact with the nuclear envelope (Supplementary Materials 1*f*, *g*; and 2*l*). In diakinesis, bivalent formation was completed with the nucleolus and the nuclear envelope still present (Supplementary Materials 1*h* and 2*m*, *n*). Compared to the S29 variety, chromosomes were densely packed in the 3R(3B) line in diakinesis (Supplementary Materials 1*h* and 2*n*).

Immunostaining analysis of meiocytes at prophase before pachytene did not show any differences between the MT dynamics observed in the 3R(3B) line and the earlier results for S29 (Loginova et al., 2020). Reticular cytoskeleton formation was observed at interphase and early prophase (Fig. 1*a*), then MTs were reorganized into radial bundles, reoriented, and moved toward the nucleus in zygotene-pachytene, while the nucleus itself migrated toward the envelope to form a "half-moon" MT structure (Fig. 1*b*).

In pachytene, a dense perinuclear ring was formed around the nucleus (Fig. 1*c*) at the center of the cell. The nucleus migrated toward the cell envelope, and cytoskeleton formed an arc-like structure in 10–90 % of the cells depending on the anther studied (Fig. 2*a*, *c*, *d*). MT nucleation density in the latter varied. In

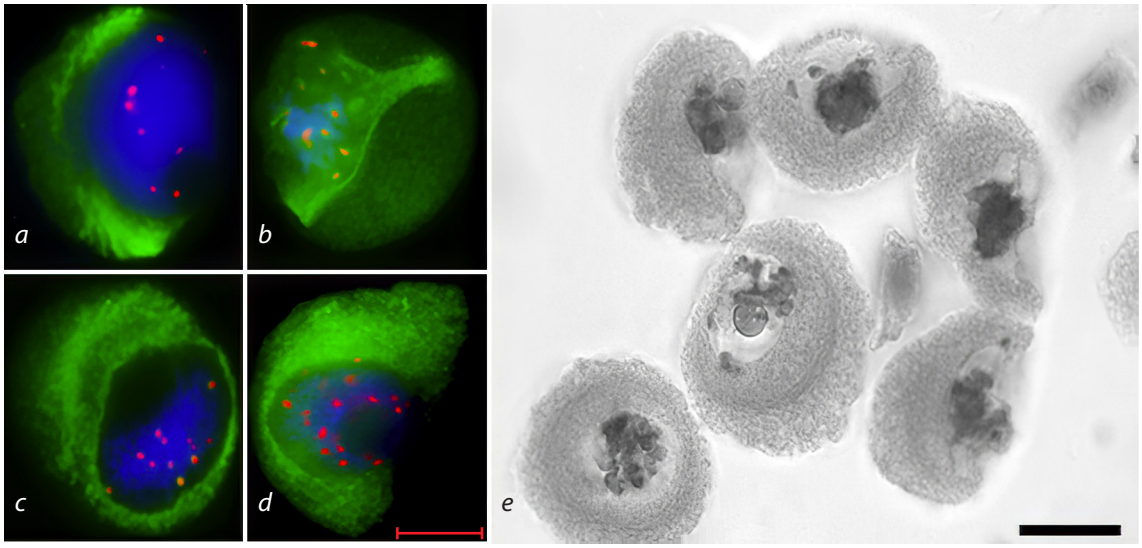


Fig. 2. Migration of the nucleus to the cell periphery and formation of arc-like MT structures at meiotic prophase in the 3R3B line. *a*, *d* – arc formed by MTs; *b* – MTs at the top of the arc form a spindle pole-like structure; *c* – incomplete migration of the nucleus to the periphery; *e* – a group of cells partially demonstrating migration of the nucleus toward the meiocyte envelope.

Immunostaining (*a*–*d*). DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm. Navashin's fluid fixation (*e*), acetocarmine staining; scale bar length is 10 μm.

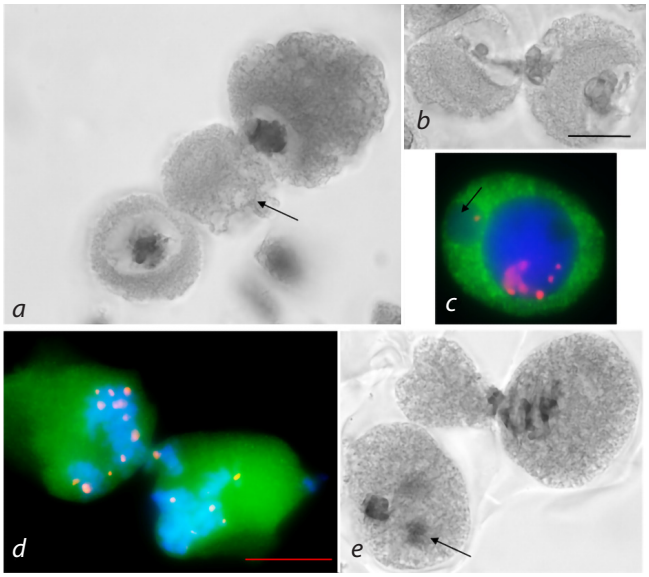


Fig. 3. Cytomixis in the wheat-rye 3R(3B) substitution line. *a* – migration of the nucleus at prophase with cytoplast formation (arrow); *b* – chromatin transfer from one meiocyte to another; *c* – micronucleus (arrow); *d* – chromosome transfer from one meiocyte to another at late prophase; *e* – cytomixis at metaphase I.

Navashin's fluid fixation, acetocarmine staining (*a*, *b*, *e*); scale bar length is 10 μm. Immunostaining (*c*, *d*). DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm.

Table 2. Cytomixis frequency in the S29 variety and the 3R(3B) line

Line/variety	Total cells analyzed	Total cells showing cytomixis or its consequences	Cell percentage, %
3R(3B)	480	24	5.0
S29	321	3	0.9

some meiocytes, MTs at the top of the arc formed a spindle pole-like structure (Fig. 2*b*).

In diplotene-diakinesis, ring disintegration occurred, and MTs were separated into beams and straightened out (Fig. 1*e–g*), 3- and 4-pole structures were formed in diakinesis (Fig. 1*g*, *h*) in both S29 variety wheat and rye (Loginova et al., 2020).

Migration of the nucleus toward the membrane in 5 % of the cells in the 3R(3B) line ended with chromatin transfer from one meiocyte to another as a result of cytomixis (Fig. 3; Table 2). Chromatin transfer occurred at prophase (Fig. 3*a*, *b*, *d*) and metaphase I (Fig. 3*e*). The cells chromatin was transferred from had reduced chromatin content or formed cytoplasts (Fig. 3*a*). The transferred chromatin formed a separate micronucleus (Fig. 3*c*) or was fused into the nucleus of the recipient cell (Fig. 3*e*).

Chromosome behavior and MT cytoskeleton dynamics in the first meiotic division in the S29 variety and the 3R(3B) line

After nuclear envelope disintegration, prophase spindle was disassembled, and at prometaphase MTs interacted with chromosome kinetochores and with each other to form central and kinetochore fibrils for the spindle apparatus in both the S29 variety and the 3R(3B) line (Fig. 4).

The ends of the microtubules converged at the poles to form the spindle apparatus and chromosomes were aligned along the equator of the cell (Fig. 5*a*).

A distinctive feature of metaphase I in the 3R(3B) line is the absence of the metaphase plate at the equator of the spindle apparatus and a different degree of chromosome compaction (Supplementary Material 3; Fig. 13). Meiocytes lacking a metaphase plate add up to 20 to 100 % depending on the anther analyzed, the average being 53.33 ± 14.62 % (Supplementary Material 3*a–e*).

Immunostaining analysis of chromosome behavior in the 3R(3B) line showed chaotic distribution of bivalents along the equatorial plane in cells lacking a metaphase plate (Fig. 6*a*) compared to the normal case (Fig. 5*a*), due to the absence of MT nucleation at kinetochores of individual bivalents (Fig. 7*a*, 8*a*) or anomalous connection of kinetochores of open and closed bivalents by MT beams (Fig. 7*a*). Meiocytes, where normal spindle apparatus could not be formed due to the absence of MT convergence at the pole (2 % meiocytes), were observed (Fig. 7*b*).

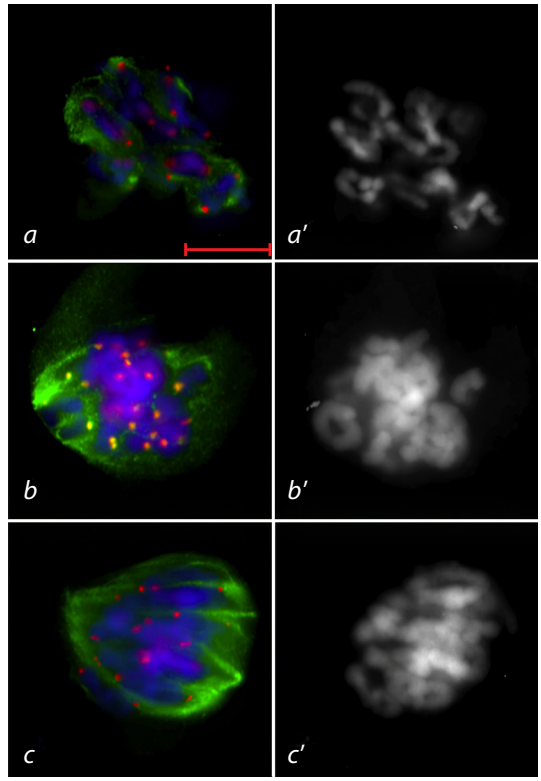


Fig. 4. Meiotic prometaphase I in the 3R(3B) line. *a* – interaction between MTs and chromosome kinetochores; *b*, *c* – formation of central and kinetochore fibrils for the spindle apparatus.

Immunostaining. DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm. DAPI staining (*a′–c′*).

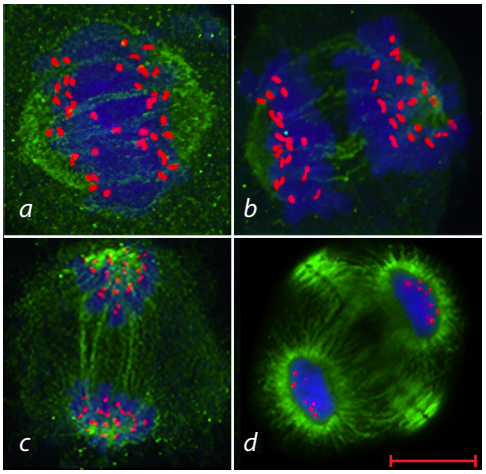


Fig. 5. MT cytoskeleton reorganization in the first meiotic division in the 3R(3B) line. *a* – metaphase, spindle apparatus is formed; *b* – anaphase, chromosomes are pulled to the poles with shortening of kinetochore fibrils of the spindle apparatus; *c* – late anaphase, only central fibrils of the spindle apparatus are present, radial cytoskeleton formation is initiated; *d* – telophase, phragmoplast formation.

Immunostaining. DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm.

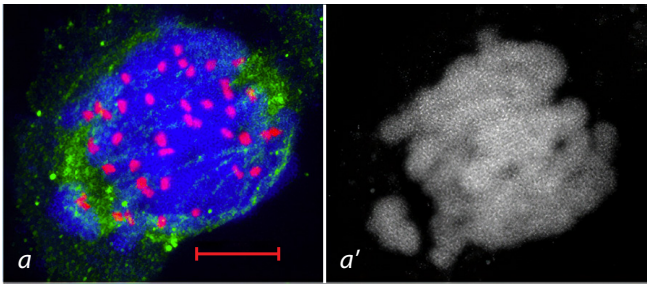


Fig. 6. Absence of the metaphase plate at metaphase I in the 3R(3B) line.

Immunostaining. DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm. DAPI staining (*a*′).

Metaphase I in the 3R(3B) line is characterized by reduced chiasmata count compared to the S29 variety (Table 3; Supplementary Material 3). According to observations, the number of rod bivalents per cell was 3.0±0.35, the number of ring bivalents per cell was 15.95±0.61, and the number of univalents was 3.79±1.0 (Table 3). No univalents were observed in the S29 variety, the number of ring bivalents was 20.92±0.04, and the number of rod bivalents was 0.08±0.04 (Table 3). Multivalents were observed in 1.2 % of meiocytes in the 3R(3B) line.

Sister kinetochores of univalent chromosomes at metaphase I were either separated or remained fused. In the first scenario, chromosomes were aligned along the equator (Fig. 8*b*), and in the second one, they were pulled randomly towards the poles before anaphase I (Fig. 8*a*). The absence of MT nucleation at the single kinetochore of a univalent (Fig. 8*b*) and a bivalent (Fig. 8*a*) could also cause metaphase plate formation abnormalities.

At early anaphase I, kinetochore fibrils of the meiotic spindle were shortened, and chromosomes were pulled to the poles in both the S29 variety and the 3R(3B) line (Supplementary Materials 4*b*, *c* and 5*b*, *c*). Chromosome distribution across the spindle apparatus did not depend on the degree of compaction (Supplementary Material 5*b*). Chromosome separation was followed by formation of a phragmoplast-cell plate structure (Supplementary Materials 4*d* and 5*d*) dividing a meiocyte into two daughter cells. The first division ends with the formation of a dyad with a radial cytoskeleton (Supplementary Materials 4*e* and 5*e*).

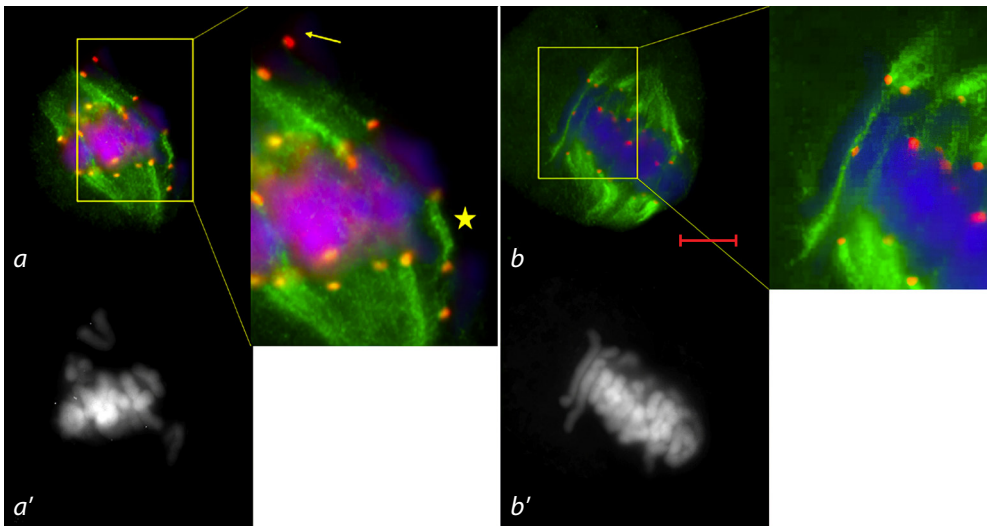


Fig. 7. Disruption of MT nucleation during spindle apparatus formation at metaphase I. *a* – absence of MT nucleation at the kinetochore of a rod bivalent (arrow), MT beams connect kinetochores of rod and ring bivalents (star); *b* – bivalents lacking MT convergence at the poles.

Immunostaining. DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm. DAPI staining (*a*′, *b*′).

Table 3. Formation of bivalents and univalents in the 3R(3B) line and the S29 wheat variety

Line/variety	Average number of ring bivalents per cell	Average number of rod bivalents per cell	Average number of univalents per cell	Average chiasmata count
3R(3B)	15.95 ± 0.6	3.00 ± 0.35	3.79 ± 1.02	34.90 ± 0.62
S29	20.92 ± 0.04***	0.08 ± 0.04***	0	41.92 ± 0.38

*** Significant differences at $p \leq 0.001$.

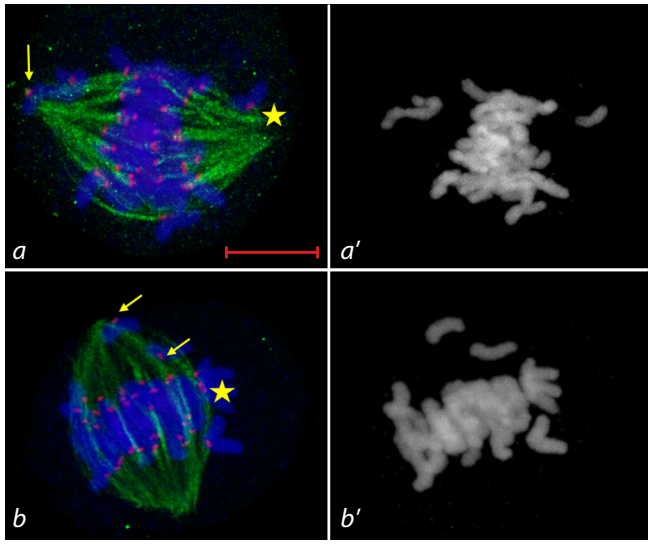


Fig. 8. Univalent distribution at metaphase I in the 3R(3B) line. *a* – MT attachment of univalents to one pole (star), absence of α-tubulin signal at the bivalent kinetochore (arrow); *b* – absence of α-tubulin signal at the univalent kinetochore (arrows), bipolar orientation of separated sister kinetochores (star).

Immunostaining. DNA is shown in blue, MTs in green, centromeric kinetochores in red; scale bar length is 5 μm. DAPI staining (*a′*, *b′*).

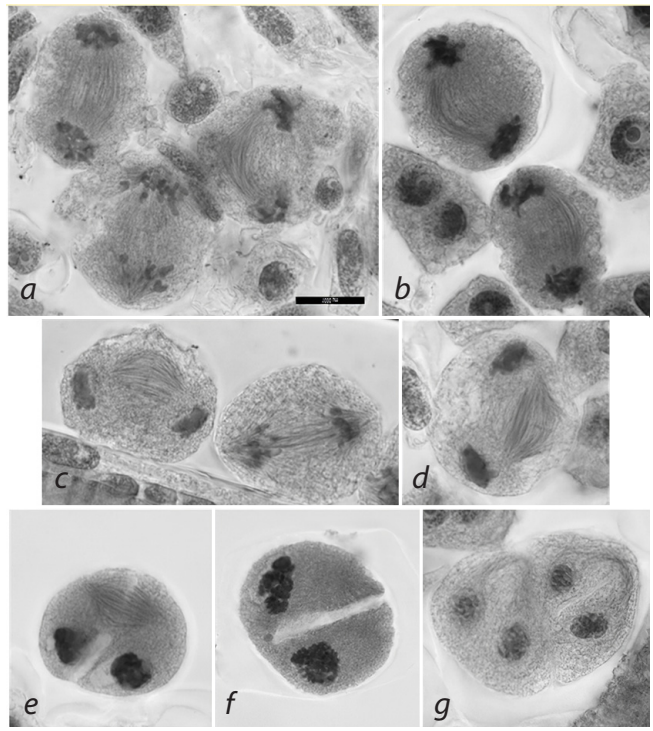


Fig. 9. C-shaped spindle formation in the 3R(3B) line. *a–c* – anaphase I; *d* – telophase I, autonomous spindle orientation; *e*, *f* – telophase I, cell wall in the form of a notch; *e* – autonomous spindle orientation; *g* – telophase II.

Navashin's fluid fixation, acetocarmine staining; scale bar length is 10 μm.

C-shaped spindles not affecting chromosome separation were observed at metaphase I and anaphase I (0 to 30 % of meiocytes, depending on the anther) and at telophase II in the 3R(3B) line (Fig. 9*a–d*, *g*).

The spindle apparatus maintained its shape after chromosome separation and was located near two telophase chromosome groups (Fig. 9*c–e*). Phragmoplast formation was disrupted, and cell wall emerged in the form of a notch not ensuring full separation of a meiocyte at telophase I (Fig. 9*e*, *f*).

Second meiotic division in the S29 variety and the 3R(3B) line

Analysis of the second division showed the presence of anthers with abnormalities in addition to normal meiotic progression in the 3R(3B) line similarly to the first division as opposed to the S29 variety. Normally, the radial cytoskeleton (Supplementary Material 6*a*) was transformed into MT beams at the second division prophase. The latter then formed the metaphase structure (Supplementary Materials 6*b*, *c* and 7*b*, *c*) and distributed sister chromatids between the poles (Supplementary Materials 6*d* and 7*d*).

Similarly to the first division, phragmoplast-cell plate system is formed, central fibrils of the spindle apparatus are preserved, and the division ends with tetrad formation (Supplementary Materials 6*e*, *f* and 7*e*, *f*). Micronuclei were detected in 4.2 % of tetrads (Fig. 10).

Asynchronous chromosome behavior was observed in the second division in the 3R(3B) line (Fig. 10, 11). Certain anthers from the same spike can simultaneously include meiocytes at different division stages: anaphase I, telophase I, metaphase II, anaphase II, and tetrads (Fig. 11). Dyads can be observed among the tetrads at telophase II (8 to 13 % of meiocytes per anther studied) (Fig. 10).

At telophase II, tetrads with unequally sized nuclei were observed in 10–20 % of meiocytes (Fig. 10, 12*a*), cytoplasts without nuclei, in 2.4 % of cells (Fig. 10, 12*c*), and triads, in 12.5 % cells (Fig. 10, 12*d–f*).

Rye chromosome 3R3R behavior in the first and second meiotic divisions

Rye chromosome 3R3R behavior was studied using FISH. At metaphase I, chromosomes 3R3R formed bivalents in 86.82 % of meiocytes, among which 21.36 % were rod biva-

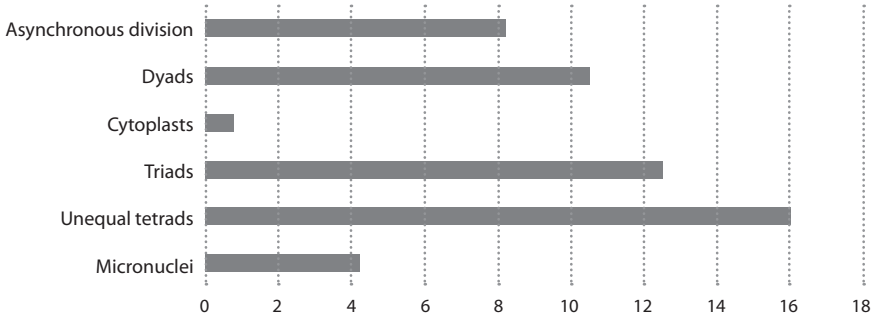


Fig. 10. Percentage of cells with various anomalies in the second meiotic division in the 3R(3B) line.

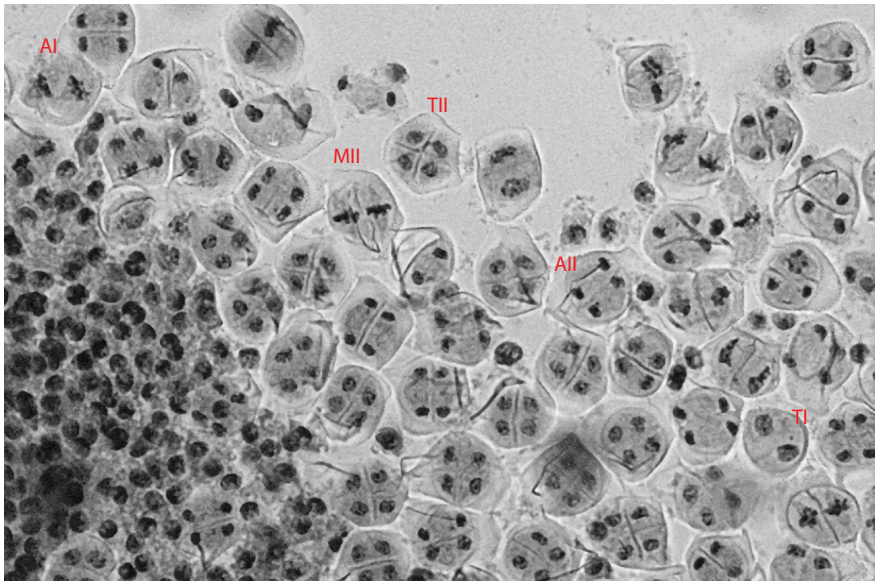


Fig. 11. Asynchronous chromosome behavior in the 3R(3B) line. Cells at different meiotic division stages within the same anther. AI – anaphase I, TI – telophase I, MII – metaphase II, AII – anaphase II, and TII – telophase II.

Navashin's fluid fixation, acetocarmine staining.

lents and 13.18 % were univalents (Fig. 13). Chromosome 3R was absent in 5.58 % of the cells.

At telophase I, separation of homologous chromosomes 3R3R was not disrupted in 98 % of meiocytes (Fig. 14*a*). At telophase II, the analysis showed the presence of chromosome 3R in all microspores of the tetrad (Fig. 14*b*), which is indicative of normal distribution pattern for both bivalents and univalents.

Discussion

Chromosome 3B is required for chiasmata formation between homologs, and its absence is not compensated by rye chromosome 3R

Chiasmata formation between homologs with simultaneous suppression of chiasmata formation between homoeologous chromosomes in bread wheat is controlled by the *TaZIP4-B2* gene identified within the *Ph1* locus localized on the long arm of chromosome 5B (Griffiths et al., 2006; Al-Kaff et al., 2008; Martín et al., 2017). However, cytogenetic studies of microsporogenesis in mutants of tetraploid and hexaploid wheats, as well as their aneuploid and deletion lines showed that genes regulating bivalent formation were also localized on wheat chromosome 3B independently from 5B (Sears, 1954; Lee et al., 1970; Lelley, 1976; Miller et al., 1983; Darrier et al., 2022; Draeger et al., 2023). For instance, nullisomic chromosome 3B in hexaploid wheat in presence of two 5B chromosomes causes reduced chiasmata count at metaphase I (asynapsis) (Sears, 1954; Lee et al., 1970; Kato, Yamagata, 1982; Darrier et al., 2022), while long arm deletions of varying sizes of chromosome 3B reduce the total chiasmata count by 35 % (Darrier et al., 2022).

Our study has shown that the distinctive feature of metaphase I in the 3R(3B) line is the reduced chiasmata count compared to the S29 variety, 34.9 ± 0.62 and 41.92 ± 0.38, respectively. Homologs of rye chromosome 3R3R also form bivalents only in 86.82 % of meiocytes, among which 21.36 % are rod bivalents. The earlier analysis of chromosome composition in the 3R(3B) line using cytogenetic and molecular methods showed the presence of two 5B chromo-

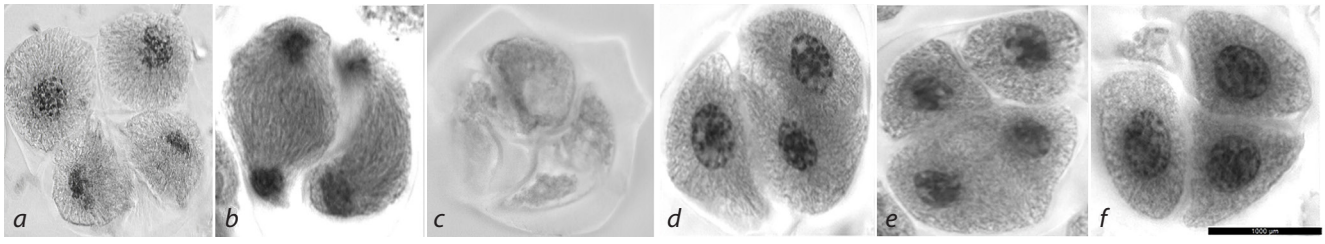


Fig. 12. Tetrad stage anomalies in the second meiotic division in the 3R(3B) line. *a* – tetrad with unequally sized nuclei; *b* – anomalous spindle in the second division, chromatin imbalance; *c* – tetrad without a nucleus; *d*, *f* – triads; *e* – absence of cell wall in one out of two cells.

Navashin's fluid fixation, acetocarmine staining; scale bar length is 10 μm.

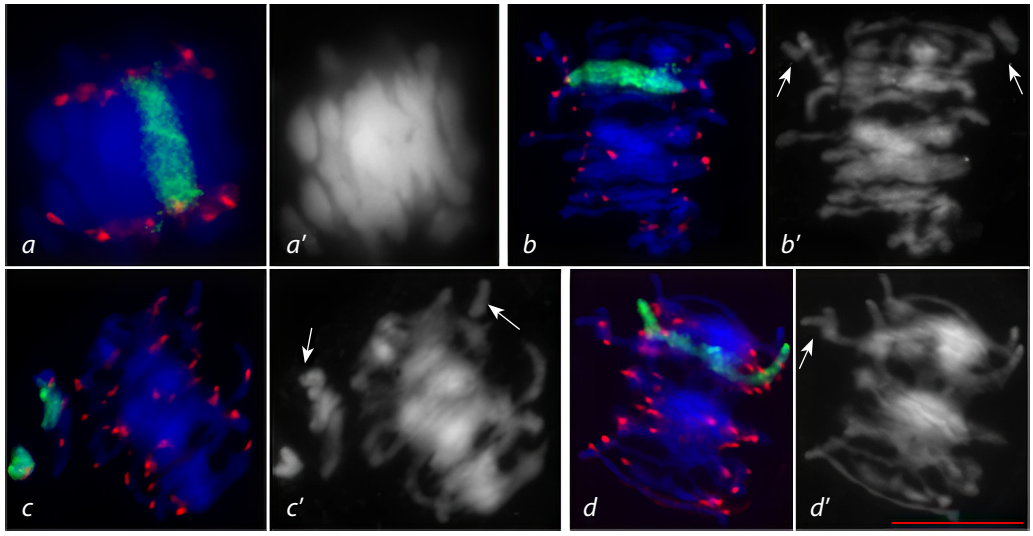


Fig. 13. Chromosome behavior at meiotic metaphase I in the 3R(3B) line. *a* – ring bivalent formation by chromosomes 3R3R, normal chromosome alignment along the equator; *b–d* – disturbance of chromosome compaction; *b* – ring bivalent formation by chromosomes 3R3R, wheat chromosome univalents (arrows); *c* – rye chromosome univalents, wheat chromosome univalents (arrows); *d* – open bivalent formation by chromosomes 3R3R, wheat chromosome univalent (arrow).

GISH: DNA is shown in blue, rye chromosomes in green, centromeric region in red. Scale bar length is 5 μm.

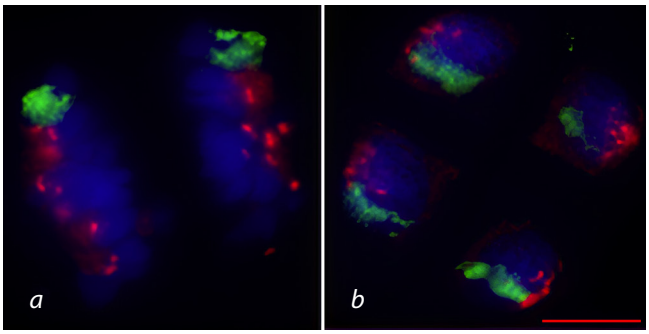


Fig. 14. Meiotic distribution of rye chromosomes in the 3R(3B) line at telophase I (a) and telophase II (b).

GISH: DNA is shown in blue, rye chromosomes in green, centromeric region in red. Scale bar length is 5 μm.

somes in the karyotype (Silkova et al., 2006). Thus, our results confirm the earlier findings with regard to the presence of genes on chromosome 3B regulating chiasmata formation independently from genes on chromosome 5B.

It has been recently confirmed that the *ZIP4* gene copies in the *Ph1* locus on chromosomes 5B (*TaZIP4*-B2), 3A (*TaZIP4-A1*), 3B (*TaZIP4-B1*), and 3D (*TaZIP4-D1*) do not compensate for the absence of each other (Rey et al., 2017; Draeger et al., 2023). The absence of *TaZIP4-B2* expression in ethyl methanesulfonate-induced TILLING *Ph1* mutants does not cause an equivalent increase in the expression of *ZIP4* homologs on homoeologous group 3 chromosomes (Rey et al., 2017). Cytogenetic analysis of chiasmata formation in TILLING mutants focusing on three copies of *ZIP4* genes in tetraploid wheat has shown that *Ttzip4-A1* produced a phenotype that is almost identical to wild wheat (Draeger et al., 2023). Significant reduction in the chiasmata count by 10 % occurs in the *Ttzip4-B1* and *Ttzip4-B2* single mutants, as well as in the *Ttzip4-A1B2* and *Ttzip4-B1B2* double mutants, but the differences between them are insignificant with only an average of 1–2 extra univalents per cell (Draeger et al., 2023). Crossovers in the *Ttzip4-A1B1* double mutants (with a single *TtZIP4-B2* copy) are reduced by 76–78 %, and the plants frequently become sterile (Draeger et al., 2023). The *TaZIP4* copies on group 3 chromosomes are also predominantly required for homologous crossovers in hexaploid wheat (Martín et al., 2021).

A set of genes is identified on chromosome 3B, of which at least eight (*CAP-E1/E2*, *DUO1*, *MLH1*, *MPK4*, *MUS81*, *RTEL1*, *SYN4*, *ZIP4*) were confirmed to be involved in recombination process (Darrier et al., 2022). Three copies of genes *CAP-E1/E2*, *MLH1*, and *MPK4-3* were characterized by the highest expression levels, while *ZIP4* expression level was significantly lower or equal to that of 3A, 3B, and 3D homoeologs. As a result, *MPK4*, *CAP-E1/E2*, and *MLH1* were picked as candidate genes responsible for chiasmata formation control (Darrier et al., 2022).

Another distinctive feature of metaphase I in the 3R(3B) line was the presence of meiocytes with decompacted chromosomes. The *AtCAP-E1*+/– and *AtCAP-E2*–/– heterozygous double mutants of Arabidopsis turned out to be the closest ones in terms of meiotic phenotype, where the *CAP-E1/E2* gene acted like a functional ortholog of the *SMC2* (*Structural Maintenance of Chromosomes 2*) gene, a subunit of the condensin complex involved in chromosome compaction (Sutani et al., 1999). The analysis of mutants showed the expression of these genes during meiosis, and heterozygous double mutants demonstrated reduced chromosome condensation at metaphase I and anaphase I (Siddiqui et al., 2003). Some authors (Darrier et al., 2022) consider anomalous condensin activity an additional factor contributing to crossover disruption.

The presence of meiotic genes on chromosome 3B was further proved by QTL mapping of *QTdes2.ndsu-3B* responsible for desynapsis in durum wheat plants with chromosome 3B long arm deletion caused by radiation exposure (Bassi et al., 2013). However, the nucleotide sequence for this deletion has not been sequenced to date and cannot be compared to the sequences of the known genes.

Our study has also demonstrated that rye chromosome 3R does not compensate for the ability of chromosome 3B to ensure normal formation of crossovers between homologs. Asynapsis between homologs as a result of chromosome 3B substitution with wheat or rye homoeologs was demonstrated earlier (Lee et al., 1970; Bassi et al., 2013). Up to 14 univalents were formed in a 3D(3B) substitution line of durum wheat, the Langdon variety, at metaphase I (Bassi et al., 2013). Substitution of wheat chromosome 3B with rye chromosome 3R in the Kharkovskaya-Dakold bread wheat line caused asynapsis between homologs in 30 % of meiocytes (Lee et al., 1970). However, the addition of a pair of rye chromosomes 3R into the karyotype of F_1 wheat-rye hybrids increased the number of bivalents at metaphase I (Lelley, 1976; Miller et al., 1983), while the lowest reduction of chiasmata count of 1.1 % was produced by chromosome 3R in the Chinese Spring-Imperial addition line (Orellana et al., 1984).

Chromosome 3R(3B) substitution causes various meiotic division abnormalities

Meiosis in the 3R(3B) line was characterized by a number of abnormalities in MT dynamics and chromosome behavior in the first and second divisions. These results can be explained by the earlier data on the co-expression of *Ttzip4-B1* and meiotic genes orthologs (Alabdullah et al., 2019). During the construction of co-expression network for the orthologs of known meiotic wheat genes associated with *TaZIP4*, three *TaZIP4* homoeologs on group 3 chromosomes 3A, 3B, and 3D (*TraesCS3A02G401700*, *TraesCS3B02G434600* and *TraesCS3D02G396500*) were clustered in the largest meiosis-related module and significantly linked to many orthologs of meiotic genes with various functions as follows: association of sister kinetochores in the first meiotic division, chromosome segregation, formation of class I and II crossovers, protection of the cohesin complex in the centromeric region, control of the meiotic cell cycle, sister chromatid cohesion, double-strand break DNA repair, synaptonemal complex, anti-crossover activity, and double-strand break formation in DNA (Alabdullah et al., 2019). However, the *TaZIP4* copy responsible for the *Ph1* phenotype (*TraesCS5B02G255100*) was not clustered in the same module (Alabdullah et al., 2019), which also confirms its alternative expression profile (Martín et al., 2018).

Our study has discovered anomalies in MT cytoskeleton dynamics in the 3R(3B) line. At metaphase I, we observed the disruptions in MT nucleation at kinetochores of certain bivalents or MT convergence at the pole, which could cause the absence of equator plate in 53.33 ± 14.62 % of meiocytes. We also observed the formation of an arc-like structure by the cytoskeleton, when the nucleus migrated toward the nuclear envelope at pachytene. A possible cause for that could be the absence of the *MPK4* (*mitogen-activated protein kinase*) gene identified on chromosome 3B (Darrier et al., 2022) and involved in MT cytoskeleton dynamics (Beck et al., 2010; Zheng et al., 2011).

Asynchronous chromosome behavior in the second division and the presence of dyads at telophase II was a significant meiosis feature in the 3R(3B) line. This meiotic phenotype matched *TAM* mutants (*tam 1*, *tam2*), where *tam1* demonstrated asynchronous meiotic division, and *tam2*, the absence of the second division and subsequent meiotic restitution. *QTug.sau-3B*, a QTL responsible for unreduced gamete production in interspecific hybrids, was identified on chromosome 3B (Hao et al., 2014) and turned out to be syntenic for the *TAM* locus in rice and *Brachypodium*, while in *Arabidopsis thaliana*, *TAM* codes for CYCA1;2 cyclin.

The absence of wheat chromosome 3B is not the only possible cause of meiotic division disturbances in the 3R(3B) line. At present, changes in gene expression levels have been detected both in wheat-alien chromosome substitution and addition lines (Rey et al., 2018; Dong et al., 2020). Disturbances in chromosome behavior in the bread wheat lines introgressed with alien chromosomes are made possible due to the formation of a new genetic background where gene expression levels change in both wheat recipients and alien donors (Rey et al., 2018; Dong et al., 2020). For instance, changes in gene expression levels were detected in all wheat chromosomes in the TA3575 line where chromosome 3B was substituted with 3Sl#2 of *Ae. longissima* (Dong et al., 2020). Transcriptome analysis showed changes in gene expression in 577 out of 1,839 genes mapped on chromosome 3B of the Chinese Spring variety (31.43 %). Most of these genes (461, 79.90 %) were not transcribed, and 100 genes (17.33 %) demonstrated reduced expression, whereas only 16 (2.77 %) genes showed increased expression. It shows that at least 34.57 % (461 out of 1,839) of the genes on the absent chromosome 3B were not genetically compensated for by introgression of chromosome 3Sl#2 of *Ae. longissimi* (Dong et al., 2020).

Conclusions

Introgression of genetic material from relatives in the form of chromosomes or their fragments into bread wheat genome is widely used in wheat breeding to transfer genes controlling valuable agronomic traits. Successful transfer of these chromosomes during hybridization is reliant on meiotic behavior of both wheat and alien chromosomes. When a wheat chromosome, the genes of which are involved in meiotic division regulation, is substituted with an alien one, the presence/absence of a compensatory effect may be observed in genes on homoeologous chromosomes of relative species and genera. It was shown earlier that wheat chromosome 3B also harbored genes regulating bivalent formation independently of 5B and that *Ttzip4-B1* was co-expressed with orthologs of meiotic genes.

In our study, we have investigated meiotic MT cytoskeleton dynamics and chromosome behavior in the 3R(3B) line with wheat chromosome 3B substituted with rye chromosome 3R. The effect of 3R(3B) substitution manifested itself not only in reduced chiasmata count compared to the S29 variety (34.9 ± 0.62 and 41.92 ± 0.38, respectively), but also in a series of anomalies in MT dynamics and chromosome behavior in the first and second divisions. The disturbances had to do with

MT nucleation at kinetochores, MT convergence at meiotic spindle poles, C-shaped spindle formation, cell wall construction, cytomixis, as well as asynchronous second division and the presence of dyads at telophase II. Thus, the results obtained show that chromosome 3B of the Saratovskaya 29 variety is involved in regulation of a series of meiotic processes, and rye chromosome 3R lacks a genetic compensatory ability to functionally replace 3B in terms of normal meiotic progression.

References

Alabdullah A.K., Borrill P., Martin A.C., Ramirez-Gonzalez R.H., Hassani-Pak K., Uauy C., Shaw P., Moore G. A co-expression network in hexaploid wheat reveals mostly balanced expression and lack of significant gene loss of homeologous meiotic genes upon polyploidization. *Front. Plant Sci.* 2019;10:1325. DOI 10.3389/fpls.2019.01325

Al-Kaff N., Knight E., Bertin I., Foote T., Hart N., Griffiths S., Moore G. Detailed dissection of the chromosomal region containing the *Ph1* locus in wheat *Triticum aestivum*: with deletion mutants and expression profiling. *Ann. Bot.* 2008;101(6):863-872. DOI 10.1093/aob/mcm252

Appels R., Eversole K., Stein N., Feuillet C., Keller B., Rogers J., Pozniak C.J., Choulet F., Distelfeld A., Poland J., … Visendi P., Cui L., Du X., Feng K., Nie X., Tong W., Wang L. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science.* 2018;361(6403):eaar7191. DOI 10.1126/science.aar7191

Bariah I., Keidar-Friedman D., Kashkush K. Where the wild things are: transposable elements as drivers of structural and functional variations in the wheat genome. *Front. Plant Sci*. 2020;11:585515. DOI 10.3389/fpls.2020.585515

Bassi F.M., Kumar A., Zhang Q., Paux E., Huttner E., Kilian A., Dizon R., Feuillet C., Xu S.S., Kianian S.F. Radiation hybrid QTL mapping of *Tdes2* involved in the first meiotic division of wheat. *Theor. Appl. Genet.* 2013;126(8):1977-1990. DOI 10.1007/s00122-013-2111-z

Beck M., Komis G., Muller J., Menzel D., Samaj J. *Arabidopsis* homologs of nucleus- and phragmoplast-localized kinase 2 and 3 and mitogen-activated protein kinase 4 are essential for microtubule organization. *Plant Cell.* 2010;22:755-771. DOI 10.1105/tpc.109.071746

Cook J.P., Wichman D.M., Martin J.M., Bruckner P.L., Talbert L.E. Identification of microsatellite markers associated with a stem solidness locus in wheat. *Crop Sci*. 2004;44(4):1397-1402. DOI 10.2135/cropsci2004.1397

Darrier B., Colas I., Rimbert H., Choulet F., Bazile J., Sortais A., Jenczewski E., Sourdille P. Location and identification on chromosome 3B of bread wheat of genes affecting chiasma number. *Plants*. 2022;11(17):2281. DOI 10.3390/plants11172281

Dong Z., Ma C., Tian X., Zhu C., Wang G., Lv Y., Friebe B., Li H., Liu W. Genome-wide impacts of alien chromatin introgression on wheat gene transcriptions. *Sci. Rep.* 2020;10(1):4801. DOI 10.1038/s41598-020-61888-1

Draeger T.N., Rey M.-D., Hayta S., Smedley M., Alabdullah A.K., Moore G., Martín A.C. *ZIP4* is required for normal progression of synapsis and for over 95 % of crossovers in wheat meiosis. *Front. Plant Sci.* 2023;14:1189998. DOI 10.3389/fpls.2023.1189998

Giorgi B. A homoeologous pairing mutant isolated in *Triticum durum* cv. Cappelli. *Mutat. Breed. Newslett.* 1978;11:4-5

Griffiths S., Sharp R., Foote T.N., Bertin I., Wanous M., Reader S., Colas I., Moore G. Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature*. 2006;439(7077):749-752. DOI 10.1038/nature04434

Hao M., Luo J., Zeng D., Zhang Li, Ning S., Yuan Z., Yan Z., Zhang H., Zheng Y., Feuillet C., Choulet F., Yen Y., Zhang L., Liu D. *QTug.sau-3B* is a major quantitative trait locus for wheat hexaploidization. *G3* (*Bethesda*). 2014;4(10):1943-1953. DOI 10.1534/g3.114.013078

Kato T., Yamagata H. Effect of 3B chromosome deficiency on the meiotic pairing between the arms of an isochromosome in common wheat. *Jpn. J. Genet.* 1982;57(4):403-406. DOI 10.1266/jjg.57.403

Kroupin P.Yu., Divashuk M.G., Karlov G.I. Gene resources of perennial wild cereals involved in breeding to improve wheat crop. *Selskokhozyaystvennaya Biologiya = Agricultural Biology*. 2019;54(3):409-425. DOI 10.15389/agrobiology.2019.3.409eng

Lee Y.H., Larter E.N., Evans L.E. Meiotic behavior of wheat-rye addition and substitution lines. *Crop Sci.* 1970;10(2):144-145. DOI 10.2135/cropsci1970.0011183X001000020006x

Lelley T. Induction of homeologous pairing in wheat by genes of rye suppressing chromosome 5B effect. *Can. J. Genet. Cytol.* 1976;18(3):485-489. DOI 10.1139/g76-059

Ling H.Q., Ma B., Shi X., Liu H., Dong L., Sun H., Cao Y., Gao Q., Zheng Sh., Li Ye, Yu Y., Du H., Qi M., Li Y., Lu H., Yu H., Cui Y., Wang N., Chen Ch., Wu H., Zhao Y., Zhang J., Li Y., Zhou W., Zhang B., Hu W., van Eijk M.J.T., Tang J., Witsenboer H.M.A., Zhao Sh., Li Zh., Zhang A., Wang D., Lian Ch. Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*. *Nature.* 2018;557(7705):424-428. DOI 10.1038/s41586-018-0108-0

Loginova D.B., Zhuravleva A.A., Silkova O.G. Random chromosome distribution in the first meiosis of F1 disomic substitution line 2R(2D) × rye hybrids (ABDR, 4x = 28) occurs without bipolar spindle assembly. *Comp. Cytogenet.* 2020;14(4):453-482. DOI 10.3897/compcytogen.v14.i4.55827

Martín A.C., Rey M.-D., Shaw P., Moore G. Dual effect of the wheat *Ph1* locus on chromosome synapsis and crossover. *Chromosoma*. 2017;126(6):669-680. DOI 10.1007/s00412-017-0630-0

Martín A.C., Borrill P., Higgins J., Alabdullah A., Ramírez-González R.H., Swarbreck D., Uauy C., Shaw P., Moore G. Genome-wide transcription during early wheat meiosis is independent of synapsis, ploidy level, and the *Ph1* locus. *Front. Plant Sci.* 2018;9:1791. DOI 10.3389/fpls.2018.01791

Martín A.C., Alabdullah A.K., Moore G. A separation-of-function *ZIP4* wheat mutant allows crossover between related chromosomes and is meiotically stable. *Sci. Rep.* 2021;11(1):21811. DOI 10.1038/s41598-021-01379-z

McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers W.J., Morris C., Appels R., Xia X.C. Catalogue of Gene Symbols for Wheat. Yokohama, 2013

Mello-Sampayo T. Genetic regulation of meiotic chromosome pairing by chromosome 3D of *Triticum aestivum*. *Nat. New Biol.* 1971;230(9):22-23. DOI 10.1038/newbio230022a0

Miller T.E., Reader S.M., Gale M.D. The effect of homoeologous group 3 chromosomes on chromosome pairing and crossability in *Triticum aestivum*. *Can. J. Genet. Cytol.* 1983;25(6):634-641. DOI 10.1139/g83-094

Mohammed Y.S., Tahir I.S., Kamal N.M., Eltayeb A.E., Ali A.M., Tsujimoto H. Impact of wheat-*Leymus racemosus* added chromosomes on wheat adaptation and tolerance to heat stress. *Breed. Sci*. 2014;63(5):450-460. DOI 10.1270/jsbbs.63.450

Munkvold J.D., Greene R.A., Bermudez-Kandianis C.E., La Rota C.M., Edwards H., Sorrells S.F., Dake T., Benscher D., Kantety R., Linkiewicz A.M., … Kianian S.F., Sandhu D., Erayman M., Gill K.S., McGuire P.E., Qualset C.O., Sorrells M.E. Group 3 chromosome bin maps of wheat and their relationship to rice chromosome 1. *Genetics*. 2004;168(2):639-650. DOI 10.1534/genetics.104.034819

Orellana J., Cermeño M.C., Lacadena J.R. Meiotic pairing in wheat–rye addition and substitution lines. *Can. J. Genet. Cytol.* 1984;26(1):25-33. DOI 10.1139/g84-005

Paux E., Röger D., Badaeva E., Gay G., Bernard M., Sourdille P., Feuillet C. Characterizing the composition and evolution of homoeolo-

gous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* 2006;48:463-474. DOI 10.1111/j.1365-313X.2006.02891.x

Ramage R.T., Hernandez-Soriano J.M. Desynaptic genes in barley. *Barley Genet. Newslett.* 1972;2:65-68

Rey M.-D., Martín A.C., Higgins J., Swarbreck D., Uauy C., Shaw P., Moore G. Exploiting the *ZIP4* homologue within the wheat *Ph1* locus has identified two lines exhibiting homoeologous crossover in wheat-wild relative hybrids. *Mol. Breed.* 2017;37(8):95. DOI 10.1007/s11032-017-0700-2

Rey M.-D., Martín A.C., Smedley M., Hayta S., Harwood W., Shaw P., Moore G. Magnesium increases homoeologous crossover frequency during meiosis in *ZIP4* (*Ph1* gene) mutant wheat-wild relative hybrids. *Front. Plant Sci.* 2018;9:509. DOI 10.3389/fpls.2018.00509

Sears E.R. The Aneuploids of Common Wheat. Columbia MO, USA: Agric. Exp. Station, Univ. of Missouri, 1954

Sears E.R. Genetics Society of Canada award of excellence lecture an induced mutant with homoeologous pairing in common wheat. *Can. J. Genet. Cytol.* 1977;19(4):585-593. DOI 10.1139/g77-063

Siddiqui N.U., Stronghill P.E., Dengler R.E., Hasenkampf C.A., Riggs C.D. Mutations in *Arabidopsis* condensin genes disrupt embryogenesis, meristem organization and segregation of homologous chromosomes during meiosis. *Development*. 2003;130(14):3283-3295. DOI 10.1242/dev.00542

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

Sutani T., Yuasa T., Tomonaga T., Dohmae N., Takio K., Yanagida M. Fission yeast condensin complex: essential roles of non-SMC subunits for condensation and Cdc2 phosphorylation of Cut3/SMC4. *Genes Dev.* 1999;13(17):2271-2283. DOI 10.1101/gad.13.17.2271

Wada B., Kusunoki F. Spindle membrane in meiosis of pollen mother cells of *Tradescantia* and in mitosis of endosperm cells of *Zephyranthes*. *Cytologia* (*Tokyo*). 1964;29(1):109-117. DOI 10.1508/cytologia.29.109

Yudina R.S., Leonova I.N., Salina E.A., Khlestkina E.K. Effect of alien genomic introgressions on the osmotic tolerance of wheat. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2014;18(4/1):643-649 (in Russian)

Zeng Q., Chen J.-G., Ellis B.E. AtMPK4 is required for male-specific meiotic cytokinesis in Arabidopsis. *Plant J.* 2011;67(5):895-906. DOI 10.1111/j.1365-313X.2011.04642.x

Zhang P., Wanlong L., Fellers J., Friebe B., Gill B.S. BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. *Chromosoma*. 2004;112(6):288-299. DOI 10.1007/s00412-004-0273-9

Conflict of interest. The authors declare no conflict of interest.
Received January 12, 2024. Revised March 25, 2024. Accepted March 26, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):377-386

DOI 10.18699/vjgb-24-43

Molecular cytogenetic characteristics of new spring bread wheat introgressive lines resistant to stem rust

O.A. Baranova [1], I.G. Adonina [2], S.N. Sibikeev [3]

[1] All-Russian Institute of Plant Protection, St. Petersburg-Pushkin, Russia
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Federal Center of Agricultural Research of the South-East Region, Saratov, Russia
baranova_oa@mail.ru

Abstract. Anticipatory wheat breeding for pathogen resistance is key to preventing economically significant crop losses caused by diseases. Recently, the harmfulness of a dangerous wheat disease, stem rust, caused by *Puccinia graminis* f. sp. *tritici*, was increased in the main grain-producing regions of the Russian Federation. At the same time, importation of the Ug99 race (TTKSK) is still a possibility. In this regard, the transfer of effective resistance genes from related species to the bread wheat breeding material followed by the chromosomal localization of the introgressions and a marker analysis to identify known resistance genes is of great importance. In this work, a comprehensive analysis of ten spring bread wheat introgressive lines of the Federal Center of Agricultural Research of the South-East Region (L657, L664, L758, L935, L960, L968, L971, L995/1, L997 and L1110) was carried out. These lines were obtained with the participation of *Triticum dicoccum*, *T. timopheevii*, *T. kiharae*, *Aegilops speltoides*, *Agropyron elongatum* and *Secale cereale*. In this study, the lines were evaluated for resistance to the Ug99 race (TTKSK) in the Njoro, Kenya. Evaluation of introgression lines in the field for resistance to the Ug99 race (TTKSK) showed that four lines were immune, two were resistant, three were moderately resistant, and one had an intermediate type of response to infection. By cytogenetic analysis of these lines using fluorescent (FISH) and genomic (GISH) *in situ* hybridization, introgressions from *Ae. speltoides* (line L664), *T. timopheevii* (lines L758, L971, L995/1, L997 and L1110), *Thinopyrum ponticum* = *Ag. elongatum* (2*n* = 70) (L664, L758, L960, L971, L997 and L1110), as well as introgressions from *T. dicoccum* (L657 and L664), *T. kiharae* (L960) and *S. cereale* (L935 and L968) were detected. Molecular markers recommended for marker-oriented breeding were used to identify known resistance genes (*Sr2*, *Sr25*, *Sr32*, *Sr1A.1R*, *Sr36*, *Sr38*, *Sr39* and *Sr47*). The *Sr36* and *Sr25* genes were observed in lines L997 and L1110, while line L664 had the *Sr39*+*Sr47*+*Sr25* gene combination. In lines L935 and L968 with 3R(3D) substitution from *S. cereale*, gene resistance was presumably identified as *SrSatu*. Thus, highly resistant to both local populations of *P. graminis* and the Ug99 race, bread wheat lines are promising donors for the production of new varieties resistant to stem rust.

Key words: *Triticum aestivum* L.; introgressive wheat lines; alien introgressions; *Puccinia graminis* f. sp. *tritici*; Ug99; *Sr* genes.

For citation: Baranova O.A., Adonina I.G., Sibikeev S.N. Molecular cytogenetic characteristics of new spring bread wheat introgressive lines resistant to stem rust. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):377-386. DOI 10.18699/vjgb-24-43

Funding. The work was supported by the Russian Science Foundation grant No. 22-26-00172 "Biological justification of genetic protection of wheat against stem rust in the Volga region".

Acknowledgements. Cytological analysis was performed at the Center for Microscopic Analysis of Biological Objects of the Siberian Branch of the Russian Academy of Sciences with the support of the budget project FWNR-2022-0017.
The authors thank the reviewers for their contributions to the peer review of this work.

Молекулярно-цитогенетическая характеристика новых интрогрессивных линий яровой мягкой пшеницы, устойчивых к стеблевой ржавчине

О.А. Баранова [1], И.Г. Адонина [2], С.Н. Сибикеев [3]

[1] Всероссийский научно-исследовательский институт защиты растений, Санкт-Петербург, Пушкин, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Федеральный аграрный научный центр Юго-Востока, Саратов, Россия
baranova_oa@mail.ru

Аннотация. Опережающая селекция пшеницы на устойчивость к патогенам – залог предотвращения экономически значимых потерь урожая от болезней. В последние годы в основных зернопроизводящих областях Российской Федерации наблюдается увеличение вредоносности опасного заболевания пшеницы – стеблевой

ржавчины (возбудитель *Puccinia graminis* f. sp. *tritici*). В то же время сохраняется опасность заноса на территорию России расы патогена Ug99 (TTKSK), которая угрожает производству зерна во всем мире. В связи с этим большое значение приобретают перенос эффективных генов резистентности от родственных видов в селекционный материал мягкой пшеницы, выявление хромосомной локализации интрогрессий и проведение маркерного анализа для идентификации известных генов устойчивости. В настоящей работе был проведен комплексный анализ десяти интрогрессивных линий яровой мягкой пшеницы селекции Федерального аграрного научного центра Юго-Востока (Л657, Л664, Л758, Л935, Л960, Л968, Л971, Л995/1, Л997 и Л1110), полученных с участием *Triticum dicoccum*, *T. timopheevii*, *T. kiharae*, *Aegilops speltoides*, *Agropyron elongatum* и *Secale cereale*. Оценка интрогрессивных линий в полевых условиях на устойчивость к расе Ug99 (TTKSK) показала, что четыре линии были иммунны, две – устойчивы, три – среднеустойчивы, а одна имела промежуточный тип реакции на заражение. Цитогенетический анализ с помощью методов флуоресцентной (FISH) и геномной (GISH) гибридизации *in situ* выявил интрогрессии от *Ae. speltoides* (линия Л664), *T. timopheevii* (линии Л758, Л971, Л995/1, Л997 и Л1110), *Thinopyrum ponticum* = *Ag. elongatum* (2*n* = 70) (Л664, Л758, Л960, Л971, Л997 и Л1110), а также интрогрессии от *T. dicoccum* (Л657 и Л664), *T. kiharae* (Л960) и *S. cereale* (Л935 и Л968). Для идентификации известных генов устойчивости (*Sr2*, *Sr25*, *Sr32*, *Sr1A.1R*, *Sr36*, *Sr38*, *Sr39* и *Sr47*) использовали молекулярные маркеры, рекомендованные для маркер-ориентированной селекции. Наличие генов *Sr36* и *Sr25* было постулировано у двух линий (Л997 и Л1110), генов *Sr39*, *Sr25* и *Sr47* – у линии Л664. У линий Л935 и Л968 с замещением 3D(3R) от *S. cereale* ген устойчивости к стеблевой ржавчине предположительно определен как *SrSatu*. Высокоустойчивые как к местным популяциям *P. graminis*, так и к расе Ug99 линии мягкой пшеницы являются перспективными донорами для создания новых устойчивых к стеблевой ржавчине сортов.

Ключевые слова: *Triticum aestivum* L.; интрогрессивные линии пшеницы; чужеродные интрогрессии; *Puccinia graminis* f. sp. *tritici*; Ug99; *Sr* гены.

Introduction

One of the conditions for increasing the yield of bread wheat is the production of varieties that are resistant to biotic and abiotic stressors. The set of the most harmful biostressors for bread wheat includes a group of rust disease pathogens: *Puccinia triticina* f. sp. *tritici* Erikss., *P. striiformis* f. sp. *tritici* Erikss., *P. graminis* f. sp. *tritici* Erikss. & Henning. These pathogens cause epiphytoties of brown, yellow and stem rust. The harmfulness of each of them can reach 50 % (Knott, 1989). The causative agents of these diseases are characterized by high virulence and great diversity in racial composition (Gultyaeva et al., 2021, 2022; Baranova et al., 2023b).

In the global production of bread wheat and under Russian conditions, a special place is occupied by stem rust (pathogen *P. graminis* f. sp. *tritici* (*Pgt*)), which can cause yield losses of more than 80 % during epiphytotic development on susceptible varieties. The well-known race of stem rust pathogen Ug99 (TTKSK) and its variants, which infect wheat varieties and lines with effective resistance genes *Sr31*, *Sr36* and *Sr24*, still pose a real threat to wheat production in the regions of the African continent, the Middle East and Asia. Due to the possibility of fungal spores spreading with air masses over vast distances, a threat of the pathogen being introduced into the territory of Eurasian countries, including Russia, remains. Over the last decade, in Europe, Kazakhstan, China and the Russian Federation, aggressive races of the fungus have appeared that are not related to the Ug99 race, but have caused severe outbreaks of the disease (Vasilova et al., 2017; Lewis et al., 2018; Baranova et al., 2021; Patpour et al., 2022).

Low diversity of stem rust resistance genes is a common problem in commercial wheat varieties around the world. The adult resistance gene *Sr57* (*Lr34/Yr18/Pm38/Bdv1*), which is part of a locus with pleiotropic action that determines nonspecific resistance to biotrophic pathogens, as well as juvenile resistance genes such as *Sr38*, *Sr6Agi*, *Sr25* and *Sr31* are used in Russian domestic varieties. The *Sr31* gene still remains effective against stem rust in the Russian Federation (Baranova et al., 2023b). Genes *Sr6Agi* and *Sr25* lose effectiveness in the Volga region, but are effective against Western Siberian populations of the fungus (Kelbin et al., 2020; Baranova et al., 2021). The *Sr38* gene is ineffective against Volga populations of the pathogen, but is recommended for breeding in Western Siberia (Skolotneva et al., 2021).

To expand the genetic basis of varieties, it is extremely important to obtain breeding material diverse for resistance genes. In general, this problem is solved by involving related species of bread wheat, mainly from the secondary and tertiary gene pools. Currently, 26 out of 63 stem rust resistance genes have been transferred from the genomes of related species (McIntosh et al., 2013, 2022). The *Ae. speltoides*, *T. timopheevii*, *T. dicoccum*, *T. ponticum*, *S. cereale* species remain important sources of valuable genes for resistance to fungal diseases and in particular to stem rust for practical breeding of bread wheat (McIntosh et al., 2013). Genes *Sr32*, *Sr39*, *Sr47* were transferred from *Aegilops speltoides* (Taush) (SS, 2*n* = 14) to the wheat genome; *Sr36*, *Sr37*, *Sr40*, from *Triticum timopheevii* Zhuk. (A^tA^t GG, 2*n* = 28); *Sr31*, *Sr27*, *Sr1A.1R*, *Sr50*, from *Secale cereale* L. (RR, 2*n* = 14) (McIntosh et al., 2013). Effectiveness against *P. graminis* and the nature and size of the introgressed material are important aspects of using these genes to develop resistant bread wheat varieties. It is important to produce combinations of currently effective *Sr* genes with each other or with genes that have partially lost their effectiveness, or with adult resistance genes.

At the Federal Center of Agricultural Research of the South-East Region (FCAR of the South-East Region), work is underway to produce new breeding material using relatives of bread wheat. Previously, lines produced with the participation of a wide range of species showed high resistance to leaf rust in the conditions of the Saratov Volga region (Gultyaeva et

Table 1. Pedigree of spring bread wheat introgressive lines

Line number	Pedigree	Source of alien genetic material
L657	L505*2//L503/3/L528//AD *T.dic*/*Ae.spelt**5S29/4/Thatcher *Lr28*	*Triticum dicoccum* Shuebl (BA), *Aegilops speltoides* Tausch (S)
L664	S55//Dobrynya/L164//Agr139/L528*2//AD *T.dic*/*Ae.spelt**5 S29//Dobrynya	*T. dicoccum* Shuebl (BA), *Ae. speltoides* Tausch (S), *Agropyron elongatum* (Host) Beauv. (source – Dobrynya variety as a carrier of translocation 7DS-7DL-7Ae#1L)
L758	L XI S29imm/L2870	*T. timopheevii*, *Ae. tauschii* (source – immune lines of Saratovskaya 29 variety (S29imm))
L935	Satu/S70//S70	*Secale cereale* L. (R) (source – triticale variety Satu)
L960	S68/*T.kiharae*//S70/3/S68	*T. kiharae* Dorof. et Migusch ((GA^{t}D)
L968	Satu/S70//S74/3/S70/4/S70	*S. cereale* L. (R), triticale variety Satu
L971	S68/*T.timopheevii**4// Dobrynya	*T. timopheevii* Zhuk. (GA^t), *Ag. elongatum* (Host) Beauv. (source – Dobrynya variety as a carrier of translocation 7DS-7DL-7Ae#1L)
L995/1	S70/Pamyati Maistrenko//S68	*T. timopheevii* and *Ae. tauschii* (source – Pamyati Maistrenko variety)
L997	S70/Pamyati Maistrenko// Dobrynya	*T. timopheevii* and *Ae. tauschii* (source – Pamyati Maistrenko variety); *Ag. elongatum* (Host) Beauv. – source of 7DS-7DL-7Ae#1L translocation, Dobrynya variety
L1110	L VI S29imm/L2032//L2032/3/L2032	*T. timopheevii*, *Ae. tauschii* (source – immune lines of Saratovskaya 29 variety (S29imm)); *Ag. elongatum* (Host) Beauv. – L2032 7DS-7DL-7Ae#1L translocation from *Ag. elongatum* (Host) Beauv.

Note. The pedigree lines indicate the following varieties of spring bread wheat: L503, L505, Dobrynya, Saratovskaya 29 (S29), Saratovskaya 55 (S55), Saratovskaya 68 (S68), Saratovskaya 70 (S70), Saratovskaya 74 (S74), as well as lines L164, L528, L2870, L2032, Agr139, L VI S29 imm, L XI S29 imm of spring bread wheat.

al., 2020). The aim of our work is a comprehensive study of new introgressive lines including assessment of resistance to the Ug99 (TTKSK) stem rust race, chromosomal localization of alien introgressions and identification of *Sr* genes using molecular markers.

Materials and methods

Plant material. Ten introgressive lines of spring bread wheat from FCAR of the South-East Region were studied. Their pedigree, indicating the donor of alien genetic material, is given in Table 1.

Cytogenetic analysis. Preparations of mitotic chromosomes were prepared from the meristem of seedling roots in accordance with the method (Badaeva et al., 2017). The FISH (fluorescence *in situ* hybridization) method using probes based on various repetitive sequences: Spelt1 (Salina et al., 1997) and Spelt52 (Salina et al., 2004), pSc119.2 (Bedbrook et al., 1980) and apAs1 (Rayburn, Gill, 1986) was used to analyze the karyotype of the lines. The FISH method described in the work of Salina et al. (Salina et al., 2006) with minor modifications was used. GISH (genomic *in situ* hybridization) using labeled *S. cereale* genomic DNA as a probe was performed according to previously published work (Schubert et al., 1998). The preparations were analyzed using an Axio Imager M1 microscope (Zeiss, Germany) equipped with a ProgRes MF CCD digital camera and Isis software (Meta Systems, Germany).

Phytopathological analysis. Resistance to race Ug99 (TTKSK) analysis was carried out at the adult plant stage using a modified Cobb scale (Peterson et al., 1948) in 2023 at the plant pathology nurseries at the International Maize and Wheat Improvement Center (CIMMYT) at the Kenya Agricultural and Livestock Research Organization (KALRO) in Njoro. The main distinguishing feature of the Ug99 race pathotypes is virulence towards carriers of the *Sr31* gene. The degree of damage to varieties with the *Sr31* gene in KALRO plant pathology nurseries in the growing season of 2023 was: for the variety Prokhorovka (*Sr31*) – 60 % (60MSS), for the variety Yugo-Vostochnaya 2 (*Sr31*) – 80 % (80S), for the variety Saratovskaya 74 (without *Sr* genes) – 80 % (80S).

Molecular genetic analysis. DNA was isolated from five-day-old wheat seedlings using cetyltrimethylammonium bromide (CTAB method) (Murray, Thompson, 1980). To identify the resistance genes *Sr2*, *Sr32*, *Sr1A.1R*, *Sr36*, *Sr38*, *Sr39*, *Sr47*, DNA markers recommended for marker-assisted selection (MAS) were used. A list of molecular markers used in the work with links to sources is presented in the Supplementary Material 1[1]. PCR was performed in duplicate on a C1000 Thermal Cycler (manufactured by BioRad). Amplification products were separated on 2 % agarose and 8 % polyacrylamide gels stained with ethidium bromide. Isogenic lines and varieties with known *Sr* genes served as a positive control; the susceptible variety Khakasskaya served as a negative control. PCR mixture was taken without adding DNA to control contamination. GeneRulerTM 50bp DNA Ladder (Thermo Scientific) was used as a molecular weight marker. Visualization of amplification products was carried out using the ChemiDoc™ (Bio-Rad) gel documentation system.

[1] Supplementary Materials 1–5 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Baranova_Engl_28_4.pdf

Results

Phytopathological analysis of spring bread wheat introgressive lines

Phytopathological screening of the lines at the stage of adult plants showed that all lines were resistant to the Ug99 race to varying degrees: four lines were immune (infection type 0), two were resistant (R), three were moderately resistant (MR) to this highly aggressive race of the fungus (Table 2). The only exception was one line, L995/1, which had an intermediate infection type (M) with 5 % of disease development.

Cytogenetic analysis of spring bread wheat introgressive lines

The identification of alien genetic material and determination of its state in the reconstructed genome of bread wheat in the form of addition or substitution chromosomes and translocations were the aim of the introgression lines cytogenetic analysis.

The main results of cytogenetic analysis are presented in Table 2 and in the Figure. Additional information indicating the probe combinations used is provided in the Supplementary Material 2.

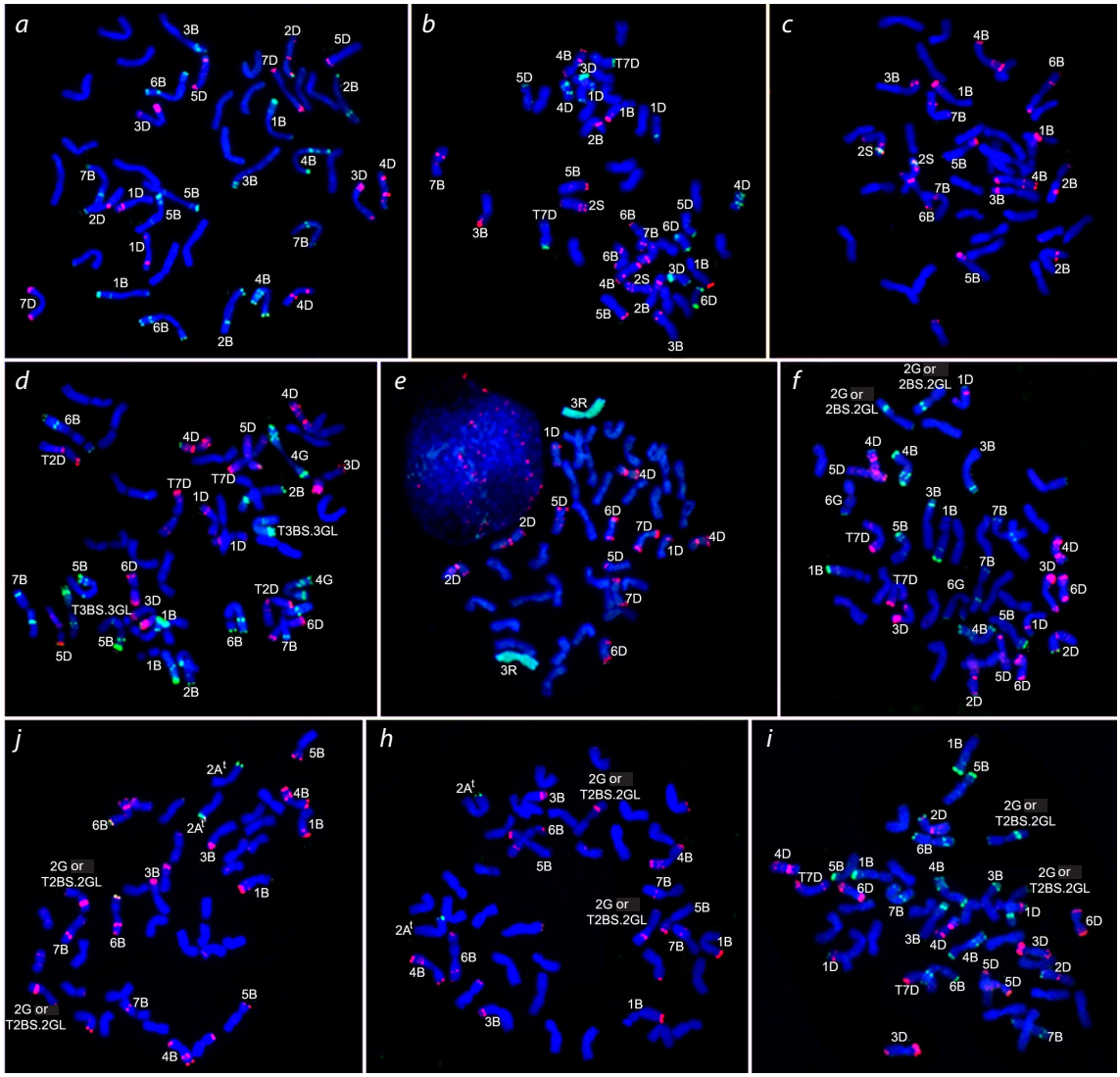
Karyotyping of the lines showed that each of them is characterized by the standard number of chromosomes for hexaploid wheat – 42. FISH was performed with probes pSc119.2 and pAs1 for each of the ten lines. The probe pSc119.2 (Bedbrook et al., 1980) is predominantly localized on the chromosomes of the bread wheat genome B, and pAs1 (Rayburn, Gill, 1986) is predominantly localized on the chromosomes of the D genome. The simultaneous use of these probes allows the identification of all chromosomes of the B and D genomes and some chromosomes of the A genome (Schneider et al., 2003). In addition, it is possible to identify the chromosomes of the G genome of *T. timopheevii* by the localization of hybridization signals with the pSc119.2 probe (Jiang, Gill, 1994). GISH with *S. cereale* DNA was used to analyze two wheat lines that had rye among ancestors. Analysis of eight lines, the ancestors of which included *Ae. speltoides*, *T. timopheevii* or *T. kiharae*, involved hybridization with probes Spelt1 and Spelt52 (performing GISH with DNA from these species is difficult due to their close relationship to bread wheat).

Table 2. Characteristics of spring bread wheat introgressive lines by translocations/substitutions, *Sr* genes and resistance to stem rust (Ug99) at the adult plants stage

Line	Cytogenetic study result	Identified *Sr* genes*	Resistance to *P. graminis* f. sp. *tritici* race Ug99 (TTKSK)**
L657	6A$^{T.\ dicoccum}$(6D)	–	5RMR
L664	2A$^{T.\ dicoccum}$(2A) or T2AS.2A$^{T.\ dicoccum}$L 2S(2D) – from *Ae. speltoides* Translocation from *Th. ponticum* to 7DL	*Sr25*, *Sr39*, *Sr47*	5RMR
L758	T2A^{t}.2A – from *T. timopheevii* Translocation from *Th. ponticum* to 7DL	*Sr25*	5R
L935	3R(3D) – from *S. cereale*	–	0
L960	2A^{t}(2A) – from *T. kiharae* T3BS.3GL – from *T. kiharae* 4G(4B) – from *T. kiharae* T2D$^{(T.\ aestivum)}$S.2D$^{(T.\ kiharae)}$L Translocation from *Th. ponticum* to 7DL	*Sr25*	5MR
L968	3R(3D) – from *S. cereale*	–	0
L971	2A^{t}(2A) – from *T. timopheevii* 2G(2B) or T2BS.2GL – from *T. timopheevii* 6G(6B) – from *T. timopheevii* Translocation from *Th. ponticum* to 7DL	*Sr25*	5R
L995/1	2A^{t}.2A – from *T. timopheevii* 2G(2B) or T2BS.2GL – from *T. timopheevii*	–	5M
L997	2A^{t}.2A – from *T. timopheevii* 2G(2B) or T2BS.2GL – from *T. timopheevii* Translocation from *Th. ponticum* to 7DL	*Sr25*, *Sr36*	0
L1110	2A^{t}.2A – from *T. timopheevii* 2G(2B) or T2BS.2GL – from *T. timopheevii* Translocation from *Th. ponticum* to 7DL	*Sr25*, *Sr36*	0

* The genes, the identification of which was confirmed by cytogenetics and pedigree analysis, are presented; the *Sr25* gene was identified previously (Baranova et al., 2023a).

** Resistance: 0 – immune infection type, R – resistant, MR – medium–resistant, RMR – intermediate type of infection between resistance and medium resistance, M – intermediate type of infection between medium resistance and medium susceptibility.



Results of FISH and GISH with different probe combinations on metaphase chromosomes of bread wheat introgressive lines.

Probes pSc119.2 (green), pAs1 (red): *a* – line L657, *b* – L664, *d* – L960, *f* – L971, *i* – L1110; probes Spelt52 (green), pSc119.2 (red): *c* – L664, *h* – L997; probes Spelt1 (green), pSc119.2 (red): *g* – L997; Rye DNA (green), pAs1 (red): *e* – line L968.

Lines L657 and L664 were obtained with the participation of *Ae. speltoides*. In the L657 line, Spelt52 repeat sites were not identified, and the Spelt1 probe is localized at the ends of the chromosome 6B arms. According to previous studies, this localization of Spelt1 occurs in bread wheat varieties (Salina et al., 2006). Consequently, we cannot speak with confidence about translocations from *Ae. speltoides* in this lineage. FISH with the pAs1 probe showed the absence of 2D chromosomes in the L664 line and revealed a pair of chromosomes with a weak signal of pSc119.2 on the short arm and two signals on the long arm (see the Figure, *b*), which also contains the Spelt52 site (see the Figure, *c*). We identified this chromosome as chromosome 2S of *Ae. speltoides* (Badaeva et al., 1996; Ruban, Badaeva, 2018). Thus, in the case of line L664, we have established chromosomal substitution 2S(2D). In addition, the results of hybridization of probes pSc119.2 and pAs1 on the chromosomes of the L657 line (see the Figure, *a*) indicate the substitution of chromosome 6D, presumably with chromosome 6A of *T. dicoccum*, a species that is present in the pedigree of this line. In the L664 line, Spelt1 sites were identified at the long arm ends of the genome A chromosomes pair, most likely chromosomes 2A. In bread wheat varieties, such localization of this probe has not been observed, but it is characteristic of tetraploid wheats, in particular *T. dicoccum*, and may indicate chromosomal substitution or translocation from a given species (present in the pedigree).

Lines L758, L960, L971, L995/1, L997, and L1110 were expected to have introgressions from the species *T. timopheevii* or *T. kiharae*. Interestingly, weak hybridization signals with the Spelt52 probe were detected on the short arms of the genome A chromosome pair, most likely chromosome 2A in all six lines (see the Figure, *h*). This localization of Spelt52 is characteristic of *T. timopheevii* or *T. kiharae* and may indicate translocations from these species. No hybridization signals with the Spelt1 probe were detected in the L758 line. Lines L995/1, L997 and L1110 carry Spelt1 blocks at the ends of the short arms of chromosome 6B (see the Figure, *g*), which is typical for a number of bread wheat varieties (Salina et al.,

2006). The localization of Spelt1 on the long arms of chromosome 2A in the lines L960, L971, L997 and L1110 (see the Figure, *g*) in combination with the localization of the Spelt52 probe on the short arms of these chromosomes (see the Figure, *h*) may indicate the substitution of chromosome 2A with $2A^t$ (from *T. timopheevii* or *T. kiharae*, respectively) in these lines. In line L960, another Spelt1 site is located on the long arm of the chromosome, which, according to the localization of the pSc119.2 probe, corresponds to chromosome 4G of *T. timopheevii*, while chromosome 4B is absent (see the Figure, *d*). The results obtained indicate that the L960 line has chromosomal substitution 4G(4B). Also, based on the localization of probe pSc119.2 in the L960 line, translocation T3BS.3GL can be assumed. The distribution of the pAs1 probe on the long arm of chromosome 2D in this line is almost identical to that in *T. kiharae*, which indicates a probable translocation $T2D^{(T.\ aestivum)}S.2D^{(T.\ kiharae)}L$ (see the Figure, *d*). It should be noted that in the L960 line, as well as in the L758, L664, L971, L997 and L1110 lines, the localization of the pAs1 probe on the long arm of chromosome 7D does not correspond to bread wheat, which indicates a translocation (see the Figure, *b*, *f*, *i*). The presence of the *Sr25* gene in four of these lines (Baranova et al., 2023a), transferred to the bread wheat genome from *Th. ponticum* (Friebe et al., 1996), and the hybridization pattern of pAs1 on the long arm of wheatgrass chromosome J^s-7 (Cui et al., 2018), allow us to conclude that chromosome 7D of these lines carries translocations from *Th. ponticum.*

In the case of the L995/1, L971 and L997, L1110 lines, according to the results of hybridization with the pSc119.2 probe, we can talk about the substitution of chromosome 2B with 2G *T. timopheevii*, 2G(2B), or about the translocation of T2BS.2GL (see the Figure, *f*–*i*). Additionally, the L971 line is expected to have a chromosomal substitution 6G(6B) (see the Figure, *f*).

Lines L935 and L968 were obtained using the Australian triticale variety Satu. GISH with rye DNA and FISH with probes pSc119.2 + pAs1 showed the substitution of 3D chromosomes with a pair of 3R chromosomes in these lines (see the Figure, *e*).

Identification of stem rust resistance genes using molecular markers

The results of *Sr* genes identification in the analyzed lines using molecular markers, confirmed by pedigree analysis and cytogenetic analysis data, are presented in Table 2. In this work, PCR fragments specific for genes *Sr32*, *Sr39*, *Sr47* (*Ae. speltoides*), *Sr36* (*T. timopheevii*) and *Sr38* (*Ae. ventricosa*) were found in different lines. All the obtained results of PCR analysis, indicating the molecular markers used, are given in the Supplementary Material 3. The diagnostic fragment of the VENTRIUP-LN2 marker for the *Sr38* gene was observed only in the L971 line (Supplementary Material 3). The presence of the *Sr36* gene was established in two lines, L997 and L1110, using the *Xstm773-2* marker (Table 2, Supplementary Material 3).

The *Sr39* gene was identified using the Sr39#22 marker. A diagnostic fragment (800 bp) was detected in five lines (Supplementary Materials 3 and 4). To identify the *Sr32* gene, the csSr32#2 marker was used. The diagnostic fragment was observed in three lines: L960, L968 and L995/1. The *Sr47* gene was identified using three markers – *Xgwm501*, *Xgpw4043*, and *Xgwm47* (Supplementary Materials 3 and 5). The diagnostic fragment of the *Xgwm501* marker (109 bp) was identified in four lines: L971, L995/1, L997 and L1110. Only the 95-bp fragment of the two diagnostic fragments of the *Xgpw4043* marker was amplified in the L657, L664, L758, and L971 lines; the 115-bp fragment was absent (Supplementary Material 5). The diagnostic fragment of the *Xgwm47* marker (165 bp) was identified only in line L664.

Previously, all the lines we analyzed were tested for the presence of the *Sr25* gene (Baranova et al., 2023a) using the Gb marker recommended for marker-based selection (Prins et al., 2001). This gene was identified in six lines (Table 2, Supplementary Material 3). According to the results of previous studies (Baranova et al., 2023a) and this work, the *Sr2*, *Sr24*, *Sr28*, *Sr31*, *Sr1A.1R* and *Sr57* genes were not found in any of the lines.

Discussion

Efficiency of molecular markers recommended for marker-based selection for identifying stem rust resistance genes

Molecular markers are widely used to identify resistance genes to various pathogens including stem rust. Among the huge number of molecular markers, the most specific ones are highlighted and recommended for marker-based selection (https://maswheat.ucdavis.edu/). However, during work with a variety of plant material, especially with introgressive lines, a researcher may encounter insufficient specificity of even a recommended marker and, as a result, false-positive gene identification. In this regard, it is desirable to conduct complex studies and confirm the presence of the desired gene along with molecular genetic analysis data, study of pedigrees, cytogenetic and phytopathological results.

During our work, introgression lines were analyzed cytogenetically and using molecular markers. Data from pedigree lines were also taken into account.

In six out of the ten studied lines (L664, L758, L960, L971, L997 and L1110), the *Sr25* gene was previously identified (Baranova et al., 2023a), which was fully confirmed by the cytogenetic analysis data in this work (Table 2). The *Sr25* gene is linked to the leaf rust resistance gene *Lr19* and is localized in the T7DS-7DL-7Ae#1L translocation from *Th. ponticum* (Friebe et al., 1994).

The identification of the *Sr36* gene using the *Xstm773-2* marker is also confirmed by cytogenetic analysis. As is known, the *Sr36* gene is localized on chromosome 2G (Friebe et al., 1996). Lines L997 and L1110, in which this gene was identified according to molecular genetic analysis, carry chromosome 2G from *T. timopheevii* (see the Figure, *g*–*i*, Table 2).

We obtained ambiguous results regarding the resistance genes *Sr32, Sr39* and *Sr47*, the source of which is *Ae. speltoides*. Cytogenetic analysis revealed genetic material from *Ae. speltoides* (substitution of chromosome 2D with chromosome 2S – 2S(2D) only in line L664, in the pedigree of which this species is present). However, the diagnostic fragment of the Sr39#22 marker (the *Sr39* gene marker) was also identified in lines L971, L995/1, L997 and L1110 (Supplementary Material 3), which lack *Ae. speltoides* in their pedigrees,

but there is genetic material from *T. timopheevii*, which was confirmed cytogenetically. Also, based on the pedigrees of the L997 and L1110 lines, Dr. Savov's synthetic (GA^{t}D) was used in crosses (*T. timopheevii* × *T. tauschii*). Thus, the diagnostic fragment of the Sr39#22 marker was amplified in lines with material from *T. timopheevii* and possibly *T. tauschii*. It should be noted that similar results for the Sr39#22 marker were obtained by E.I. Gultyaeva and colleagues (Gultyaeva et al., 2014). Their study noted that, despite the fact that this marker is widely used to identify the *Sr39/Lr35* gene, its diagnostic fragment was amplified in wheat samples with material from *T. timopheevii* and *T. tauschii*: for example, in the variety Pamyati Maistrenko, which was used to obtain lines L995/1 and L997 (Table 1).

The diagnostic fragment of the *Sr32* gene marker csSr32#2 (152 bp) was identified in lines L960, L968 and L995/1 with genetic material from *T. kiharae* and *Th. ponticum* (line L960), *S. cereale* L. (line L968) and *T. timopheevii* (line L995/1) (Supplementary Material 3). Based on all of the above, we did not take into account the results obtained for this marker of the *Sr32* gene, considering them a clear example of a false-positive gene identification.

Another gene from *Ae. speltoides* is *Sr47*, which we identified using three markers: *Xgwm501*, *Xgwm47* and *Xgpw4043*, the results were also ambiguous. The diagnostic fragment (109 bp) of the *Xgwm501* marker was clearly detected in lines L971, L995/1, L997 and L1110 (Supplementary Material 3), which were described above. The Pamyati Maistrenko variety and Dr. Savov's synthetic (GA^{t}D) – *T. timopheevii* × *T. tauschii* are present in the pedigrees of those lines. As can be seen from cytogenetic analysis, the genetic material from *Ae. speltoides* is not present in them (Table 2). As for the *Xgpw4043* marker, the diagnostic fragment of 95 bp was observed in lines L657, L664, L758 and L971, while the second diagnostic fragment of 115 bp was absent. This situation was described in the article (Klindworth et al., 2012), where in some wheat lines with the *Sr47* gene the 115-bp fragment was not amplified or differed in staining intensity and only the 95-bp fragment was amplified and the authors strongly recommended to pick up several markers for *Sr47* gene identification. The diagnostic fragment of the *Xgwm 47* marker was identified only in the L664 line, which has *Ae. speltoides* in its pedigree, and cytogenetic analysis revealed chromosomal substitution 2S(2D). Also, in this line, diagnostic fragment of the Sr39#22 marker (the *Sr39*/*Lr35* gene) was identified. It should be noted that both the *Sr47* and *Sr39* genes are localized on chromosome 2S of *Ae. speltoides*, with *Sr39* on the short arm and *Sr47* on the long arm (Klindworth et al., 2012). Since the L664 line has a 2S(2D) substitution, it is quite possible that it has both the *Sr39* and *Sr47* genes.

The *Sr38* gene comes from the species *Ae. ventricosa* Tausch. and is linked to the genes for resistance to brown (*Lr37*) and yellow (*Yr17*) rust (Bariana, McIntosh, 1993). Despite the absence of this species in the pedigrees of the introgressive lines (Table 1), the diagnostic fragment of the VENTRIUP-LN2 marker to the *Sr38* gene was present in the L971 line. Cytogenetic analysis did not reveal genetic material from *Ae. ventricosa* in this line, therefore we can conclude that in this case there was a false positive identification of the *Sr38* gene.

Thus, attention should be paid to the shown in our study insufficient specificity of the markers for the following genes: *Sr32* – csSr32#2, *Sr38* – VENTRIUP-LN2, *Sr39* – Sr39#22, *Sr47* – *Xgwm501* and *Xgpw4043*. During identification of resistance genes using molecular markers in wheat samples with alien genetic material, this point must be taken into account. It is necessary to once again note the importance of combining different approaches when conducting the analysis of introgressive forms to increase its effectiveness.

Characteristics of new spring bread wheat introgressive lines resistant to stem rust

The Supplementary Material 3 presents the results of assessing the resistance of the analyzed lines to the Volga region populations of the stem rust pathogen at the seedling stage, which we obtained earlier (Baranova et al., 2023a). It was shown that two lines (L657 and L971) were susceptible to the Tatarstan population of the fungus collected from the Nadira variety, while the L971 line was heterogeneous in resistance. Lines L758 and L960 were susceptible to the Saratov population collected from the Voevoda variety. Six lines showed resistance to both populations of the pathogen (L664, L935, L968, L995/1, L997, L1110). Thus, the characterization of introgressive lines will be based on the data from previous works (Baranova et al., 2023a, b) and the results obtained in this study.

All the lines were highly resistant to race Ug99 (TTKSK) with the exception of L995/1 as assessed by KALRO (Kenya) (Table 2). According to FAO data, to date the following genes remain effective versus the Ug99 race: *Sr28*, *Sr29*, *SrTmp* (*T. aestivum* L.), *Sr2*, *Sr13*, *Sr14* (*T. turgidum* L.), *Sr22*, *Sr35* (*T. monococcum* L.), *Sr37* (*T. timopheevii* Zhuk.), *Sr32*, *Sr39*, *Sr47*, (*Ae. speltoides* Tausch.), *Sr33*, *Sr45* (*Ae. tauschii* Coss.), *Sr40* (*T. araraticum* Jakubz.), *Sr25*, *Sr26*, *Sr43* (*Ag. elongatum* Host.), *Sr44* (*Ag. intermedium* Host.), *Sr27* and *Sr1A.1R* (*S. cereale* L.) (http://www.fao.org/agriculture/crops/rust/stem/stem-pathotypetracker/stem-effectivesrgenes/en). The *SrSatu* gene is also effective against the Ug99 race (Olivera et al., 2013).

Based on a previous analysis of the pathogen populations virulence from the Nadira and Voevoda varieties (Baranova et al., 2023b), only the *Sr32* gene is effective against both populations of the fungus among the genes, the presence of which could be assumed in the studied lines. However, this gene was not identified in any of the lines. In addition to it, the *Sr39* gene, identified only in the L664 line, is effective against the Tatarstan population of the pathogen. Consequently, the lines resistance to the pathogen is determined by other unstudied genes or combinations of genes.

In three lines resistant to the Volga region populations of the fungus (L935, L968 and L995/1), molecular genetic analysis failed to identify known resistance genes. Two of them (lines L935 and L968 (Table 1)) carry genetic material from *S. cereale*. Their pedigrees include the triticale variety Satu and they have chromosomal substitution 3R(3D) according to cytogenetic analysis (Table 2, the Figure, *e*). They also turned out to be immune to the Ug99 race. The L968 line is also immune to yellow rust (according to the KALRO assessment), i. e. it has resistance to yellow and stem rust pathogens. The *SrSatu* gene is localized on chromosome 3R of rye and is closely linked to the *LrSatu* gene. In addition, the *Sr27* gene

is localized on chromosome 3R (Singh, McIntosh, 1988). According to McIntosh (1995), the *Sr27* and *SrSatu* genes are allelic to each other and are highly effective against the stem rust pathogen populations. In earlier work, S.J. Singh and R.A. McIntosh, based on genetic analysis of F_2 and F_3 hybrids of stem rust-resistant varieties Satu (*SrSatu*) and Coorong (*Sr27*) with susceptible triticale varieties, showed that the resistance of each variety is determined by one dominant gene and the *SrSatu* and *Sr27* genes are allelic or closely linked (Singh, McIntosh, 1988). This article shows that the Satu variety used in the crosses did not carry the *Sr27* gene. In our studies, the response type was "1" to the population of *P. graminis* f. sp. *tritici* collected from the spring bread wheat variety Voevoda, and "2+" to the population from the variety Nadira for the *Sr27* gene (Baranova et al., 2023b), while lines L935 and L968 showed the infection type of either "0" or "1" (Supplementary Material 3). On the other hand, the *Sr27* gene is effective against the Ug99 race, but the lines containing it are resistant or moderately resistant (R, MR) (Jin et al., 2007); in our study, the lines were immune, the infection type was "0" (Table 2). Thus, taking into account the pedigree, as well as data from cytogenetic and phytopathological analyses, there is reason to believe that the L935 and L968 lines carry the *SrSatu* gene.

Resistance to the Volga region populations of the fungus in line L995/1 (Supplementary Material 3) is most likely determined by unidentified genes on chromosome 2G of *T. timopheevii* (chromosomal substitution 2G(2B), or translocation T2BS.2GL) (Table 2). The Pamyati Maistrenko variety is a donor of alien introgressions and resistance to the stem rust pathogen in the L995/1 line. As is known, this variety inherited chromosomal substitution 2B(2G) from the line of spring bread wheat "Saratovskaya 29 immune L10", which has age-related resistance to the stem rust pathogen (Laikova et al., 2013). The L995/1 line showed an intermediate type of resistance to the Ug99 race (5M).

Resistance genes, the presence of which is confirmed by cytogenetic analysis (Table 2) and pedigree analysis (Table 1), were identified in lines L664, L997 and L1110. Line L664 was resistant to both populations of the fungus collected from the spring bread wheat varieties Voevoda and Nadira and highly resistant (5RMR) to the Ug99 race. The following genes have been identified in this line: *Sr25* (confirmed by the presence of a translocation on 7DL from *Th. ponticum*, T7DS-7DL-7Ae#1L); *Sr39* and *Sr47*, which are located on chromosome 2S of *Ae. speltoides* (McIntosh et al., 2013). Population of *P. graminis* f. sp. *tritici* from the Voevoda variety is virulent to the *Sr39* and *Sr25* genes; however, the L664 line is highly resistant, which may be determined by an additional resistance gene from *Ae. speltoides* – *Sr47*. The type of infection for lines with the *Sr47* gene, infected with the Ug99 race, is "2–" (Klindworth et al., 2012), which correlates with the results of assessing the L664 line for resistance to Ug99 – 5RMR (Table 2). Thus, there is reason to conclude that L664 carries an effective combination of *Sr25* + *Sr39* + *Sr47* genes.

It should be noted that the use of *Sr25* gene sources in breeding for resistance to stem rust is traditional for breeding centers in the Volga region. In 2009, it was reported that a fungal isolate virulent to this gene had been identified in India (Jain et al., 2009). Unfortunately, this previously highly effective gene has been losing efficiency in recent years in the Volga region (Baranova et al., 2021; Baranova et al., 2023b). However, *Sr25* is still effective against the Ug99 race and may be valuable for breeding in combination with other genes such as *Sr31*, *Sr35* and *Sr36*, and in this case with the *Sr39* and *Sr47* genes. Lines with genetic material from *Ae. speltoides* with the *Sr39* + *Sr47* genes are very promising for breeding due to their effectiveness against the Ug99 race (Klindworth et al., 2012). Among the domestic varieties of Russia, the Chelyaba 75 variety, selected by the Chelyabinsk Research Institute of Agriculture with genetic material from *Ae. speltoides* (the *Sr39* gene), stands out and has not only group resistance to leaf and stem rusts and smut (according to the originator), but also resistance to the Ug99 race (Shamanin et al., 2011).

Lines L997 and L1110 are resistant to both Volga populations of *P. graminis* f. sp. *tritici*. Genes *Sr25* (T7DS-7DL-7Ae#1L from *Th. ponticum*) and *Sr36* (2G(2B)), or translocation T2BS.2GL were identified in those lines (Table 2, the Figure, *g–i*). Resistance to the stem rust pathogen in these lines is determined by the *Sr25* + *Sr36* genes combination and, probably, by unidentified gene(s) from *T. timopheevii* on chromosome $2A^t$. This is confirmed by the immunity of these lines to the Ug99 race, as well as the immunity of the L997 line to the yellow rust pathogen.

The resistance of the L960 line to the stem rust population from the Nadira variety and its average resistance to the Ug99 race (5MR) are most likely associated with unidentified genes from *T. kiharae* (Table 2).

Line L971 is moderately resistant to race Ug99 (Table 2). It turned out to be heterogeneous in resistance to the population of *P. graminis* f. sp. *tritici* collected from the spring bread wheat variety Nadira, and resistant to the fungal population from the Voevoda variety (Supplementary Material 3). Since it is known that both populations are virulent to the *Sr25* gene previously identified in this line, it can be assumed that it has other resistance genes, most likely localized on chromosome $2A^t$ of *T. timopheevii* and/or on chromosome 2G (Table 2). Moreover, the resistance gene(s) differ from the genes previously transferred from *T. timopheevii*, *Sr36* (T2B/2G#1) and *Sr40* (T2BL/2G#2S), by greater efficiency, since populations of *P. graminis* f. sp. *tritici* collected from varieties Nadira and Voevoda are virulent to them (Baranova et al., 2023b).

The resistance of the L758 line to the Ug99 race (5R) is determined by the *Sr25* gene – T7DS-7DL-7Ae#1L from *Th. ponticum* (Table 2).

Line L657 is resistant to the stem rust population collected from the Voevoda wheat variety and to the Ug99 race (5RMR), but does not carry any of the tested *Sr* genes. It is possible that unidentified or unknown *Sr* genes that determine the resistance of this line are localized on chromosome 6A of *T. dicoccum*, which replaced its chromosome 6D (Table 2, the Figure, *a*).

Thus, as a result of phytopathological, molecular genetic and cytogenetic analyses, lines that were immune and resistant to the Volga region populations of the fungus, as well as to the Ug99 race (Table 2), with effective combinations of resistance

genes: *Sr25* + *Sr39* + *Sr47* (L664), *Sr25* + *Sr36* (L997 and L1110) and with a gene, previously identified as *SrSatu* (L935 and L968), were isolated. Lines L657, L960, and L971 may be sources of new stem rust resistance genes.

Conclusion

Cytogenetic analysis together with identification using DNA markers of *Sr* genes and phytopathological evaluation of resistance to the Ug99 *P. graminis* f. sp. *tritici* race in introgressive lines of spring bread wheat made it possible to: determine the nature of alien introgressions; establish the degree of resistance to the pathogen; identify effective *Sr* genes. As a result, comprehensive characterization of ten introgressive lines of spring bread wheat resistant to the Ug99 race was obtained, which allows their targeted use in the breeding of spring bread wheat for resistance to the stem rust pathogen.

References

Badaeva E.D., Friebe B., Gill B.S. Genome differentiation in *Aegilops*. 1. Distribution of highly repetitive DNA sequences on chromosome of diploid species. *Genome*. 1996;39(2):293-306. DOI 10.1139/g96-040

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): Application Guide. Springer protocols handbooks. Berlin; Heidelberg: Springer-Verlag, 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)

Baranova O.A., Sibikeev S.N., Konkova E.A. Analysis of resistance to stem rust and identification of *Sr* genes in introgressive lines of spring bread wheat. *Trudy po Prikladnoy Botanike, Genetike i Selektsii = Proceedings on Applied Botany, Genetics and Breeding*. 2023a;184(1):177-186. DOI 10.30901/2227-8834-2023-1-177-186 (in Russian)

Baranova O., Solyanikova V., Kyrova E., Kon'kova E., Gaponov S., Sergeev V., Shevchenko S., Mal'chikov P., Dolzhenko D., Bespalova L., Ablova I., Tarhov A., Vasilova N., Askhadullin Dm., Askhadullin Dn., Sibikeev S. Evaluation of resistance to stem rust and identification of *Sr* genes in Russian spring and winter wheat cultivars in the Volga region. *Agriculture*. 2023b;13(3):635. DOI 10.3390/agriculture13030635

Bariana H.S., McIntosh R.A. Cytogenetic studies in wheat. XV. Location of rust resistance genes in VPM1 and their genetic linkage with other disease resistance genes in chromosome 2A. *Genome*. 1993;36(3):476-482. DOI 10.1139/g93-065

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

Cui Y., Zhang Y., Qi J., Wang H., Wang R.R.C., Bao Y., Li X. Identification of chromosomes in *Thinopyrum intermedium* and wheat *Th. intermedium* amphiploids based on multiplex oligonucleotide probes. *Genome.* 2018;61(7):515-521. DOI 10.1139/gen-2018-0019

Friebe B., Jiang J., Knott D.R., Gill B.S. Compensation indexes of radiation-induced wheat *Agropyron elongatum* translocations conferring resistance to leaf rust and stem rust. *Crop Sci.* 1994;34(2):400-404. DOI 10.2135/cropsci1994.0011183X003400020018x

Friebe B., Jiang J., Raupp W.J., McIntosh R.A., Gill B.S. Characterization of wheat-alien translocations conferring resistance to diseases and pests current status. *Euphytica*. 1996;91:59-87. DOI 10.1007/BF00035277

Gultyaeva E.I., Orina A.S., Gannibal Ph.B., Mitrofanova O.P., Odintsova I.G., Laikova L.I. The effectiveness of molecular markers for the identification of *Lr28*, *Lr35*, and *Lr47* genes in common wheat. *Russ. J. Genet*. 2014;50(2):131-139. DOI 10.1134/S1022795414020069

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* Eriks.) in Lower Volga region. *Selskokhozyaystvennaya Biologiya = Agricultural Biology.* 2020;55(1):27-44. DOI 10.15389/agrobiology.2020.1.27eng

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

Gultyaeva E., Shaydayuk E., Kosman E. Virulence diversity of *Puccinia striiformis* f. sp. *tritici* in common wheat in Russian regions in 2019–2021. *Agriculture*. 2022;12(11):1957. DOI 10.3390/agriculture12111957

Jain S.K., Prashar M., Bhardwaj S.C., Singh S.B., Sharma Y.P. Emergence of virulence to *Sr25* of *Puccinia graminis* f. sp. *tritici* on wheat in India. *Plant Dis*. 2009;93(8):840. DOI 10.1094/PDIS-93-8-0840B

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

Jin Y., Singh R.P., Ward R.W., Wanyera R., Kinyua M., Njau P., Fetch T., Pretorius Z.A., Yahyaoui A. Characterization of seedling infection types and adult plant infection responses of monogenic *Sr* gene lines to race TTKS of *Puccinia graminis* f. sp. *tritici*. *Plant Dis.* 2007;91(9):1096-1099. DOI 10.1094/PDIS-91-9-1096

Kelbin V.N., Skolotneva E.S., Salina E.A. Challenges and prospects for developing genetic resistance in common wheat against stem rust in Western Siberia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2020;24(8):821-828. DOI 10.18699/VJ20.679 (in Russian)

Klindworth D.L., Niu Z., Chao S., Friesen T.L., Faris J.D., Cai X., Xu S.S. Introgression and characterization of a goatgrass gene for a high level of resistance to Ug99 stem rust in tetraploid wheat. *G3* (*Bethesda*). 2012;2(6):665-673. DOI 10.1534/g3.112.002386

Knott D.R. The Wheat Rust – Breeding for Resistance. (Ser. Monographs on Theoretical and Applied Genetics). Berlin: Springer-Verlag, 1989. DOI 10.1007/978-3-642-83641-1

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 from synthetic hexaploid *Triticum timopheevii* Zhuk. × *Aegilops tauschii* Coss. *Russ. J. Genet.* 2013;49(1):89-97. DOI 10.1134/S1022795413010067

Lewis C.M., Persoons A., Bebber D.P., Kigathi R.N., Maintz J., Findlay K., Bueno-Sancho V., Corredor-Moreno P., Harrington S.A., Kangara N., Berlin A., García R., German S.E., Hanzalová A., Hodson D.P., Hovmøller M.S., Huerta-Espino J., Imtiaz M., Mirza J.I., Justesen A.F., Niks R.E., Omarani A., Patpour M., Pretorius Z.A., Roohparvar R., Sela H., Singh R.P., Steffenson B., Visser B., Fenwick P.M., Thomas J., Wulff B.B.H., Saunders D.G.O. Potential for re-emergence of wheat stem rust in the United Kingdom. *Commun. Biol*. 2018;1:13. DOI 10.1038/s42003-018-0013-y

McIntosh R.A., Wellings C.R., Park R.F. Wheat Rusts: An atlas of resistant genes. CSIRO Publication, Collingwood, 1995. DOI 10.1007/978-94-011-0083-0

McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers W.J., Morris C., Appels R., Xia X.C. Catalogue of Gene Symbols for Wheat. 12th International Wheat Genetics Symposium. 8–13 September 2013. Yokohama, Japan, 2013. https://shigen.nig.ac.jp/wheat/komugi/genes/macgene/2013/GeneCatalogueIntroduction.pdf

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;68:68-81

Murray M.G., Thompson W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res*. 1980;8(19):4321-4325. DOI 10.1093/nar/8.19.4321

Olivera P.D., Pretorius Z.A., Badebo A., Jin Y. Identification of resistance to races of *Puccinia graminis* f. sp. *tritici* with broad virulence in triticale (×*Triticosecale*). *Plant Dis*. 2013;97:479-484. DOI 10.1094/PDIS-05-12-0459-RE

Patpour M., Hovmøller M.S., Rodriguez-Algaba J., Randazzo B., Villegas D., Shamanin V.P., Berlin A., Flath K., Czembor P., Hanzalova A., Sliková S., Skolotneva E.S., Jin Y., Szabo L., Meyer K.J.G., Valade R., Thach T., Hansen J.G., Justesen A.F. Wheat stem rust back in Europe: diversity, prevalence and impact on host resistance. *Front. Plant Sci.* 2022;13:882440. DOI 10.3389/fpls.2022.882440

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

Prins R., Groenewald J.Z., Marais G.F., Snape J.W., Koebner R.M.D. AFLP and STS tagging of *Lr19*, a gene conferring resistance to leaf rust in wheat. *Theor. Appl. Genet.* 2001;103(4):618-624. DOI 10.1007/PL00002918

Rayburn A.L., Gill B.S. Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. *Plant Mol. Biol. Rep*. 1986;4(2):104-109. DOI 10.1007/BF02732107

Ruban A.S., Badaeva E.D. Evolution of the S-genomes in *Triticum-Aegilops* alliance: evidences from chromosome analysis. *Front. Plant Sci*. 2018;9:1756. DOI 10.3389/fpls.2018.01756

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

Salina E.A., Adonina I.G., Vatolina T.Y., Kurata N. Comparative analysis of the composition and organization of two subtelomeric repeat families in *Aegilops speltoides* Tausch and related species. *Genetica*. 2004;122(3):227-237. DOI 10.1007/s10709-004-5602-7

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

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

Schubert I., Shi F., Fuchs J., Endo T.R. An efficient screening for terminal deletions and translocations of barley chromosomes added to common wheat. *Plant J*. 1998;14(4):489-495. DOI 10.1046/j.1365-313X.1998.00125.x

Shamanin V.P., Morgunov A.I., Chursin A.S., Merkeshina N.N., Shtubey T.Yu., Levshunov M.A., Karakoz I.I. Does stem rust pose a threat to the wheat harvest in Western Siberia. *Uspekhi Sovremennogo Estestvoznaniya = Advances in Current Natural Sciences*. 2011;2:56-60 (in Russian)

Singh S.J., Mclntosh R.A. Allelism of two genes for stem rust resistance in triticale. *Euphytica*. 1988;38:185-189. DOI 10.1007/BF00040190

Skolotneva E.S., Kelbin V.N., Shamanin V.P., Boyko N.I., Aparina V.A., Salina E.A. The gene *Sr38* for bread wheat breeding in Western Siberia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2021;25(7):740-745. DOI 10.18699/VJ21.084

Vasilova N.Z., Askhadullin Dam.F., Askhadullin Dan.F. Stem rust epiphytotic on soft spring wheat in Tatarstan. *Zashchita i Karantin Rasteniy = Plant Protection and Quarantine*. 2017;2:27-28 (in Russian)

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

Received November 7, 2023. Revised March 5, 2024. Accepted March 5, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):387-397

DOI 10.18699/vjgb-24-44

Transgenerational effect of prenatal stress on behavior and lipid peroxidation in brain structures of female rats during the estral cycle

A.V. Vyushina, A.V. Pritvorova, S.G. Pivina, N.E. Ordyan

Pavlov Institute of Physiology, Russian Academy of Sciences, St. Petersburg, Russia
PritvorovaAV@infran.ru

Abstract. The effect of stress in pregnant female Wistar rats on the behavior and lipid peroxidation (LP) in the neocortex, hippocampus and hypothalamus in the female F2 generation during the ovarian cycle was investigated. We subjected pregnant females to daily 1-hour immobilization stress from the 15th to the 19th days of pregnancy. Further, family groups were formed from prenatally stressed and control male and female rats of the F1 generation: group 1, the control female and male; group 2, the control female and the prenatally stressed male; group 3, the prenatally stressed female and the control male; group 4, the prenatally stressed female and male. The females of the F2 generation born from these couples were selected into four experimental groups in accordance with the family group. At the age of 3 months, behavior of rats was studied in the "open field" test in two stages of the ovarian cycle – estrus and diestrus. After 7–10 days, the rats were decapitated and the neocortex, hypothalamus and hippocampus were selected to determine the level of diene and triene conjugates, Schiff bases and the degree of lipid oxidation (Klein index). In F2 females with one prenatally stressed parent, there was no interstage difference in locomotor-exploratory activity and anxiety. If both F1 parents were prenatally stressed, female F2 rats retained interstage differences similar to the control pattern, while their locomotor-exploratory activity and time spent in the center of an "open field" decreased in absolute values. In the neocortex of F2 females in groups with prenatally stressed mothers, the level of primary LP products decreased and the level of Schiff bases increased in the estrus stage. In the hippocampus of F2 females in the groups with prenatally stressed fathers, the level of Schiff bases decreased in the estrus stage, and the level of primary LP products increased in group 2 and decreased in group 4. In the hypothalamus of F2 females in the groups with prenatally stressed mothers, the level of Schiff bases increased in the estrus stage and decreased in the diestrus; in addition, in group 3, the level of primary LP products in the estrus stage increased. Thus, we demonstrated the influence of prenatal stress of both F1 mother and F1 father on the behavior and the level of LP in the neocortex, hippocampus and hypothalamus in female rats of the F2 generation in estrus and diestrus.
Key words: prenatal stress; F2 generation; behavior; lipid peroxidation; estrus; diestrus.

For citation: Vyushina A.V., Pritvorova A.V., Pivina S.G., Ordyan N.E. Transgenerational effect of prenatal stress on behavior and lipid peroxidation in brain structures of female rats during the estral cycle. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):387-397. DOI 10.18699/vjgb-24-44

Funding. The study was supported by the state funding allocated to the Pavlov Institute of Physiology of the Russian Academy of Sciences (2024–2028) No. 1023032400236-8-3.1.4.

Трансгенерационное влияние пренатального стресса на поведение и перекисное окисление липидов в структурах мозга у самок крыс в течение эстрального цикла

А.В. Вьюшина, А.В. Притворова, С.Г. Пивина, Н.Э. Ордян

Институт физиологии им. И.П. Павлова Российской академии наук, Санкт-Петербург, Россия
PritvorovaAV@infran.ru

Аннотация. Исследовано влияние стресса у беременных самок крыс Вистар на поведение и показатели перекисного окисления липидов (ПОЛ) в неокортексе, гиппокампе и гипоталамусе у поколения самок F2 в течение эстрального цикла. Беременных самок подвергали ежедневному 1-часовому иммобилизационному стрессу с 15-го по 19-й день беременности. Далее из рожденных пренатально стрессированных и контрольных самцов и самок крыс поколения F1 формировали семейные группы: группа 1 – контрольные самка и самец, группа 2 – контрольная самка и пренатально стрессированный самец, группа 3 – пренатально стрессированная самка и контрольный самец, группа 4 – пренатально стрессированные самка и самец. Рожденных от этих семейных

пар самок поколения F2 отбирали в четыре экспериментальные группы в соответствии с семейной группой. В возрасте трех месяцев у крыс исследовали показатели поведения в тесте «открытое поле» в двух стадиях полового цикла – эструсе и диэструсе. Через 7–10 дней крыс декапитировали и производили отбор неокортекса, гипоталамуса и гиппокампа для определения уровня диеновых и триеновых конъюгатов, оснований Шиффа и степени окисленности липидов (индекса Клейна). У самок F2 с одним пренатально стрессированным родителем отсутствует межстадиальная разница в локомоторно-исследовательской активности и тревожности. Если оба родителя F1 являются пренатально стрессированными, самки крыс F2 сохраняют межстадиальные различия, схожие с контрольным паттерном, при этом по абсолютным значениям у них снижаются локомоторно-исследовательская активность и время нахождения в центре открытого поля. В неокортексе у самок F2 в группах с пренатально стрессированными матерями снижается уровень первичных продуктов ПОЛ и повышается уровень оснований Шиффа в стадии эструса. В гиппокампе у самок F2 в группах с пренатально стрессированными отцами снижается уровень оснований Шиффа в стадии эструса, а уровень первичных продуктов ПОЛ повышается в группе 2 и снижается в группе 4. В гипоталамусе у самок F2 в группах с пренатально стрессированными матерями уровень оснований Шиффа повышается в стадии эструса и снижается в диэструсе, кроме того, в группе 3 повышается уровень первичных продуктов ПОЛ в стадии эструса. Таким образом, выявлено влияние пренатального стресса как матери F1, так и отца F1 на показатели поведения и уровень ПОЛ в неокортексе, гиппокампе и гипоталамусе у самок крыс поколения F2 в эструсе и диэструсе.
Ключевые слова: пренатальный стресс; поколение F2; поведение; перекисное окисление липидов; эструс; диэструс.

Introduction

It has now been established that an increase in the level of maternal glucocorticoids during pregnancy causes changes in the neuroendocrine and immune systems of the offspring. Elevated level of maternal glucocorticoids promotes excess production of reactive oxygen species (ROS), with the organ-specific stress response depending on the relative balance between ROS generation and the antioxidant capacity of the cell (Dennery, 2010; Thompson, Al-Hasan, 2012). Disruption of this balance leads to oxidative stress and contributes to epigenetic changes in prenatally stressed offspring. Epigenetic changes are maintained in a number of mitotic divisions of somatic cells, and can also be transmitted to the next generations if these changes occurred in germ cells (Dyban, 1988; Rodgers, Bale, 2015; Yao et al., 2021). Thus, numerous negative effects of maternal stress identified in the first generation may be sustained in subsequent generations (Essex et al., 2013; Provençal, Binder, 2015).

In females, both epigenetic changes and maternal behavior influence offspring, so studying the transgenerational effects of stress in male rodents has advantages compared to females (Brunton, 2013; Bale, 2014). In this regard, transgenerational changes caused by stress in females remain insufficiently studied. Currently, researchers are paying special attention to alteration of fertility in F2 and subsequent generations. Thus, a number of authors (Zhang et al., 2020; Piquer et al., 2022) found experimentally that prenatal nonphysiological influences of varying genesis affect fertility in both the second and third generations. This is expressed as a change in a number of morphometric parameters of the ovaries and uterus, biochemical parameters such as the level of corticosterone, luteinizing hormone, follicle-stimulating hormone, insulin and other metabolic parameters in the blood serum, as well as in disruption of the estrous cycle. Other studies have found that prenatal stress has a transgenerational effect on the processes of free radical oxidation of biomolecules in various tissues (Aiken et al., 2019).

Epigenetic changes have the potential to influence endocrine programming and brain development in the fetus over multiple generations. The authors of the review (Babenko et al., 2015) emphasize the complex relationship between the effects of prenatal stress on changes in microRNA expression, DNA methylation in the placenta and brain and an increased risk of developing mental illness.

It can be assumed that prenatal stress, which causes epigenetic changes, becomes one of the most potent factors affecting mental health. Moreover, such changes affect, among other things, various structures of the brain associated with the neuroendocrine system and cognitive abilities. In the research (Huerta-Cervantes et al., 2021), it is noted that cognitive impairment in female rats at different stages of the ovarian cycle may be associated with disorders in the processes of lipid peroxidation in the hippocampus and neocortex.

Lipid peroxidation is not only a universal modifier of the properties of biological membranes, but also an important physiological regulator of their structure, a factor that establishes and supports the activity of enzymes, channel-former molecules, and receptors. The intensity of free radical processes of lipid peroxidation, which are under the control of endogenous oxidants, is associated with the composition and physical state of phospholipids of biological membranes (their fluidity), with their sensitivity to ligand signals and extreme influences. It is also extremely important for the regulatory and informational properties of membranes in normal cellular metabolism. Oxidative processes that affect the composition and viscosity of the lipid layer of membranes can regulate cellular metabolism (Baraboy et al., 1992).

Behavioral changes at different stages of the ovarian cycle in rats are associated with preparation for successful reproductive function. The changes in fertility may influence the behavioral features of different stages of the ovarian cycle. Thus, it is of interest to study the transgenerational effect of prenatal stress on behavior and lipid peroxidation in the brain structures of female rats during the ovarian cycle.

Matherials and methods

For this study, the Wistar rats from the I.P. Pavlov Institute of Physiology of the Russian Academy of Sciences (St. Pe-

tersburg) were used. The recommendations on the ethics of working with animals proposed by Directive 2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific purposes were respected. The breeding protocol is presented in Figure 1.

Breeding of experimental groups of F2 offspring. *Stage 1*. The Wistar rats of the F0 generation weighing 280–300 g and three months old were used at the beginning of this stage. Rats were housed in standard cages for laboratory mice and rats M-6 ("Proflab", Russia) and received rat water and chow for laboratory animals LBK-120 (Tosno, Russia) *ad libitum*. The animals were kept at a controlled temperature (22–24 °C). A 12:12 h light-dark cycle was maintained. Female rats were coupled with males; fertilization was confirmed by the detection of spermatozoa in a vaginal swab and indicated as day zero of pregnancy. Pregnant females were randomly divided into two groups: control pregnant rats to breed control F1 offspring (n = 12) and rats that were further stressed to breed prenatally stressed F1 offspring (n = 12).

Stage 2. To breed prenatally stressed F1 offspring, pregnant F0 females were exposed to one-hour immobilization stress under high-light conditions from the 15th to the 19th day of gestation (Ordyan, Pivina, 2003). The stress was performed at the same time of day from 14.00 to 15.00 h. The days for stressing were chosen due to the fact that it is during this period that the integration of all parts of neuroendocrine regulation takes place and the formation of the hypothalamic-pituitary-adrenocortical system is completed (Rice, Barone, 2000).

Control and stressed pregnant females were housed 4–5 individuals in a cage. At the 20th day of pregnancy, the dams were housed in individual cages. The resulting offspring were counted on the 2nd day of life, taking into account the number of males and females in litters, and the litters themselves were aligned to 8–10 pups with an equal sex ratio. The pups were housed with their mother for 30 days and received water and chow *ad libitum*. Further, prenatally stressed F1 offspring were placed in cages, separating males and females. The control F1 offspring, born to intact F0 dames, were also aligned on the 2nd day of life and separated from the mother at the age of 30 days – males and females separately.

Stage 3. For breeding F2 offspring, animals were randomly selected by an independent person. The family groups of F1 offspring were formed from one male and three females, so that the animals were not siblings. These family groups were formed in such a way that the F2 offspring resulted in four experimental groups: group 1 (k+k) – offspring obtained from F1 females and F1 males from the control group that were not exposed to any influences, group 2 (k+2) – offspring obtained from F1 females of the control group and F1 males of the prenatally stressed group, group 3 (2+k) – offspring obtained from F1 females of the prenatally stressed group and F1 males from the control group, group 4 (2+2) – offspring obtained from F1 females and F1 males from the prenatally stressed group.

The resulting F2 offspring were counted on the 2nd day of life, taking into account the number of males and females in litters, and the litters themselves, as in the case of F1 offspring, were aligned to 8–10 pups with an equal sex ratio. The pups were housed with their mother for 30 days and received water and chow *ad libitum*. Next, the F2 offspring were placed in cages of 5–7 individuals, separating males and females. Females aged 3 months were used for further studies.

Previously, the rats were subjected to handling and trained: vaginal swabs were taken from them for 3 weeks. Next, the

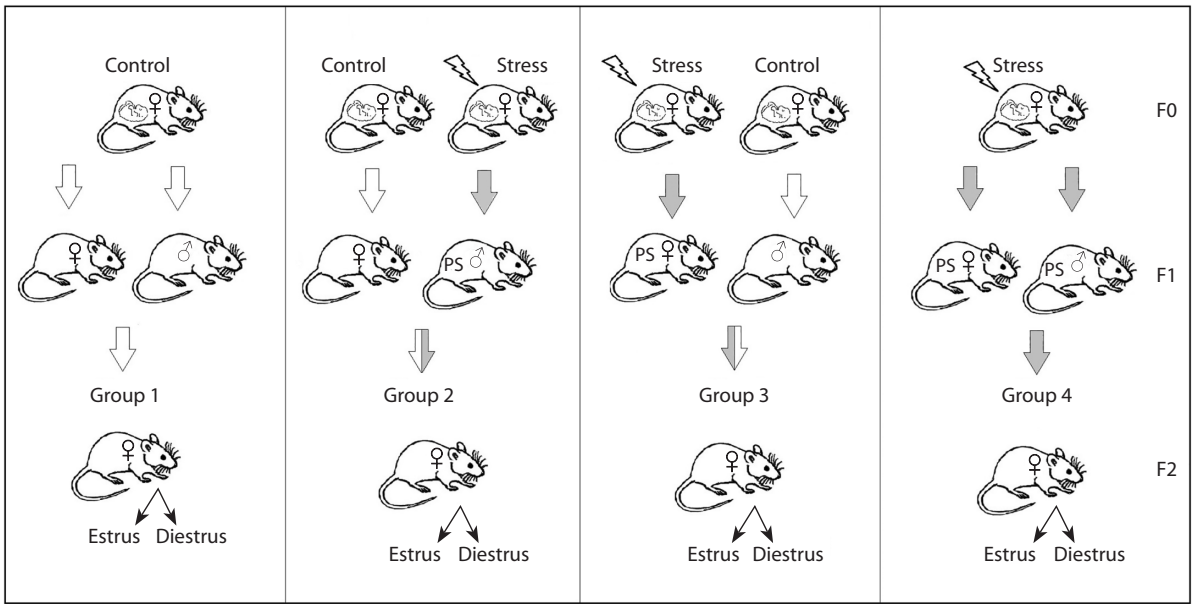


Fig. 1. Experimental design.

Group 1 – F2 offspring born from unexposed parents; group 2 – F2 offspring born from a mother not exposed to any influences and a prenatally stressed (PS) father; group 3 – F2 offspring born from a prenatally stressed mother and a father not exposed to any influences; group 4 – F2 offspring born from prenatally stressed parents.

animals' behavior was tested; immediately after testing, vaginal swabs were taken from rats and the stage of the ovarian cycle was determined. Two weeks after behavior testing, the rats were decapitated. Immediately after decapitation, control vaginal swabs were taken again.

Behavior testing. Open field. The "open field" test (OF) was a rectangular Plexiglas box (90 × 90 × 50 cm), the floor of which was divided into squares (15 × 15 cm). The illumination of the box was 300 lx. Testing occurred for 5 min from 10.00 to 13.00 h. The rat was placed in the center of the box and the total time in the center, the number of crossed squares (horizontal motor activity or locomotor activity), the number of vertical positions (vertical motor activity or research activity), the time of the grooming reaction and the time of immobility (freezing) were recorded. Indicators of horizontal and vertical motor activity indicate locomotor research activity. The total time in the center, the time of immobility and the reaction time of grooming indicate the degree of anxiety in rats.

Determination of lipid peroxidation products. Rats were decapitated and the neocortex, hippocampus and hypothalamus were isolated on the ice. Next, lipids were extracted from the samples using the Folch method.

To determine the level of conjugated diene (CD) and triene (CT), the Klein index the dry lipid extract was dissolved in a methanol : heptane (2:1) mixture and the optical density was measured – the CD level at 230 nm and the CT level at 274 nm. The content of conjugated dienes and trienes was expressed in units of optical density per 1 mg of phospholipids (Arutyunyan et al., 2000). The fluorescent intensity of Schiff bases was determined by the fluorimetric method at a maximum excitation of 365 nm and a maximum emission of 425 nm (Bidlack, Tappel, 1973), expressed in relative units of fluorescence per 1 mg of phospholipids.

The amount of phospholipids was estimated by the content of nonorganic phosphorus by the Bartlett method. The method is based on the reaction of nonorganic phosphate with ammonium molybdate, resulting in phosphoric-molybdenum acid, which is then reduced by eikonogen to form colored molybdenum oxides, the optical density of which is measured at 830 nm.

To determine the degree of lipid oxidation, the Klein index was calculated as follows: the optical density of lipid extracts was determined at 215 nm and the ratio of optical densities at 233 and 215 nm was calculated.

To determine the level of conjugated diene and triene, the phospholipids and the Klein index, a BioTek PowerWave HT spectrophotometer (USA) was used. The determination of the Schiff bases level was carried out using a Hitachi MPF-4 spectrofluorimeter (Japan).

All reagents used in biochemical analyses were purchased from "Vecton" (Russia), with the exception of eikonogen (Merck, Germany).

Statistical analysis. For statistical analysis, STATISTICA 8.0 software package (StatSoft Inc.) was used. The normality of the distribution of values in the samples was determined using the Shapiro–Wilk test. Normally distributed data were analyzed with a parametric Student's *t*-test and non-normally distributed data were analyzed with a non-parametric Mann–Whitney U-test. Data that are presented as Mean ± SEM are indexes of behavior, and those presented as medians (Me) and interquartile range (IQR) between the values of 25 and 75 percentiles are biochemical indexes. To identify differences between ovarian cycle stages in each of the four studied groups, comparisons of each indicator in the estrus and diestrus stages were performed. And also the values of each index in groups 2, 3 and 4 were compared with group 1 both in estrus and in diestrus. The differences were considered statistically significant at $p < 0.05$.

Results

Behavior

Differences in behavioral indicators between ovarian cycle stages are observed in the group of control animals (Fig. 2).

The indicators of horizontal motor activity, vertical motor activity and time of presence in the center of OF in females in estrus are higher than in females in diestrus, while the time of immobility and grooming time are lower in females in estrus. It can be concluded that in the control group, females in estrus have increased motor and research activity and decreased anxiety (according to the indicators in the OF test), which, apparently, represents an evolutionarily appropriate strategy related to sexual behavior.

In the k+2 group, horizontal and vertical motor activities, as well as time of presence in the center in female F2 rats, do not differ in estrus and diestrus. The indicators "freezing" and "grooming reaction time" demonstrate an interstadial difference. F2 females in diestrus become more active and less anxious compared to the control. In F2 females, research activity decreases in estrus, the freezing time increases, the time present in the center increases; that is, the females become less anxious. Based on the changes described above, we assume that F2 females, during the most favorable period for mating, become less mobile compared to the control and, accordingly, the likelihood of meeting a partner decreases. At the same time, females in diestrus, according to the studied behavior indicators, approach females in the estrus stage in the control group.

In group 2+k, the indicators of horizontal and vertical motor activity and time in the center in female F2 rats do not differ in estrus and diestrus. The indicators "freezing" and "grooming reaction time" demonstrate the difference between estrus and diestrus, which is inverted with respect to the control. F2 females in diestrus are less active, spend less time on grooming and prefer to be in the center of the OF, i. e. they have reduced anxiety compared to the control. Research activity decreases and the time of immobility increases in F2 females in estrus compared to the control; in addition, females of this group are less mobile in estrus than in diestrus. Thus, in this group, too, females in estrus have a distortion of the behavior associated with finding a partner.

In the 2+2 group, female F2 rats have shown differences between the stages of the ovarian cycle in all indicators except horizontal motor activity. F2 females in diestrus are less mobile and more anxious compared to the controls. In F2 females in estrus, there is also a decrease in locomotor and research activity and an increase in anxiety compared to the control. Nevertheless, the interstadial ratio in females of this

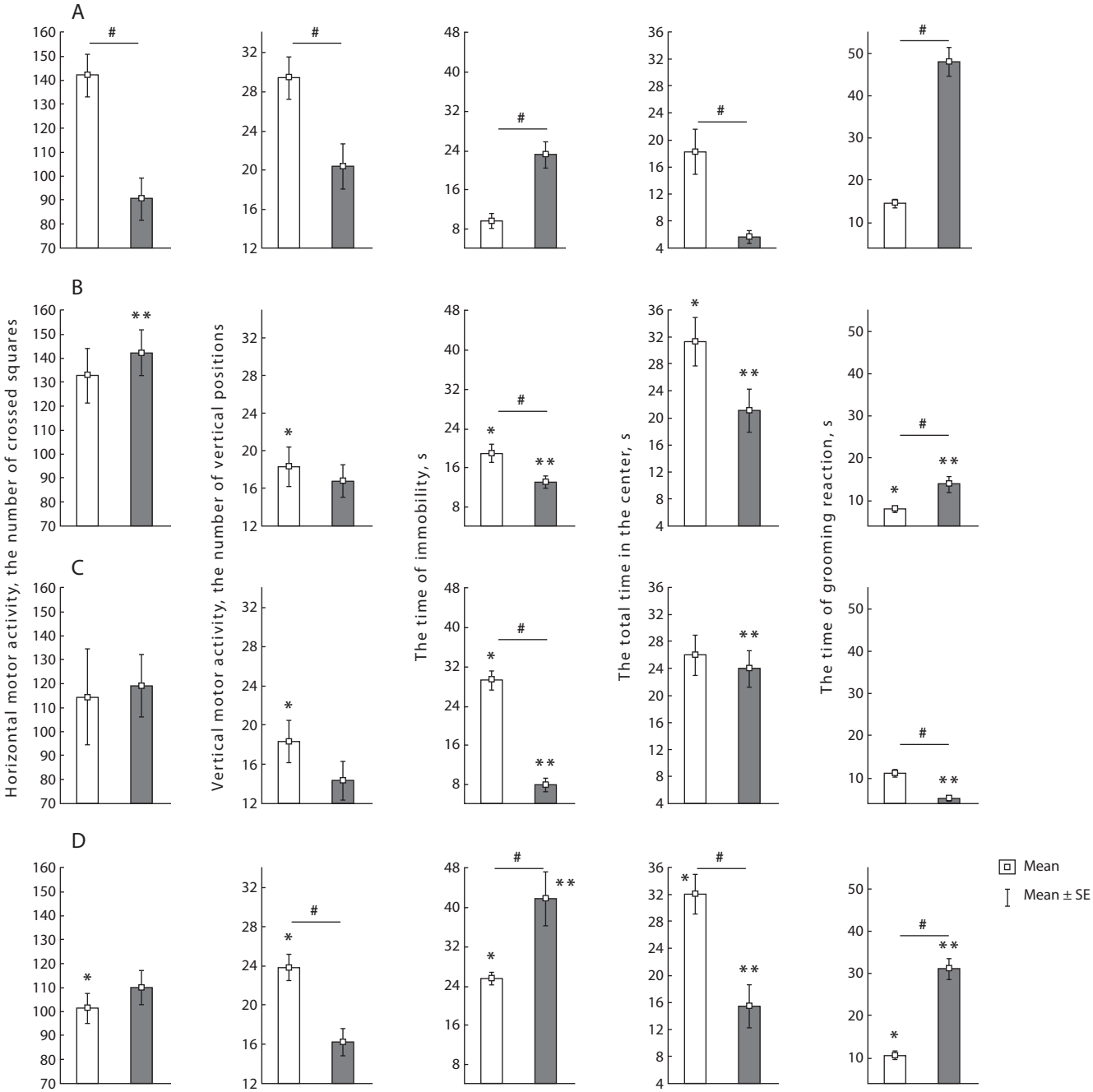


Fig. 2. Rats behavior in the "open field" test.

Light bars are rats in estrus, dark bars are rats in diestrus. Panel A – group 1 (control rats); panel B – group 2 (control mother and PS father); panel C – group 3 (PS mother and control father); panel D – (PS parents). # – statistically significant differences between rats in diestrus and estrus, $p < 0.05$; * – statistically significant differences from the control in animals in estrus, $p < 0.05$; ** – statistically significant differences from the control in animals in diestrus, $p < 0.05$.

group remains similar to the interstadial ratio in the control group; however, according to the absolute values of behavior indicators, females of the 2+2 group differ from the control group by an increase in anxiety.

Neocortex

In the control group, there are differences between the stages of the cycle only in Schiff bases: in the estrus stage, the level of the end products of lipid peroxidation reactions is 2 times lower than in the diestrus stage (Fig. 3).

In the k+2 group, the Schiff bases level in diestrus is two times smaller than in the control group; however, the indicators of other studied products of lipid peroxidation do not differ from the level of the control group. There are no interstadial differences.

In the 2+k group, in F2 females, relative to the control indicators, the level of CD and CT – the initial products of lipid peroxidation – is lower, and the level of Schiff bases (the final products) is higher. It should be noted that the CD indicators and the Klein index show an interstadial difference.

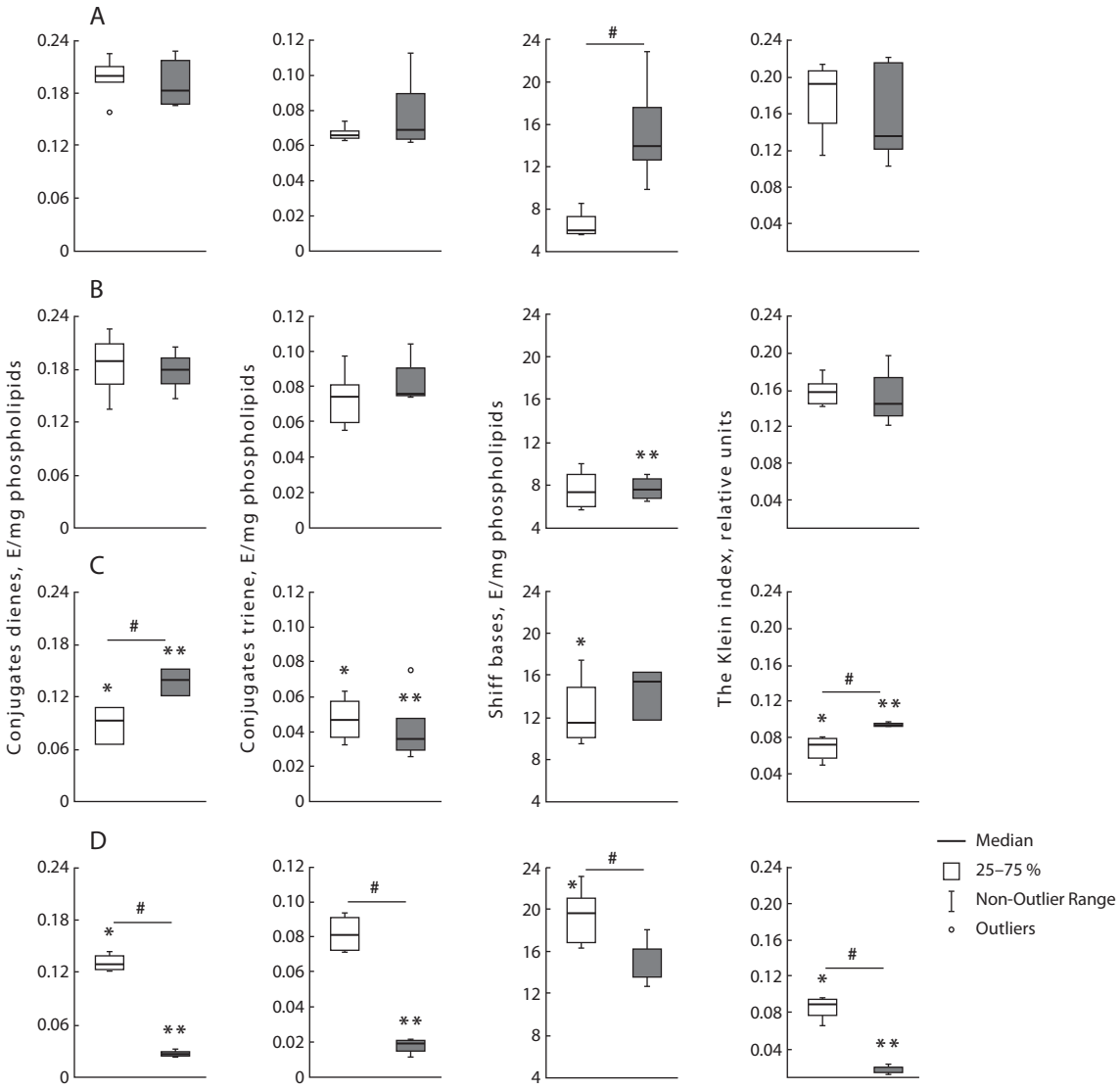


Fig. 3. CD, CT, Schiff bases levels and Klein index in the neocortex of female rats in the experimental groups.

Light bars are rats in estrus, dark bars are rats in diestrus. Panel A – group 1 (control rats); panel B – group 2 (control mother and PS father); panel C – group 3 (PS mother and control father); panel D – (PS parents). # – statistically significant differences between rats in diestrus and estrus, $p < 0.05$; * – statistically significant differences from the control in animals in estrus, $p < 0.05$; ** – statistically significant differences from the control in animals in diestrus, $p < 0.05$.

The levels of CD, CT and the Klein index in the 2+2 group are lower relative to the corresponding control indicators in both stages of the cycle, with the exception of CT in estrus. The Schiff bases level in estrus is 3 times higher than in the control group. It can be concluded that in the 2+2 group, the level of lipid peroxidation indicators is lower compared to the control, especially in the diestrus stage, excluding the Schiff bases. It should be noted that all the studied lipid peroxidation indicators of this group demonstrate a difference between estrus and diestrus.

Thus, the k+k and k+2 groups are similar in their profile of levels of CD, CT and the Klein index in the neocortex, whereas groups 2+2 and 2+k differ by a decrease in these indicators of lipid peroxidation. At the same time, the Schiff bases indicators of these groups in the estrus stage are significantly higher than the control indicators, whereas in diestrus, only the k+2 group is characterized by an acute decrease in this indicator.

Hippocampus

In the control group, there are differences between the stages of the cycle only in the Schiff bases level: in the estrus stage, the level of the end products of lipid peroxidation is two times higher than in the diestrus stage (Fig. 4).

In the k+2 group, the CD level in estrus and diestrus is higher than in the control group, the CT level in diestrus is higher than in the control group. The Schiff bases level in estrus is two times lower than in the control group. The values of the Klein index in diestrus are higher than in the control. It can be concluded that the level of indicators of the initial lipid peroxidation products is higher compared to the control group, especially in diestrus, but this group is characterized by the absence of differences between estrus and diestrus.

In the 2+k group, the Schiff bases level in the diestrus stage is lower relative to the control group. In diestrus, there are no differences between CD, CT and the Klein index relative

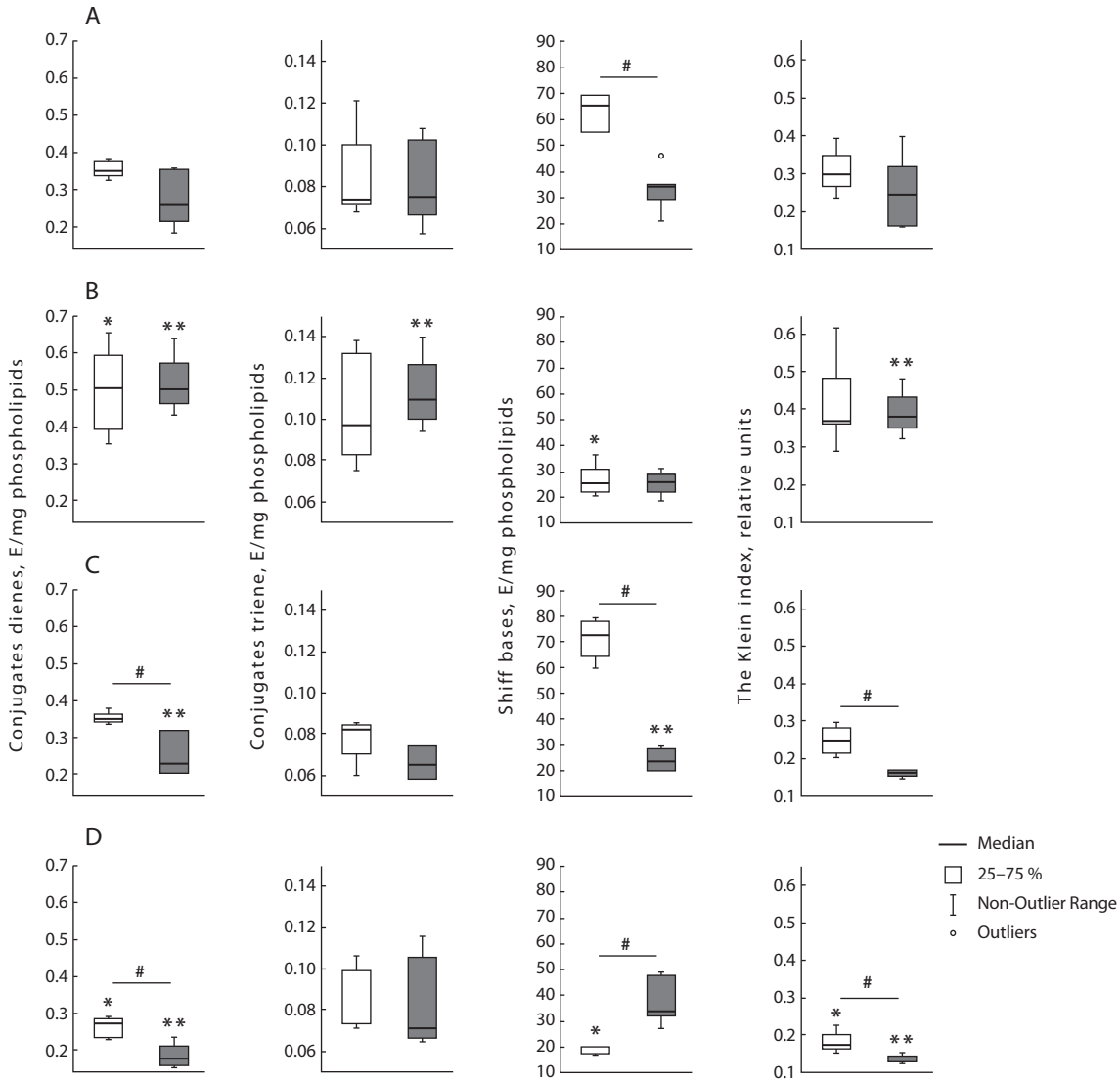


Fig. 4. CD, CT, Schiff bases levels and Klein index in the hippocampus of female rats in the experimental groups.

Light bars are rats in the estrus stage, dark bars are rats in the diestrus stage. Panel A – group 1 (control rats); panel B – group 2 (control mother and PS father); panel C – group 3 (PS mother and control father); panel D – (PS parents). # – statistically significant differences between rats in diestrus and estrus, $p < 0.05$; * – statistically significant differences from the control in animals in estrus, $p < 0.05$; ** – statistically significant differences from the control in animals in diestrus, $p < 0.05$.

to the control indicators. At the same time, there is a difference between estrus and diestrus for Schiff bases, CD, and the Klein index.

The CD levels in the 2+2 group in both stages of the ovarian cycle are lower relative to the values in the control group. The CT level is characterized by the absence of both differences between the ovarian cycle stages and differences from the control group. The Schiff bases level in estrus is 3 times lower compared to the control group, and there is no difference in diestrus compared to the control. The Klein index in estrus and diestrus is lower relative to the respective control indicators. In group 2+2, the level of lipid peroxidation indicators is lower compared to the control and there are differences between estrus and diestrus in the levels of both initial and final lipid peroxidation products. It is noteworthy that the 2+k and 2+2 groups demonstrate a difference between estrus and diestrus in lipid peroxidation indicators.

Hypothalamus

In the control group, the levels of CT and Schiff bases differ between estrus and diestrus. The level of these lipid peroxidation products is higher in diestrus compared to estrus (Fig. 5).

In the k+2 group, the Schiff bases level in diestrus is lower than in the control group. The CD level and the Klein index in diestrus are 3 times higher than the control level, and the CT level does not differ. Interstadial differences are observed for all the studied lipid peroxidation indicators in k+2 group.

In the 2+k group, the Schiff bases level in estrus is 10 times higher relative to the control level, but in diestrus it is lower. The levels of CD and CT in estrus are higher compared to the control group. The Schiff bases level has a difference between estrus and diestrus inverted with respect to the control.

The Schiff bases level in the 2+2 group in estrus is 3 times higher than in the control group, whereas in diestrus, the Schiff bases level is 2 times lower than in the control. There are no

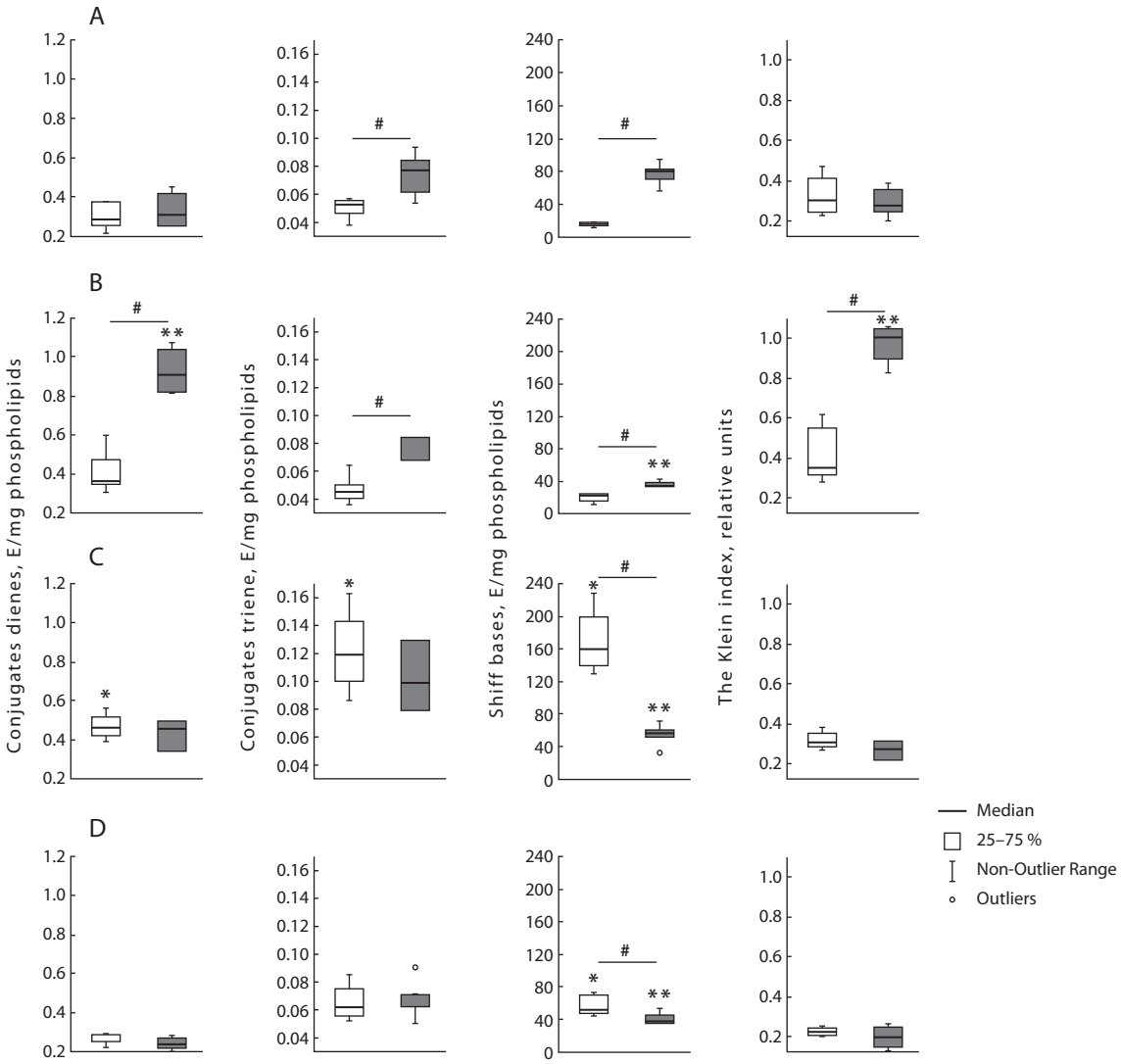


Fig. 5. CD, CT, Schiff bases levels and Klein index in the hypothalamus in female rats in the experimental groups.

Light bars are rats in estrus, dark bars are rats in diestrus. Panel A – group 1 (control rats), panel B – group 2 (control mother and PS father), panel C – group 3 (PS mother and control father), panel D – (PS parents). # – statistically significant differences between rats in diestrus and rats in estrus, $p < 0.05$; * – statistically significant differences from the control in animals in estrus, $p < 0.05$; ** – statistically significant differences from the control in animals in diestrus, $p < 0.05$.

differences in other lipid peroxidation indicators between estrus and diestrus. Also there are no differences in these indicators of the control group.

The similarity of the k+k and 2+2 groups in the level of initial and intermediate lipid peroxidation indicators is noteworthy. At the same time, the Schiff bases levels of the 2+2 group in estrus and diestrus are inverted relative to the control. Also noteworthy is the multiple increase in comparison with the control of the initial lipid peroxidation products and the oxidation index (Klein index) in diestrus in the k+2 group, and the Schiff bases level in estrus in the 2+k group.

Discussion

An increase in maternal glucocorticoid levels during pregnancy can lead to sustainable epigenetic changes. In the research (Gilbert et al., 2012; Matthews, Phillips, 2012), it is noted that epigenetic changes may be sustained over subsequent generations.

Due to the changed social status of women in recent decades, there has been a shift in the reproductive age to later age cohorts. The presence of a significant number of assisted reproductive technologies (ART) allows women to procreate even in the case of significant pathologies. However, the negative consequences of such pathological pregnancies for future generations are currently only beginning to be understood (Aiken et al, 2015; Sanches-Garrido et al., 2022).

In the review by A.L. Levinson et al. (2022), the authors note the multifactorial nature of hormonal effects in the reproductive dysfunction problem. When analyzing the failures of the use of ART, the importance of not only hypothalamic disorders, but also the influence of paracrine factors of the ovary has been revealed. At the same time, in our laboratory's study, it was found that the negative effect of prenatal stress on the morphometric parameters of the uterus associated with a disturbance of the cycle of sex hormones is noted only in young female rats aged 3 months, while in older animals such

disruptions are leveled (Pivina et al., 2010). However, in a study by B. Piquer et al. (2022), a decrease in the fertility of female rats has been noted, expressed in impaired fertilization and the number of pups born after prenatal stress up to the F4 generation. The same authors have shown a disturbance of the morphometric parameters of the ovaries and uterus and disorders of the ovarian cycle before the F4 generation. As A.L. Levinson et al. (2022) note, "Despite the fact that research on the effects of psycho-emotional stress is widely represented in both the medical and scientific biological literature devoted to experiments on laboratory models, these two areas are developing largely independently". It can be suggested that failures in the use of ART may be caused, among other things, by the transgenerational effects of stress.

In various experimental models of transgenerational transfer of epigenetic changes, fertility disorders of female offspring over several generations are noted (Guilbert et al., 2012; Moisiadis et al., 2017; Adams, Smith, 2020). In addition, the effect of prenatal stress on transgenerational changes in male and female offspring is different (Grundwald, Brunton, 2015; Zaidan, Gaisler-Salomon, 2015; Zhang et al., 2020; Huerta-Cervantes et al., 2021). At the same time, both prenatally stressed mothers and fathers have an impact.

According to our data, the behavior of control group 1 females in diestrus is characterized by increased anxiety indicators, whereas in estrus anxiety is reduced and locomotor and research activity is increased. These data correspond to earlier studies (Mora et al., 1996; Marcondes et al., 2001; Miller et al., 2021), where behavioral changes in estrus and diestrus are associated with a different hormonal profile. Obviously, the relationship is the result of the fact that receptors for sex hormones are present in the structures of the brain, causing evolutionarily appropriate behavioral reactions associated with reproduction (Reznikov et al., 2004).

A change in the lipid peroxidation level is considered an important indicator of membrane destabilization (Levitsky, Gubsky, 1994) and can cause an alteration of the molecular structure of membranes, which in turn is expressed as a change in behavior (Moisiadis et al., 2017). At the same time, lipid peroxidation processes, occurring within the physiological norm, represent a mechanism for regulating the physicochemical state of membranes and, accordingly, structures associated with membranes – receptors and ion channels (Halliwell, Gutteridge, 2007). The data obtained in our study allow us to conclude (based on changes in the Schiff bases level) that in estrus there is a decrease in viscosity and an increase in plasticity of membranes in the neocortex and the reverse changes occur in the hippocampus in rats. In the hypothalamus in estrus, changes in the level of lipid peroxidation, and, accordingly, changes in the physicochemical state of the membranes are similar to those in the neocortex, but more pronounced. Apparently, changes in behavior in estrus compared to diestrus require appropriate changes in membranes and related structures (receptors and ion channels).

When interpreting the results, the model of "father prenatal stress" is more understandable for studying the mechanisms of epigenetic transfer than "mother prenatal stress", which has additional effects on F2 offspring by maternal behavior, childbirth and lactation (Bale, 2015). The detection of an altered phenotype in the case of an experiment with a prenatally stressed father can be considered valid evidence of transgenerational transfer of epigenetic changes in the second generation (Dunn et al., 2011).

Our study showed that in group 2, where one of the parents is a prenatally stressed father, behavior at different stages of the ovarian cycle does not correspond to the goal of reproductive behavior, accordingly, we can make an indirect conclusion about a disturbance of the hormonal regulation of sexual behavior. Considering the results of the lipid peroxidation change, we see that the physico-chemical properties of the neocortex membranes in rats in diestrus are characterized by an increased level of membrane plasticity in terms of the Schiff bases level in comparison with the control. In the hippocampus, changes in different lipid peroxidation products are multidirectional compared with the control, but taking into account such an indicator as the Klein index, which characterizes the degree of lipid oxidation, it can be concluded that the level of lipid peroxidation in diestrus increases compared with the control group. At the same time, in the hypothalamus, changes in lipid peroxidation indicators in estrus compared with diestrus are similar to the control, but more expressed.

Thus, changes in the level of lipid peroxidation in the studied brain structures probably also contribute to changes in the sexual behavior of females in a group where one of the parents is a prenatally stressed father. It can be concluded that the father's prenatal stress makes an important contribution to the reproductive pattern of the daughters, including through biochemical processes associated with the oxidation of biomolecules.

Maternal prenatal stress is an additional stressor because F2 offspring are cared for by a female with impaired maternal behavior (Graf et al., 2012). However, the behavior of group 3, where the mother was subjected to prenatal stress, demonstrates a distortion of the behavior of females in both estrus and diestrus, similar to group 2. Lipid peroxidation processes in the neocortex demonstrate an imbalance between the initial and final products, due to which changes occur compared to the control. Thus, a difference between estrus and diestrus appears in the indicators of conjugated dienes and the Klein index when these indicators decrease relative to the control ones. While the indicator of the final products of the lipid peroxidation – Schiff bases – in the estrus stage exceeds the values in the control group, as a result of which the interstage difference disappears.

In the hippocampus, the main changes relate to a decrease in the lipid peroxidation indices in diestrus relative to the control, whereas in the hypothalamus, on the contrary, the lipid peroxidation indices increase in estrus relative to the control values. Apparently, behavioral disorders at different stages of the ovarian cycle may occur in this group, including as a result of changes in the physico-chemical properties of the membranes of the researched structures: an imbalance in the neocortex and interstage distortions of the lipid peroxidation processes in the hippocampus and hypothalamus (an increase in the plasticity of the hippocampal membranes in diestrus and a decrease in the plasticity of the hypothalamus membranes in estrus). Thus, a prenatally stressed mother affects changes in the reproductive pattern of daughters differently at diffe-

rent stages of the cycle. Likely, possible epigenetic changes in F2 females are also influenced by disturbances in maternal behavior in prenatally stressed F1 females.

There is information in the literature on the cumulative effect of prenatal stress of both parents on offspring (Adams, Smith, 2020). In our studies in group 4, where both parents were exposed to prenatal stress, the behavior of female F2 offspring shows an increase in anxiety in both estrus and diestrus. Lipid peroxidation in the neocortex of this group undergoes changes compared to the control group, due to which there is a significant interstadial difference in all indicators of lipid peroxidation. A similar profile of interstadial differences is observed in the hippocampus. In the hypothalamus, Schiff bases levels are inverted by stage relative to the control group. It can be assumed that one of the reasons for the increase in anxiety, regardless of the stage of the ovarian cycle in this group, may be a change in lipid peroxidation processes in the neocortex and hippocampus. Perhaps the cumulative effect of prenatal stress of both parents is manifested in this group by an unambiguous change in behavior at both stages of the cycle and impairment of lipid peroxidation in the neocortex.

Conclusion

Our results revealed the transgenerational effect of prenatal stress on the processes of lipid peroxidation and the behavior of female rats of the F2 generation, depending on the stage of the ovarian cycle.

The prenatal stress of the father or mother changes the processes of lipid peroxidation in the neocortex, hippocampus and hypothalamus of a female rat of the F2 generation in such a way that physico-chemical properties of the membranes of these brain structures do not correspond to the goals of the ovarian cycle stages. While the prenatal stress of the father causes the greatest changes in the hypothalamus lipid peroxidation processes, and the prenatal stress of the mother – those in the neocortex. The behavior in both cases does not meet the objectives of reproduction.

Prenatal stress of both parents of the female rats of the F2 generation has the greatest effect on the changes in lipid peroxidation processes in the studied brain structures, reducing the intensity of lipid peroxidation. The behavior is characterized by increased anxiety in both the researched ovarian cycle stages.

References

Adams R.C.M., Smith C. *In utero* exposure to maternal chronic inflammation transfers pro-inflammatory profile to generation F2 via sex-specific mechanisms. *Front. Immunol.* 2020;11:48. DOI 10.3389/fimmu.2020.00048

Aiken C.E., Tarry-Adkins J.L., Ozanne S.E. Transgenerational developmental programming of ovarian reserve. *Sci. Rep.* 2015;5:16175. DOI 10.1038/srep16175

Aiken C.E., Tarry-Adkins J.L., Spiroski A., Nuzzo A.M., Ashmore T.J., Rolfo A., Sutherland M.J., Camm E.J., Giussani D.A., Ozanne S.E. Chronic gestational hypoxia accelerates ovarian aging and lowers ovarian reserve in next-generation adult rats. *FASEB J.* 2019;33(6):7758-7766. DOI 10.1096/fj.201802772R

Arutyunyan A.V., Dubinina E.E., Zybina N.N. Methods of Evaluation of Free-radical Oxidation and the Antioxidant System. Saint Petersburg: Foliant Publ., 2000 (in Russian)

Babenko O., Kovalchuk I., Metz G.A.S. Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. *Neurosci. Biobehav. Rev.* 2015;48:70-91. DOI 10.1016/j.neubiorev.2014.11.013

Bale T.L. Lifetime stress experience: transgenerational epigenetics and germ cell programming. *Dialogues Clin. Neurosci*. 2014;16(3):297-305. DOI 10.31887/DCNS.2014.16.3/tbale

Bale T.L. Epigenetic and transgenerational reprogramming of brain development. *Nat. Rev. Neurosci.* 2015;16(6):332-344. DOI 10.1038/nrn3818

Baraboy V.A., Brekhman I.I., Golotin V.G., Kudriashov Yu.B. Peroxidation and Stress. Saint Petersburg: Nauka Publ., 1992 (in Russian)

Bidlack W.R., Tappel A.L. Fluorescent products of phospholipids during lipid peroxidation. *Lipids.* 1973;8(4):203-207. DOI 10.1007/BF02544636

Brunton P.J. Effects of maternal exposure to social stress during pregnancy: consequences for mother and offspring. *Reproduction*. 2013;146(5):R175-R189. DOI 10.1530/REP-13-0258

Dennery P.A. Oxidative stress in development: nature or nurture? *Free Radic. Biol. Med.* 2010;49(7):1147-1151. DOI 10.1016/j.freeradbiomed.2010.07.011

Dunn G.A., Morgan C.P., Bale T.L. Sex-specificity in transgenerational epigenetic programming. *Horm. Behav.* 2011;59(3):290-295. DOI 10.1016/j.yhbeh.2010.05.004

Dyban A.P. Early Development of Mammals. Leningrad, 1988 (in Russian)

Essex M.J., Boyce W.T., Hertzman C., Lam L.L., Armstrong J.M., Neumann S.M.A., Kobor M.S. Epigenetic vestiges of early developmental adversity: childhood stress exposure and DNA methylation in adolescence. *Child Dev.* 2013;84(1):58-75. DOI 10.1111/j.1467-8624.2011.01641.x

Graf A.V., Dunaeva T.Y., Maklakova A.S., Maslova M.V., Sokolova N.A., Trofimova L.K. Transgenerational consequences of acute antenatal stress in pregnant rats. *Rossiyskiy Fiziologicheskiy Zhurnal imeni Ivana Mikhaylovicha Sechenova = Russian Journal of Physiology.* 2012;98(3):331-341 (in Russian)

Grundwald N.J., Brunton P.J. Prenatal stress programs neuroendocrine stress responses and affective behaviors in second generation rats in a sex-dependent manner. *Psychoneuroendocrinology.* 2015;62:204-216. DOI 10.1016/j.psyneuen.2015.08.010

Guilbert F., Lumineau S., Kotrschal K., Mostl E., Richard-Yris M., Houdelier C. Trans-generational effects of prenatal stress in quail. *Proc. Biol. Sci.* 2012;280(1753):20122368. DOI 10.1098/rspb.2012.2368

Halliwell B., Gutteridge J.M.C. Free Radicals in Biology and Medicine. New York: Oxford University Press, 2007

Huerta-Cervantes M., Peña-Montes D.J., López-Vázquez M.A., Montoya-Pérez R., Cortés-Rojo C., Olvera-Cortés M.E., Saavedra-Molina A. Effects of gestational diabetes in cognitive behavior, oxidative stress and metabolism on the second-generation off-spring of rats. *Nutrients.* 2021;13(5):1575. DOI 10.3390/nu13051575

Levinson A.L., Igonina T.N., Rozhkova I.N., Brusentsev E.Yu., Amstislavsky S.Ya. Psycho-emotional stress, folliculogenesis, and reproductive technologies: clinical and experimental data. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2022;26(5):431-441. DOI 10.18699/VJGB-22-53 (in Russian)

Levitsky E.L., Gubsky J.I. Free radical damage of the nuclear genetic apparatus of cells. *Ukrainskiy Biokhimicheskiy Zhurnal = The Ukrainian Biochemical Journal.* 1994;66(4):18-30 (in Russian)

Marcondes F.K., Miguel K.J., Melo L.L., Spadari-Bratfisch R.C. Estrous cycle influences the response of female rats in the elevated plus-maze test. *Physiol. Behav*. 2001;74(4-5):435-440. DOI 10.1016/s0031-9384(01)00593-5

Matthews S.G., Phillips D.I. Transgenerational inheritance of stress pathology. *Exp. Neurol*. 2012;233(1):95-101. DOI 10.1016/j.expneurol.2011.01.009

Miller C.R., Halbing A.A., Patisaul H.B., Meitzen J. Interaction of the estrous cycle, novelty and light on female and male rat open field locomotor ant anxiety-related behaviors. *Physiol. Behav.* 2021;228:113203. DOI 10.1016/j.physbeh.2020.113203

Moisiadis V.G., Constantinof A., Kostaki A., Szyf M., Matthews S. Prenatal glucocorticoid exposure modifies endocrine function and behavior for 3 generations following maternal and paternal transmission. *Sci. Rep.* 2017;7(1):11814. DOI 10.1038/s41598-017-11635-w

Mora S., Dussaubat N., Diaz-Veliz G. Effects of the estrous cycle and ovarian hormones on behavioral indices of anxiety in female rats. *Psychoneuroendocrinology.* 1996;21(7):609-620. DOI 10.1016/s0306-4530(96)00015-7

Ordyan N.E., Pivina S.G. Behavioral characteristics and stress reaction of the pituitary-adrenal system in prenatally stressed rats. *Rossiyskiy Fiziologicheskiy Zhurnal imeni Ivana Mikhaylovicha Sechenova = Russian Journal of Physiology.* 2003;89(1):52-59 (in Russian)

Piquer B., Ruz F., Barra R., Lara H.E. Gestational sympathetic stress programs the fertility of offspring: a rat multi-generation study. *Int. J. Environ. Res. Public Health.* 2022;19(5):3044. DOI 10.3390/ijerph19053044

Pivina S.G., Rakitskaya V.V., Shamolina T.S., Ordyan N.E. Change of the uterus morphometric parameters in the prenatally stressed rats. *Rossiyskiy Fiziologicheskiy Zhurnal imeni Ivana Mikhaylovicha Sechenova = Russian Journal of Physiology.* 2010;96(6):621-626 (in Russian)

Provençal N., Binder E.B. The effect of early life stress on the epigenome: from the womb to adulthood and even before. *Exp. Neurol.* 2015;268:10-20. DOI 10.1016/j.expneurol.2014.09.001

Reznikov A.G., Pishak V.P., Nosenko N.D., Tkachyuk S.S., Myslitskiy V.F. Prenatal Stress and Neuroendocrine Pathology. Chernivtsi: Medacademia Publ., 2004 (in Russian)

Rice D., Barone S. Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environ. Health Perspect*. 2000;108(Suppl. 3):511-533. DOI 10.1289/ehp.00108s3511

Rodgers A.B., Bale T.L. Germ cell origins of posttraumatic stress disorder risk: the transgenerational impact of parental stress experience. *Biol. Psychiatry.* 2015;78(5):307-314. DOI 10.1016/j.biopsych.2015.03.018

Sanches-Garrido M.A., Garcia-Galiano D., Tena-Sempere M. Early programming of reproductive health and fertility: novel neuroendocrine mechanisms and implications in reproductive medicine. *Hum. Reprod. Update.* 2022;28(3):346-375. DOI 10.1093/humupd/dmac005

Thompson L.P., Al-Hasan Y. Impact of oxidative stress in fetal programming. *J. Pregnancy.* 2012;2012:582748. DOI 10.1155/2012/582748

Yao S., Lopes-Tello J., Sferruzzi-Perri A.N. Developmental programming of the female reproductive system – a review. *Biol. Reprod.* 2021;104(4):745-770. DOI 10.1093/biolre/ioaa232

Zaidan H., Gaisler-Salomon I. Prereproductive stress in adolescent female rats affects behavior and corticosterone levels in second-generation offspring. *Psychoneuroendocrinology.* 2015;58:120-129. DOI 10.1016/j.psyneuen.2015.04.013

Zhang H.-L., Yi M., Li D., Li R., Zhao Y., Qiao J. Transgenerational inheritance of reproductive and metabolic phenotypes in PCOS rats. *Front. Endocrinol*. 2020;11:144. DOI 10.3389/fendo.2020.00144

Conflict of interest. The authors declare no conflict of interest.
Received September 14, 2023. Revised December 19, 2023. Accepted February 9, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):398-406

DOI 10.18699/vjgb-24-45

Effect of amisulpride on the expression of serotonin receptors, neurotrophic factor BDNF and its receptors in mice with overexpression of the aggregation-prone [R406W] mutant tau protein

E.M. Kondaurova[#], A.A. Komarova[#], T.V. Ilchibaeva, A.Ya. Rodnyy, E.A. Zalivina, V.S. Naumenko

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
kond_em@bionet.nsc.ru

Abstract. Serotonin 5-HT$_7$ receptors (5-HT7R) are attracting increasing attention as important participants in the mechanisms of Alzheimer's disease and as a possible target for the treatment of various tau pathologies. In this study, we investigated the effects of amisulpride (5-HT7R inverse agonist) in C57BL/6J mice with experimentally induced expression of the gene encoding the aggregation-prone human Tau[R406W] protein in the prefrontal cortex. In these animals we examined short-term memory and the expression of genes involved in the development of tauopathy (*Htr7* and *Cdk5*), as well as biomarkers of neurodegenerative processes – the *Bdnf* gene and its receptors TrkB (the *Ntrk2* gene) and p75^{NTR} (the *Ngfr* gene). In a short-term memory test, there was no difference in the discrimination index between mice treated with AAV-Tau[R406W] and mice treated with AAV-EGFP. Amisulpride did not affect this parameter. Administration of AAV-Tau[R406W] resulted in increased expression of the *Htr7*, *Htr1a*, and *Cdk5* genes in the prefrontal cortex compared to AAV-EGFP animals. At the same time, amisulpride at the dose of 10 mg/kg in animals from the AAV-Tau[R406W] group caused a decrease in the *Htr7*, *Htr1a* genes mRNA levels compared to animals from the AAV-Tau[R406W] group treated with saline. A decrease in the expression of the *Bdnf* and *Ntrk2* genes in the prefrontal cortex was revealed after administration of AAV-Tau[R406W]. Moreover, amisulpride at various doses (3 and 10 mg/kg) caused the same decrease in the transcription of these genes in mice without tauopathy. It is also interesting that in mice of the AAV-EGFP group, administration of amisulpride at the dose of 10 mg/kg increased the *Ngfr* gene mRNA level. The data obtained allow us to propose the use of amisulpride in restoring normal tau protein function. However, it should be noted that prolonged administration may result in adverse effects such as an increase in *Ngfr* expression and a decrease in *Bdnf* and *Ntrk2* expression, which is probably indicative of an increase in neurodegenerative processes.

Key words: Alzheimer's disease; tau protein; amisulpride; 5-HT$_7$ receptor; Cdk5 kinase; *Bdnf*; *Ngfr*; *Ntrk2*; mice.

For citation: Kondaurova E.M., Komarova A.A., Ilchibaeva T.V., Rodnyy A.Ya., Zalivina E.A., Naumenko V.S. Effect of amisulpride on the expression of serotonin receptors, neurotrophic factor BDNF and its receptors in mice with overexpression of the aggregation-prone [R406W] mutant tau protein. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):398-406. DOI 10.18699/vjgb-24-45

Funding. The work was supported by the Russian Science Foundation (grant No. 22-15-00011).

Acknowledgements. The cost of animal housing was compensated by basic research project No. FWNR-2022-0023.

Действие амисульприда на экспрессию серотониновых рецепторов, нейротрофического фактора BDNF и его рецепторов при сверхэкспрессии склонного к агрегации тау-белка с мутацией [R406W] у мышей

Е.М. Кондаурова[#], А.А. Комарова[#], Т.В. Ильчибаева, А.Я. Родный, Е.А. Заливина, В.С. Науменко

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
kond_em@bionet.nsc.ru

Аннотация. Серотониновые рецепторы 5-HT$_7$ (5-HT7R) привлекают все больше внимания в качестве одного из важных звеньев в механизмах развития болезни Альцгеймера и возможной мишени для лечения различных тау-патологий. В настоящей работе исследовано влияние амисульприда (обратный агонист 5-HT7R) в мо-

[#] Equal contributed

дели экспериментального повышения экспрессии гена, кодирующего склонный к агрегации белок человека Tau[R406W], в префронтальной коре мышей линии C57BL/6J на кратковременную память и экспрессию генов, участвующих в развитии таупатии (*Htr7* и *Cdk5*), а также биомаркеров нейродегенеративных процессов – гена *Bdnf* и его рецепторов TrkB (ген *Ntrk2*) и $p75^{NTR}$ (ген *Ngfr*). В тесте на кратковременную память мыши не было обнаружено разницы по индексу дискриминации между мышами, которым вводили AAV-Tau[R406W], и мышами с AAV-EGFP. Амисульприд не повлиял на данный показатель. Введение AAV-Tau[R406W] привело к повышению экспрессии генов *Htr7*, *Htr1a* и *Cdk5* в префронтальной коре по сравнению с животными группы AAV-EGFP. При этом амисульприд в дозе 10 мг/кг у животных группы AAV-Tau[R406W] вызвал снижение уровня мРНК генов *Htr7* и *Htr1a* по сравнению с животными группы AAV-Tau[R406W], которым вводили физиологический раствор. Выявлено снижение экспрессии генов *Bdnf* и *Ntrk2* в префронтальной коре после введения AAV-Tau[R406W]. При этом амисульприд в различных дозах (3 и 10 мг/кг) вызывал такое же снижение транскрипции этих генов у мышей без таупатии. Интересно также, что у мышей группы AAV-EGFP после введения амисульприда в дозе 10 мг/кг повышался уровень мРНК гена *Ngfr*. Полученные данные позволяют рассматривать амисульприд в качестве агента для восстановления нормальной функции тау-белка. Однако следует учитывать возможный негативный эффект амисульприда при длительном применении, отражающийся в увеличении экспрессии гена *Ngfr* и снижении экспрессии генов *Bdnf* и *Ntrk2*, что может указывать на усиление нейродегенеративных процессов.

Ключевые слова: болезнь Альцгеймера; тау-белок; амисульприд; 5-HT_7-рецептор; киназа Cdk5; *Bdnf*; *Ngfr*; *Ntrk2*; мыши.

Introduction

It is well known that tau protein plays an important role in the maintenance of axonal structure and growth, as well as regulates the formation of neuronal polarity, axonal transport and neuroplasticity (Arendt et al., 2012). However, hyperphosphorylated tau protein loses its normal ability to stabilize microtubules in cells and aggregates in pathomorphological structures – paired helical filaments and neurofibrillary tangles (Grundke-Iqbal et al., 1986). This leads to dysfunction of tau protein and causes various tauopathies, including Alzheimer's disease (AD).

Currently, more than 50 different pathogenic mutations of the *MAPT* gene encoding tau protein have been detected. Most of these mutations are in exons and occur in regions encoding the C-terminal microtubule-binding domain (Strang et al., 2019). Mutations in coding sequences are mostly missense mutations, although there are also data on deletions (Rovelet-Lecrux et al., 2009). The most common manifestations of these mutations are impaired binding to microtubules and, as a consequence, their dysfunction, while the effect on the tau protein aggregation *in vivo* is observed only for some mutations (Xia et al., 2019).

Tau[R406W] is one of the *MAPT* gene mutations that promotes protein aggregation due to the reduced ability of the phosphorylated form to bind to microtubules (Perez et al., 2000). This mutation (located in exon 13 of the *MAPT* gene) results in the replacement of arginine with tryptophan at position 406 (p.R406W) and causes familial frontotemporal lobar degeneration with tau pathology (FTLD-tau). The frequency of the p.R406W mutation is 0.62 % among patients with FTLD-tau and 0.26 % among patients with AD (Gossye et al., 2023). The location of this mutation near the MTBR (microtubule-binding region) may affect the ability of this region to cause conformational changes in the neighboring MTBR (Xia et al., 2019). An important fact is that the R406W mutation is located near to key amino acid residues (Ser396, Ser404) that are phosphorylated in tau protein during the formation of pathological paired helical filaments (Hutton et al., 1998).

On the other hand, it is known that the brain serotonin (5-HT) system also plays an important role in the pathological development and clinical manifestations of primary tauopathies, including frontotemporal dementia, progressive supranuclear palsy and corticobasal degeneration (Huey et al., 2006; Murley, Rowe, 2018). The function of the 5-HT system is realized through numerous receptors. Nowadays, there is a growing number of studies investigating the role of 5-HT receptors in the mechanisms of tauopathies and AD development (Eremin et al., 2023).

In this regard, the 5-HT_7 receptor (5-HT7R) has attracted particular attention. Recent studies have demonstrated that the constitutive activity of 5-HT7R induces hyperphosphorylation of tau protein and its subsequent aggregation through interaction with CDK5 kinase. Moreover, administration of the highly selective 5-HT7R inverse agonist SB-269970 prevents receptor-induced accumulation and hyperphosphorylation of tau protein (Labus et al., 2021).

Also, it has been shown that amisulpride (a drug with antipsychotic, antidepressant and procognitive effects), a strong inverse agonist of 5-HT7R, is able to affect the hyperphosphorylation of tau protein. The therapeutic potential of amisulpride in preventing/dispersing tau aggregation and tau-mediated pathology has been confirmed *in vitro* (in Tau-BiFC HEK293 cells and in human cortical neurons with the Tau[R406W] mutation) and *in vivo* (in mice overexpressing human mutant Tau [R406W] protein in the prefrontal cortex, and in transgenic mice expressing human mutant Tau[P301L] protein). In these animal models of tauopathy, treatment with amisulpride prevented tau protein hyperphosphorylation, aggregation, and neurotoxicity, and reversed memory impairment in both mouse strains (Jahreis et al., 2023).

In addition, it was shown that chronic administration of amisulpride in OXYS rats (a model of sporadic AD) (Stefanova et al., 2015) reduced phosphorylation of tau protein in the cortex and hippocampus of 3-month-old animals (Molobekova et al., 2023). Besides, in the hippocampus of 1- and 3-month-old rats, amisulpride also reduced the mRNA level of the *Cdk5* kinase gene (Molobekova et al., 2023).

It is well known that the progression of tauopathies and AD causes the development of nerve cells atrophy in the cerebral cortex, hippocampus and other subcortical structures (Bettens et al., 2010). Thus, it was shown that the Tau[R406W] mutation causes disturbances in genes associated with neurogenesis and synaptic function in mouse neurons (Minaya et al., 2023). Among the biomarkers of neurodegenerative processes, the brain-derived neurotrophic factor (BDNF) is well known. The decrease in BDNF mRNA and protein levels in the cerebral cortex and hippocampus was shown in AD (Hock et al., 2000). BDNF-induced neuronal growth and development are mediated by its receptors, tyrosine kinase receptor B (TrkB) and common neurotrophin receptor p75 ($p75^{NTR}$), which bind with BDNF and proBDNF, respectively. Accumulating evidence indicates the cross-talk between 5-HT and BDNF, suggesting that both systems may control each other's functions by acting through shared intracellular signaling pathways. Balance in the functioning of the 5-HT and BDNF systems appears to be fundamental for the development of a normal phenotype (Popova, Naumenko, 2019).

Thus, the aim of the study was to investigate the effects of amisulpride in mice with experimentally induced expression of the Tau[R406W] gene (using an adeno-associated viral construct *in vivo*) in the prefrontal cortex on short-term memory and on the expression of genes that are involved in the development of tauopathy (*Htr7* and *Cdk5*), as well as the gene of BDNF and its receptors (*Ntrk2* (encodes TrkB) and *Ngfr* (encodes $p75^{NTR}$)).

Materials and methods

Animals. Experiments were carried out on 2-month-old C57BL/6J male mice. Work with animals was performed at the Center for Genetic Resources of Laboratory Animals, Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, under standard conditions of a conventional vivarium (grant of the Russian Ministry of Science and Higher Education No. RFMEFI62119X0023). Animals were kept and tested in accordance with the Instructions for the Care and Use of Laboratory Animals (National Institute of Health's Guide for the Care and Use of Laboratory Animals, NIH Publications, 2010).

Plasmids. Plasmids carrying the Tau[R406W] and EGFP (EGFP – enhanced green fluorescent protein) genes or only EGFP under control of the synapsin promoter were obtained from Professor E.G. Ponimaskin (MHH, Hannover, Germany).

Cell transfection. Packaging of pAAV_SynH1-2_Tau[R406W]-EGFP and pAAV_SynH1-2_EGFP plasmids to adeno-associated viral (AAV) capsids was performed by their co-transfection with AAV-DJ and pHelper plasmids (Cell Biolabs, Inc., USA) into HEK293FT cells that were incubated according to the protocol described previously (Kondaurova et al., 2021). Viral particles were collected after 48 h according to the protocol described previously (Grimm et al., 2003). The number of viral particles obtained was determined by quantitative real-time PCR analysis and diluted to a concentration of 10^9 viral particles/μl.

Stereotactic injection. Before the procedure, the animals were anesthetized with a mixture of 2,2,2-tribromoethanol and 2-methyl-2-butanol and placed in a stereotaxic frame (TSE Systems, Germany). Briefly, the scalp was opened, and two holes were drilled in the skull: AP: +1.5 mm, LR: ±1 mm, DV: 1 mm (http://labs.gaidi.ca/mouse-brain-atlas/?ml=1.5&ap=-2&dv=2). Mice of both groups (36 males "AAV-Tau[R406W]" and "AAV-EGFP") were bilaterally injected with the AAV-Tau[R406W]-EGFP or AAV-EGFP viral construct into the prefrontal cortex. After the bilateral injections of the virus, the incision was closed with interrupted silk sutures, and the animal was placed in a warm cage and monitored closely (Kondaurova et al., 2021).

Pharmacological administration. Seven days after AAV administration, each group was divided into three subgroups (12 mice per subgroup). The effect of chronic amisulpride administration (Sanofi-Aventis, France) was assessed after 4 weeks of intraperitoneal administration in the doses of 3 and 10 mg/kg for the first and second subgroups, respectively, in a volume of 10 μl/g. Animals of the third subgroup were treated with the same volume of saline. Two days before the experiments, mice were placed in individual cages to remove the group effect.

"Open field" test. This test was carried out in a circular arena (40 cm in diameter) surrounded by a white plastic wall (25 cm high) and illuminated through a mat and semitransparent floor with two halogen lamps of 12 W each placed 40 cm under the floor (Kulikov et al., 2008). Each mouse was placed near the wall and tested for 5 min. The animal's behavior was recorded for 5 minutes using a camera located at a distance of 80 cm from the arena. The arena was treated with 70 % alcohol after each test. The video stream from the camera was analyzed frame by frame using the original EthoStudio software (Khotskin et al., 2019). The path length was measured automatically (horizontal activity).

Short-term memory test ("recency test"). This test was performed within the framework of the "open field" test paradigm. At the first stage of the "recency test", the animals were familiarized with two identical objects (plastic cubes measuring 5 × 5 cm – "old object"), located in the center of the arena at a distance of 8–10 cm from each other and 10 cm from the walls of the arena. Animals were tested for 10 min. 90 min after the first test, two other objects (plastic cups with a diameter of 4 cm and a height of 5 cm – "new object") were presented. These objects were located in the center of the arena at a distance of 8–10 cm from each other and 10 cm from the walls of the arena. Animals were tested for 10 min. 90 min after the second stage, one of the presented objects was replaced with the first object (plastic cubes, 5 × 5 cm). The animal was tested for 10 minutes. The time required to approach the new and old objects was assessed. Then discrimination index was calculated using the formula: (time for the "new object" – time for the "old object") / total time for both objects.

Dissection of brain samples. In 48 h after behavioral testing, animals were removed from the experiment by decapitation. Immediately after euthanasia, the brain was removed and the necessary brain structures (prefrontal cortex, hippocampus) were excised on ice, frozen in liquid nitrogen. Until further procedures, the structures were stored in a low-temperature refrigerator at –80 °C.

Fluorescence Microscopy. At least 6 weeks after AAV injection, one or two mice from each group were transcardially

perfused for 2 min with 20 mL of phosphate-buffered saline (PBS) and 20 mL of a 4 % paraformaldehyde solution for 10 min under anesthesia. The brain was removed and post-fixed with 4 % paraformaldehyde for 16 h and immersed in 30 % sucrose in PBS for 2 days. Sequential 12 µm slices were prepared on a cryostat (Thermo Scientific, Germany). Cell nuclei were stained with a bis-benzimide solution (Hoechst 33258 dye, 5 µg/mL in PBS; Sigma-Aldrich, Germany). The sections were then mounted in antiquenching Fluoromount G medium (Southern Biotechnology Associates, USA) followed by examination using an Olympus IX83P2ZF confocal microscope.

RNA isolation. The brain structures were homogenized in 300 µl TRIzol Reagent (Life Technologies, USA) according to the manufacturer's protocol. The total RNA was dissolved in 24.5 µl of water treated with diethyl pyrocarbonate (DEPC). To eliminate possible genomic DNA contaminations, 0.5 µl of DNase (RNase-free DNase, Promega, USA, 1,000 p. u./ml) was added. The samples were incubated for 15 minutes at 37 °C, and then for 10 minutes at 65 °C. The RNA concentrations were determined using a Nanodrop2000c spectrophotometer (Thermo Fisher Scientific), and diluted to 125 ng/µl. The RNA was stored at –80 °C.

Real-time RT-PCR. The gene expression was determined using a quantitative reverse transcription-polymerase chain reaction (RT-PCR) developed in our laboratory (Kulikov et al., 2005; Naumenko, Kulikov, 2006; Naumenko et al., 2008). Two types of standards were used: external and internal. An internal standard (housekeeping genes *Polr2a* (RNA polymerase II gene) and *B2m* (β2-microglobulin gene)) was used to monitor reverse transcription and as a basis for calculating the mRNA levels of the target genes. Mouse DNA of a known concentration served as an external standard, which made it possible to control the PCR and determine the number of mRNA copies of the studied genes in the samples. To determine the mRNA levels, we used the ratio of the cDNA level of the studied genes to the geometric mean level of cDNA of the *rPol2a* and *B2m* genes.

Primers for cDNA amplification were selected based on sequences published in the EMBL nucleotide database and synthesized at the Bioset company (Novosibirsk, Russia). PCR was carried out on a Real-time CFX96 Touch cycler (Bio-Rad, USA) in accordance with the following protocol: 3 min at 95 °C; 40 cycles with three stages: 10 sec at 95 °C, 30 sec at the primer annealing temperature, 20 sec at 72 °C (Supplementary Material 1)[1].

Statistical Analysis. Statistical analysis was performed using GraphPadPrism 9.1.0. To search and exclude outliers form the analysis, the ROUT method (Q = 0.05) was used. The normal distribution of samples was tested using the Kolmogorov–Smirnov and Shapiro–Wilk tests. According to these criteria, all data have normal distribution. To identify differences between groups, a two-way ANOVA with post-hoc Fisher's multiple comparison was carried out. The results were presented as m ± SEM (m – mean; SEM – standard error of the mean). The statistical significance value was set at $p < 0.05$.

[1] Supplementary Materials 1 and 2 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Kond_Engl_28_4.pdf

Results

Fluorescence microscopy of mouse brain sections was used to verify the correct injection of the constructs. Brain sections of the prefrontal cortex area showed fluorescence (emission at 510 nm) when excited by light with a wavelength of 488 nm, which confirmed the successful expression of the viral construct into the brain structure (Fig. 1).

Based on the presence of the *MAPT* gene product (on average at PCR cycle 28), we confirmed the gene expression in the prefrontal cortex of mice treated with the AAV-Tau[R406W] construct, while transcription of this gene was not observed in control mice (AAV-EGFP) (Fig. 2).

In the "open field" test, the locomotor activity of mice was affected by both factors – the AVV construct ($F_{1.57} = 3.598$, $p = 0.063$) and the administration of amisulpride ($F_{2.57} = 4.580$, $p = 0.014$), as well as by the interaction of these factors ($F_{2.57} = 3.520$, $p = 0.036$). AAV-EGFP animals showed increased locomotor activity not only compared to AAV-Tau[R406W] mice ($p = 0.012$), but also compared to AAV-EGFP mice treated with amisulpride at the dose of 3 mg/kg ($p = 0.003$) and 10 mg/kg ($p = 0.013$) (Fig. 3).

Neither amisulpride ($F_{2.26} = 1.8$, $p > 0.05$), nor AAV administration ($F_{1.26} = 1.111$, $p > 0.05$) and their interaction ($F_{2.26} = 1.6$, $p > 0.05$) had a significant effect on the discrimination index values in the "recency test" (Fig. 4).

A significant effect of interaction of AAV and amisulpride was found in the *Htr7* mRNA level in the prefrontal cortex ($F_{2.44} = 7.059$, $p = 0.002$). Administration of the AAV-Tau[R406W] construct caused an increase in the *Htr7* gene transcription ($p = 0.020$). At the same time, we observed a restoration of the cortical *Htr7* gene mRNA level to normal values in the mice from the AAV-Tau[R406W] group that were treated by amisulpride at the dose of 10 mg/kg ($p = 0.014$) (Fig. 5*a*). Amisulpride administration at the concentration of 3 mg/kg did not evoke a similar effect ($p = 0.157$). Interesting that when the drug was administered at the dose of 10 mg/kg to AAV-EGFP mice the increased *Htr7* mRNA levels was observed ($p = 0.004$) (Fig. 5*a*).

For the *Htr1a* gene in the prefrontal cortex (Fig. 5*b*), similar differences were observed: the effect of interaction between the AAV and amisulpride factors ($F_{2.52} = 3.359$, $p = 0.043$) (Supplementary Material 2), a decrease in receptor gene transcription upon amisulpride treatment of AAV-Tau[R406W] mice ($p = 0.005$) at the dose of 10 mg/kg and an increase in the *Htr1a* mRNA levels in amisulpride (10 mg/kg) treated AAV-EGFP mice ($p = 0.044$).

When analyzing the *Cdk5* gene mRNA level in the prefrontal cortex, an interaction of factors was found ($F_{2.48} = 7.182$, $p = 0.002$) (Supplementary Material 2). We showed that the *Cdk5* mRNA level in mice from the AAV-Tau[R406W] group that were not exposed to amisulpride treatment was increased compared to the AAV-EGFP group ($p = 0.021$), while the effect of amisulpride at the dose of 10 mg/kg led to a decrease in *Cdk5* gene expression ($p = 0.004$) compared to the AAV-EGFP group. In addition, *Cdk5* transcription was increased by 10 mg/kg amisulpride in AAV-EGFP mice compared to AAV-EGFP saline-treated animals ($p = 0.001$) (Fig. 6).

We investigated the effect of amisulpride on the mRNA level of the brain-derived neurotrophic factor *Bdnf* and its

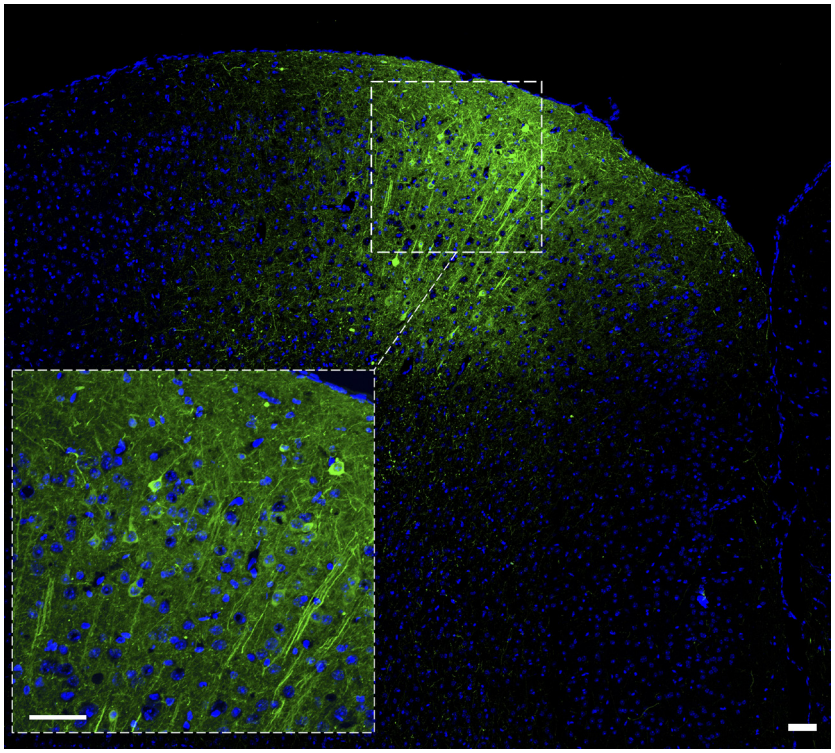


Fig. 1. Microphotographs of the mouse prefrontal cortex section demonstrating successful AAV-mediated cell transfection as indicated by EGFP (enhanced green fluorescent protein) expression.

Scale bar: 50 μm.

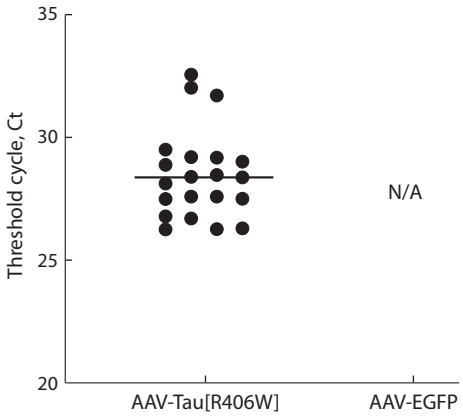


Fig. 2. Expression (Ct) of the *MAPT* gene in the prefrontal cortex of mice overexpressing Tau[R406W].

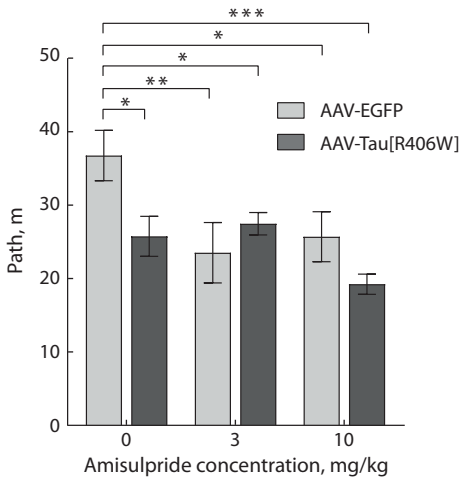


Fig. 3. Changes in the motor activity of mice in the "open field" test.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

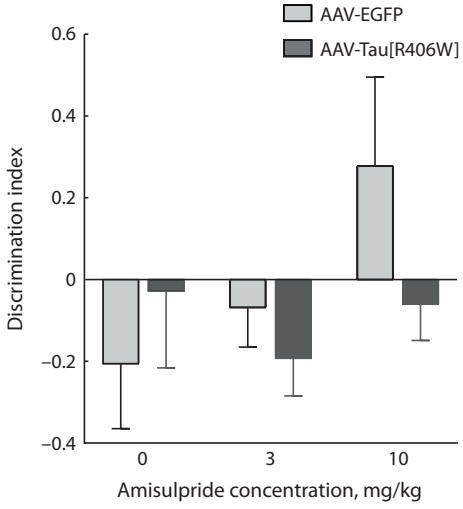


Fig. 4. Effect of amisulpride on the discrimination index in the "recency test".

receptors *Ntrk2* (encodes the TrkB receptor) and *Ngfr* (encodes the $p75^{NTR}$ receptor). In the prefrontal cortex for *Bdnf* mRNA, the effect of AAV administration, amisulpride treatment and their interaction was observed. For the *Ntrk2* gene, only AAV administration and the interaction of the AAV and amisulpride factors were found (Supplementary Material 2). *Bdnf* ($p < 0.001$) and *Ntrk2* ($p < 0.001$) mRNA levels were decreased by mutant Tau[R406W] overexpression compared with AAV-EGFP in saline-treated mice.

In addition, a decrease in the level of *Bdnf* ($p < 0.001$) and *Ntrk2* ($p = 0.037$) mRNAs was observed when the drug was administered at the dose of 3 mg/kg to AAV-EGFP mice, as well as when amisulpride was administered at the dose of 10 mg/kg to AAV-EGFP mice (for *Bdnf* ($p = 0.004$) and *Ntrk2* ($p = 0.045$)) (Fig. 7*a*, *c*). At the same time, the effect of interaction of the AAV and amisulpride treatment factors was observed for the *Ngfr* gene mRNA level in the prefrontal cortex ($F_{2.43} = 4.752$, $p = 0.014$) (Supplementary Material 2). The *Ngfr* gene expression in the cortex of AAV-EGFP mice increased upon administration of amisulpride at the concentration of 10 mg/kg ($p = 0.002$); however, mice overexpressing Tau[R406W] showed a decrease in the mRNA level of this gene when exposed to the same dose of the drug ($p = 0.002$) (Fig. 7*b*).

No statistically significant differences in the expression of all investigated genes were found in the hippocampus. In Supplementary Material 2, data from a two-factor analysis of variance are presented.

Discussion

Here it was shown that locomotor activity in the "open field" test was reduced both in the AAV-Tau[R406W] group that received saline, and in groups that were treated by amisulpride at different doses compared to the AAV-EGFP group that received saline. This data are consistent with the results of the work by K. Jahreis et al.: they showed that administration of Tau[R406W] vector

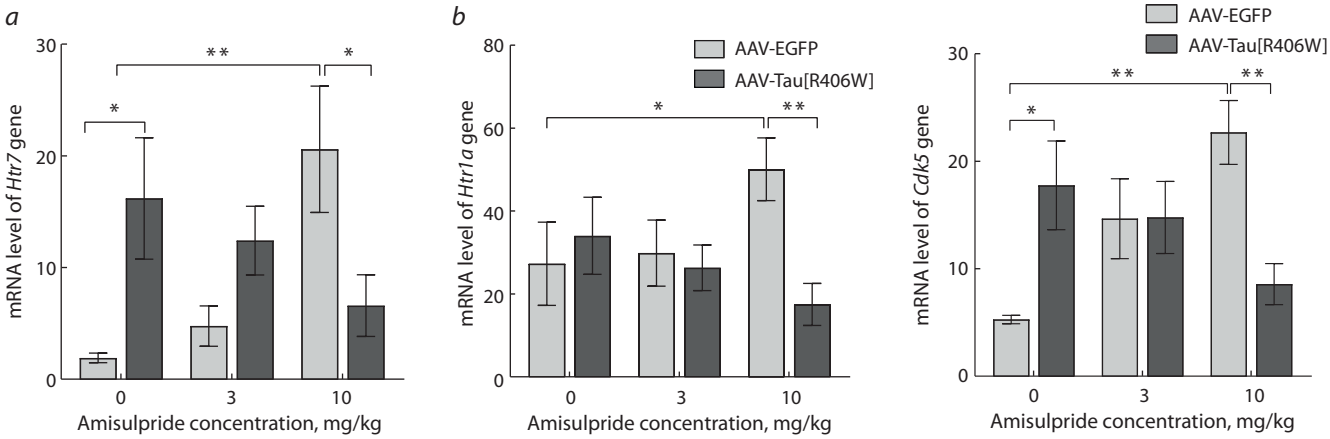


Fig. 5. Effect of amisulpride on the transcription of the *Htr7* (*a*) and *Htr1a* (*b*) genes in the prefrontal cortex of mice overexpressing Tau[R406W].

Values are normalized to the geometric mean of *Polr2a* and *B2m* mRNA.
* $p < 0.05$; ** $p < 0.01$.

Fig. 6. Effect of amisulpride on *Cdk5* gene transcription in the prefrontal cortex of tau [R406W] overexpressing mice.

Values are normalized to the geometric mean of *Polr2a* and *B2m* mRNA.
* $p < 0.05$; ** $p < 0.01$.

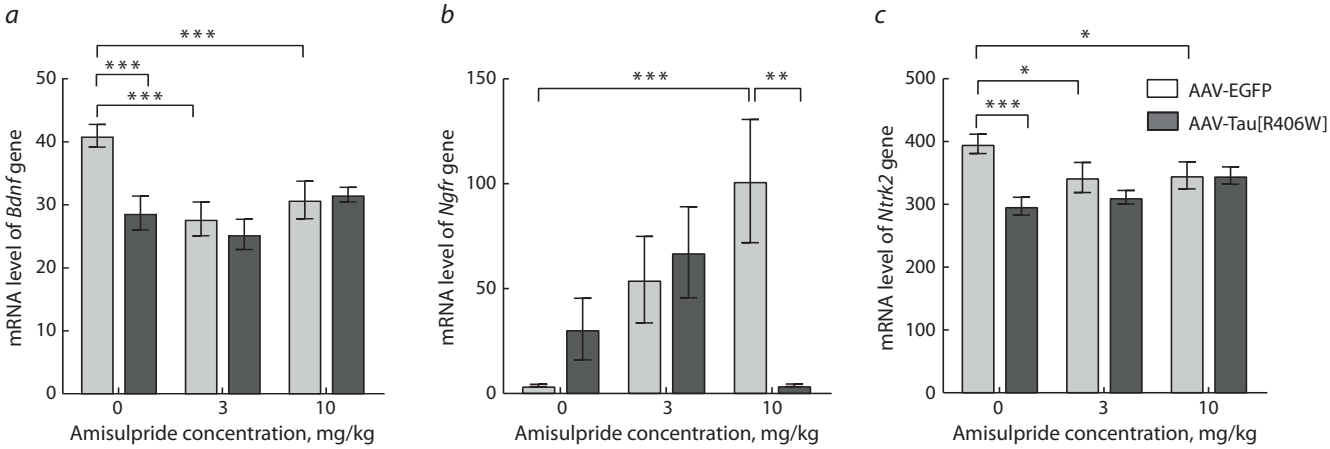


Fig. 7. Effect of amisulpride on the transcription of the *Bdnf* (*a*), *Ngfr* (*b*) and *Ntrk2* (*c*) genes in the prefrontal cortex of mice overexpressing Tau[R406W].

Values are normalized to the geometric mean of *Polr2a* and *B2m* mRNA.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

and treatment with amisulpride (1 mg/kg, 16 days) reduced locomotor activity in the "open field" test compared to control (Jahreis et al., 2023).

In the current study, amisulpride failed to produce a significant effect on the short-term memory of animals treated with AAV-Tau[R406W], in contrast to the paper of K. Jahreis et al., who showed an increase in the discrimination index in mice treated with AAV-Tau[R406W] and amisulpride (Jahreis et al., 2023). The lack of a significant amisulpride effect on short-term memory in our experiment may be due to a shorter recovery period after vector administration and before the beginning of amisulpride therapy. In our study, it was seven days, unlike the work of K. Jahreis et al., in which the recovery period took three weeks. Thus, the stage of tauopathy development is probably important for the amisulpride therapy of cognitive abilities.

We found that administration of AAV-Tau[R406W] leads to increased mRNA levels of the *Htr7* and *Cdk5* genes in the prefrontal cortex compared to control animals. These findings are likely due to a neuroprotective response involving increased levels of 5-HT7R, which is known to be involved in the regulation of neuronal morphology, neurite outgrowth, dendritic spines, and synaptogenesis (Kobe et al., 2012). However, a recent study has shown a reduced *Htr7* gene mRNA level in the anterior prefrontal cortex in postmortem brain samples from AD patients (Solas et al., 2021). This discrepancy can be explained by long-term neurodegenerative processes in the brains of AD patients, while in our work the effect of the mutant tau protein lasted only six weeks.

The increased transcription of the *Cdk5* gene in AAV-Tau[R406W] mice is consistent with a study of J. Labus and coauthors, who showed that CDK5 is responsible for the pathological effect of 5-HT7R on tau protein hyperphosphorylation (Labus et al., 2021). At the same time, the combined decrease in the mRNA levels of both *Htr7* and *Cdk5* in AAV-Tau[R406W] mice treated with amisulpride to values

similar to those in control animals confirms the proposed mechanism of 5-HT7R inverse agonists action in restoring normal tau protein function *in vivo*. The increase of the *Htr7*, *Htr1a* and *Cdk5* mRNA levels after amisulpride administration at the dose of 10 mg/kg in AAV-EGFP mice is probably a compensatory response to inhibition of the 5-HT_7 receptor by amisulpride. The effect of amisulpride on the *Cdk5* mRNA level is in good agreement with the data obtained on OXYS rats: in healthy one-month-old rats, amisulpride also increased the *Cdk5* mRNA level in the cortex (Molobekova et al., 2023).

It is known that 5-HT_{1A} (5-HT1AR) and 5-HT7R receptors can form heterodimers *in vitro* and *in vivo.* Such heterodimerization leads to agonist-mediated internalization of 5-HT_{1A} receptors (Renner et al., 2012). Chronic activation of 5-HT7R causes desensitization of these receptors and also reduces the level and functional activity of 5-HT_{1A} receptors in the frontal cortex, without affecting the level of 5-HT7R (Kondaurova et al., 2017). It has also been shown that overexpression of 5-HT7R in the midbrain leads to changes in 5-HT1AR gene expression depending on the mouse strain. In mice of the C57Bl/6J strain, a decrease in the 5-HT1AR gene mRNA level was detected in the frontal cortex, while in ASC (antidepressant sensitive cataleptics) mice, the expression of this gene was reduced in the hippocampus (Rodnyy et al., 2022). Amisulpride, as an inverse agonist of 5-HT_7 receptors, suppresses receptor constitutive activity and, perhaps, can thus influence the mRNA levels of the 5-HT7R gene in a negative feedback manner. It is interesting to note that chronic administration of amisulpride at the dose of 10 mg/kg led to an increase in the expression of both the 5-HT7R gene and the 5-HT1AR gene, which may be due to the mutual regulation of these receptors through their heterodimerization.

BDNF is one of the most studied neurotrophic factors. It plays an important role in the growth and maturation of brain cells at all stages of development, and is involved in the regulation of synaptic transmission and plasticity in adulthood (Edelmann et al., 2015). In the context of AD, BDNF depletion is associated with tau protein phosphorylation and aggregation, Aβ accumulation, neuroinflammation, and neuronal death (Pisani et al., 2023). BDNF stimulation leads to dephosphorylation of tau protein through TrkB activation and phosphatidylinositol 3-kinase (PI3K) signaling (Elliott et al., 2005).

In our study, we found a decrease in the mRNA levels of BDNF and its receptor TrkB in the cortex after administration of AAV-Tau[R406W]. These data are in agreement with the decrease in the BDNF level observed in AD (Song et al., 2015). In addition, we showed that amisulpride administration at different doses also reduces the mRNA levels of these genes in both AAV-EGFP and AAV-Tau[R406W] mice. These results contradict previous findings indicating that amisulpride increases BDNF levels in human neuroblastoma SH-SY5Y cells (Park et al., 2011). However, there is evidence that amisulpride does not affect the *Bdnf* mRNA level in another cell model – in T98G glioma cells (Jóźwiak-Bębenista et al., 2017). The work of E.N. Rizos et al. also did not reveal any effect of amisulpride on the BDNF level in the blood serum of patients with schizophrenia (Rizos et al., 2010). At the same time, an increase in the expression and phosphorylation of TrkB was detected 30 min after activation of 5-HT7R (Samarajeewa et al., 2014).

On the one hand, it can be assumed that the mechanisms of amisulpride action *in vitro* and *in vivo* are different. On the other hand, it has been shown that in human neuroblastoma SH-SY5Y cells, the elongation of nerve fibers caused by incubation with 5-HT, nerve growth factor (NGF) or brain-derived neurotrophic factor BDNF is blocked by 5-HT7R antagonists. The knockdown of the *Htr7* gene also reduces the length of nerve fibers, whereas 5-HT7R activation by agonists increases the expression of the NGF and BDNF genes (Chang et al., 2022).

A recent paper by L.L. Shen and colleagues has shown that knockout of $p75^{NTR}$ receptor leads to a reduction in Aβ-induced tau hyperphosphorylation and neurodegeneration both in healthy mice and in a mouse model of human tauopathy, involving CDK5 and GSK3β kinases (Shen et al., 2019). These data suggest that $p75^{NTR}$ receptor at least partially mediates Aβ peptide-triggered tau pathology. However, in our study, overexpression of Tau[R406W] did not have a significant effect on the $p75^{NTR}$ receptor mRNA level. At the same time, we found that amisulpride increases transcription of the $p75^{NTR}$ receptor gene in AAV-EGFP mice. There are other literature data on the negative effects of long-term amisulpride administration through a decrease in choline acetyltransferase (ChAT). G.B. Huang et al. demonstrated that long-term amisulpride administration (45 days) in rats reduced the number of ChAT-positive cells in the prefrontal cortex but not in hippocampus, which may have a negative effect on cognitive function (Huang et al., 2012).

Conclusion

Thus, the utilization of amisulpride in mice with Tau[R406W] overexpression led to a decrease in the *Htr7* and *Cdk5* genes mRNA level in the prefrontal cortex, which allowed us to suggest the drug as an agent for restoring normal tau protein function. However, the drug administration in mice without tauopathy caused a decrease in the *Bdnf* and *Ntrk2* genes mRNA levels in the frontal cortex. At the same time, the levels of *Htr7*, *Htr1a* and *Cdk5* mRNAs were increased in AAV-EGFP mice that were treated with the amisulpride. These changes probably reflect the negative effect of chronic amisulpride administration, which is also indirectly confirmed by an increase in the expression of the $p75^{NTR}$ receptor gene, which is known to initiate apoptotic processes in the brain.

References

Arendt D.H., Smith J.P., Bastida C.C., Prasad M.S., Oliver K.D., Eyster K.M., Summers T.R., Delville Y., Summers C.H. Contrasting hippocampal and amygdalar expression of genes related to neural plasticity during escape from social aggression. *Physiol. Behav.* 2012;107(5):670-679. DOI 10.1016/j.physbeh.2012.03.005

Bettens K., Sleegers K., Van Broeckhoven C. Current status on Alzheimer disease molecular genetics: from past, to present, to future. *Hum. Mol. Genet.* 2010;19(R1):R4-R11. DOI 10.1093/hmg/ddq142

Chang W.Y., Yang Y.T., She M.P., Tu C.H., Lee T.C., Wu M.S., Sun C.H., Hsin L.W., Yu L.C. 5-HT(7) receptor-dependent intestinal neurite outgrowth contributes to visceral hypersensitivity in irritable bowel syndrome. *Lab. Invest.* 2022;102(9):1023-1037. DOI 10.1038/s41374-022-00800-z

Edelmann E., Cepeda-Prado E., Franck M., Lichtenecker P., Brigadski T., Lessmann V. Theta burst firing recruits BDNF release and signaling in postsynaptic CA1 neurons in spike-timing-dependent LTP. *Neuron*. 2015;86(4):1041-1054. DOI 10.1016/j.neuron.2015.04.007

Elliott E., Atlas R., Lange A., Ginzburg I. Brain-derived neurotrophic factor induces a rapid dephosphorylation of tau protein through a PI-3Kinase signalling mechanism. *Eur. J. Neurosci*. 2005;22(5):1081-1089. DOI 10.1111/j.1460-9568.2005.04290.x

Eremin D.V., Kondaurova E.M., Rodny A.Ya., Molobekova K.A., Kudlay D.A., Naumenko V.S. Serotonin receptors – a potential target for the treatment of Alzheimer's disease. *Biokhimiya = Biochemistry*. 2023;88(12):2399-2421. DOI 10.31857/S032097252312 0059 (in Russian)

Gossye H., Van Mossevelde S., Sieben A., Bjerke M., Hendrickx Van de Craen E., van der Zee J., De Deyn P.P., De Bleecker J., Versijpt J., van den Ende J., Deryck O., Bourgeois P., Bier J.C., Goethals M., Vandenberghe R., Engelborghs S., Van Broeckhoven C. Patients carrying the mutation p.R406W in *MAPT* present with non-conforming phenotypic spectrum. *Brain*. 2023;146(4):1624-1636. DOI 10.1093/brain/awac362

Grimm D., Kay M.A., Kleinschmidt J.A. Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. *Mol. Ther*. 2003;7(6):839-850. DOI 10.1016/s1525-0016(03)00095-9

Grundke-Iqbal I., Iqbal K., Tung Y.C., Quinlan M., Wisniewski H.M., Binder L.I. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc. Natl. Acad. Sci. USA*. 1986;83(13):4913-4917. DOI 10.1073/pnas.83.13.4913

Hock C., Heese K., Hulette C., Rosenberg C., Otten U. Region-specific neurotrophin imbalances in Alzheimer disease: decreased levels of brain-derived neurotrophic factor and increased levels of nerve growth factor in hippocampus and cortical areas. *Arch. Neurol*. 2000;57(6):846-851. DOI 10.1001/archneur.57.6.846

Huang G.B., Zhao T., Li C.R., Sui Z.Y., Kang N.I., Han E.H., Chung Y.C. Choline acetyltransferase expression in rat prefrontal cortex and hippocampus after acute and chronic exposure to amisulpride, haloperidol, and risperidone. *Neurosci. Lett*. 2012;528(2):131-136. DOI 10.1016/j.neulet.2012.09.024

Huey E.D., Putnam K.T., Grafman J. A systematic review of neurotransmitter deficits and treatments in frontotemporal dementia. *Neurology*. 2006;66(1):17-22. DOI 10.1212/01.wnl.0000191304.55196.4d

Hutton M., Lendon C.L., Rizzu P., Baker M., Froelich S., Houlden H., Pickering-Brown S. … Oostra B.A., Hardy J., Goate A., van Swieten J., Mann D., Lynch T., Heutink P. Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*. 1998;393(6686):702-705. DOI 10.1038/31508

Jahreis K., Brüge A., Borsdorf S., Müller F.E., Sun W., Jia S., Kang D.M., Boesen N., Shin S., Lim S., Koroleva A., Satala G., Bojarski A.J., Rakuša E., Fink A., Doblhammer-Reiter G., Kim Y.K., Dityatev A., Ponimaskin E., Labus J. Amisulpride as a potential disease-modifying drug in the treatment of tauopathies. *Alzheimers Dement*. 2023;19(12):5482-5497. DOI 10.1002/alz.13090

Jóźwiak-Bębenista M., Jasinska-Stroschein M., Kowalczyk E. The differential effects of neuroleptic drugs and PACAP on the expression of BDNF mRNA and protein in a human glioblastoma cell line. *Acta Neurobiol. Exp*. 2017;77(3):205-213

Khotskin N.V., Plyusnina A.V., Kulikova E.A., Bazhenova E.Y., Fursenko D.V., Sorokin I.E., Kolotygin I., Mormede P., Terenina E.E., Shevelev O.B., Kulikov A.V. On association of the lethal yellow (A^Y) mutation in the *agouti* gene with the alterations in mouse brain and behavior. *Behav. Brain Res*. 2019;359:446-456. DOI 10.1016/j.bbr.2018.11.013

Kobe F., Guseva D., Jensen T.P., Wirth A., Renner U., Hess D., Muller M., Medrihan L., Zhang W., Zhang M., Braun K., Westerholz S., Herzog A., Radyushkin K., El-Kordi A., Ehrenreich H., Richter D.W., Rusakov D.A., Ponimaskin E. 5-HT_7R/G_{12} signaling regulates neuronal morphology and function in an age-dependent manner. *J. Neurosci*. 2012;32(9):2915-2930. DOI 10.1523/JNEUROSCI.2765-11.2012

Kondaurova E.M., Bazovkina D.V., Naumenko V.S. 5-HT_{1A}/5-HT_7 receptor interplay: Chronic activation of 5-HT_7 receptors decreases the functional activity of 5-HT_{1A} receptor and its content in the mouse brain. *Molecular Biology*. 2017;51(1):136-142. DOI 10.1134/S0026893316060108

Kondaurova E.M., Plyusnina A.V., Ilchibaeva T.V., Eremin D.V., Rodnyy A.Y., Grygoreva Y.D., Naumenko V.S. Effects of a *Cc2d1a/Freud-1 Knockdown* in the hippocampus on behavior, the serotonin system, and BDNF. *Int. J. Mol. Sci*. 2021;22(24):13319. DOI 10.3390/ijms222413319

Kulikov A.V., Naumenko V.S., Voronova I.P., Tikhonova M.A., Popova N.K. Quantitative RT-PCR assay of 5-HT_{1A} and 5-HT_{2A} serotonin receptor mRNAs using genomic DNA as an external standard. *J. Neurosci. Methods*. 2005;141(1):97-101. DOI 10.1016/j.jneumeth.2004.06.005

Kulikov A.V., Tikhonova M.A., Kulikov V.A. Automated measurement of spatial preference in the open field test with transmitted lighting. *J. Neurosci. Methods*. 2008;170(2):345-351. DOI 10.1016/j.jneumeth.2008.01.024

Labus J., Röhrs K.F., Ackmann J., Varbanov H., Müller F.E., Jia S., Jahreis K., Vollbrecht A.L., Butzlaff M., Schill Y., Guseva D., Böhm K., Kaushik R., Bijata M., Marin P., Chaumont-Dubel S., Zeug A., Dityatev A., Ponimaskin E. Amelioration of Tau pathology and memory deficits by targeting 5-HT7 receptor. *Prog. Neurobiol*. 2021;197:101900. DOI 10.1016/j.pneurobio.2020.101900

Minaya M.A., Mahali S., Iyer A.K., Eteleeb A.M., Martinez R., Huang G., Budde J., Temple S., Nana A.L., Seeley W.W., Spina S., Grinberg L.T., Harari O., Karch C.M. Conserved gene signatures shared among *MAPT* mutations reveal defects in calcium signaling. *Front. Mol. Biosci*. 2023;10:1051494. DOI 10.3389/fmolb.2023.1051494

Molobekova C.A., Kondaurova E.M., Ilchibaeva T.V., Rodnyy A.Y., Stefanova N.A., Kolosova N.G., Naumenko V.S. Amisulpride decreases tau protein hyperphosphorylation in the brain of OXYS rats. *Curr. Alzheimer Res*. 2023;20(7):496-505. DOI 10.2174/1567205020666230828144651

Murley A.G., Rowe J.B. Neurotransmitter deficits from frontotemporal lobar degeneration. *Brain*. 2018;141(5):1263-1285. DOI 10.1093/brain/awx327

Naumenko V.S., Kulikov A.V. Quantitative assay of 5-HT_{1A} receptor gene expression in the brain. *Molecular Biology*. 2006;40(1):30-36. DOI 10.1134/S0026893306010067

Naumenko V.S., Osipova D.V., Kostina E.V., Kulikov A.V. Utilization of a two-standard system in real-time PCR for quantification of gene expression in the brain. *J. Neurosci. Methods*. 2008;170(2):197-203. DOI 10.1016/j.jneumeth.2008.01.008

Park S.W., Seo M.K., Cho H.Y., Lee J.G., Lee B.J., Seol W., Kim Y.H. Differential effects of amisulpride and haloperidol on dopamine D2 receptor-mediated signaling in SH-SY5Y cells. *Neuropharmacology*. 2011;61(4):761-769. DOI 10.1016/j.neuropharm.2011.05.022

Perez M., Lim F., Arrasate M., Avila J. The FTDP-17-linked mutation R406W abolishes the interaction of phosphorylated tau with microtubules. *J. Neurochem*. 2000;74(6):2583-2589. DOI 10.1046/j.1471-4159.2000.0742583.x

Pisani A., Paciello F., Del Vecchio V., Malesci R., De Corso E., Cantone E., Fetoni A.R. The role of BDNF as a biomarker in cognitive and sensory neurodegeneration. *J. Pers. Med*. 2023;13(4):652. DOI 10.3390/jpm13040652

Popova N.K., Naumenko V.S. Neuronal and behavioral plasticity: the role of serotonin and BDNF systems tandem. *Expert Opin. Ther. Targets*. 2019;23(3):227-239. DOI 10.1080/14728222.2019.1572747

Renner U., Zeug A., Woehler A., Niebert M., Dityatev A., Dityateva G., Gorinski N., Guseva D., Abdel-Galil D., Frohlich M., Doring F., Wischmeyer E., Richter D.W., Neher E., Ponimaskin E.G.

Heterodimerization of serotonin receptors 5-HT_{1A} and 5-HT_7 differentially regulates receptor signalling and trafficking. *J. Cell Sci*. 2012;125(Pt. 10):2486-2499. DOI 10.1242/jcs.101337

Rizos E.N., Papadopoulou A., Laskos E., Michalopoulou P.G., Kastania A., Vasilopoulos D., Katsafouros K., Lykouras L. Reduced serum BDNF levels in patients with chronic schizophrenic disorder in relapse, who were treated with typical or atypical antipsychotics. *World J. Biol. Psychiatry*. 2010;11(2-2):251-255. DOI 10.3109/15622970802182733

Rodnyy A.Y., Kondaurova E.M., Bazovkina D.V., Kulikova E.A., Ilchibaeva T.V., Kovetskaya A.I., Baraboshkina I.A., Bazhenova E.Y., Popova N.K., Naumenko V.S. Serotonin 5-HT7 receptor overexpression in the raphe nuclei area produces antidepressive effect and affects brain serotonin system in male mice. *J. Neurosci. Res*. 2022; 100(7):1506-1523. DOI 10.1002/jnr.25055

Rovelet-Lecrux A., Lecourtois M., Thomas-Anterion C., Le Ber I., Brice A., Frebourg T., Hannequin D., Campion D. Partial deletion of the *MAPT* gene: a novel mechanism of FTDP-17. *Hum. Mutat.* 2009;30(4):E591-E602. DOI 10.1002/humu.20979

Samarajeewa A., Goldemann L., Vasefi M.S., Ahmed N., Gondora N., Khanderia C., Mielke J.G., Beazely M.A. 5-HT7 receptor activation promotes an increase in TrkB receptor expression and phosphorylation. *Front. Behav. Neurosci*. 2014;8:391. DOI 10.3389/fnbeh.2014.00391

Shen L.L., Li W.W., Xu Y.L., Gao S.H., Xu M.Y., Bu X.L., Liu Y.H., Wang J., Zhu J., Zeng F., Yao X.Q., Gao C.Y., Xu Z.Q., Zhou X.F., Wang Y.J. Neurotrophin receptor p75 mediates amyloid β-induced tau pathology. *Neurobiol. Dis*. 2019;132:104567. DOI 10.1016/j.nbd.2019.104567

Solas M., Van Dam D., Janssens J., Ocariz U., Vermeiren Y., De Deyn P.P., Ramirez M.J. 5-HT_7 receptors in Alzheimer's disease. *Neurochem. Int*. 2021;150:105185. DOI 10.1016/j.neuint.2021.105185

Song J.H., Yu J.T., Tan L. Brain-derived neurotrophic factor in Alzheimer's disease: risk, mechanisms, and therapy. *Mol. Neurobiol*. 2015;52(3):1477-1493. DOI 10.1007/s12035-014-8958-4

Stefanova N.A., Muraleva N.A., Korbolina E.E., Kiseleva E., Maksimova K.Y., Kolosova N.G. Amyloid accumulation is a late event in sporadic Alzheimer's disease-like pathology in nontransgenic rats. *Oncotarget*. 2015;6(3):1396-1413. DOI 10.18632/oncotarget.2751

Strang K.H., Golde T.E., Giasson B.I. *MAPT* mutations, tauopathy, and mechanisms of neurodegeneration. *Lab. Invest*. 2019;99(7):912-928. DOI 10.1038/s41374-019-0197-x

Xia Y., Sorrentino Z.A., Kim J.D., Strang K.H., Riffe C.J., Giasson B.I. Impaired tau-microtubule interactions are prevalent among pathogenic tau variants arising from missense mutations. *J. Biol. Chem*. 2019;294(48):18488-18503. DOI 10.1074/jbc.RA119.010178

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

Received February 9, 2024. Revised March 26, 2024. Accepted March 28, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):407-415

DOI 10.18699/vjgb-24-46

Dopamine receptors and key elements of the neurotrophins (BDNF, CDNF) expression patterns during critical periods of ontogenesis in the brain structures of mice with autism-like behavior (BTBR) or its absence (C57BL/6J)

P.D. Pravikova, M.A. Arssan, E.A. Zalivina, E.M. Kondaurova, E.A. Kulikova, I.I. Belokopytova, V.S. Naumenko

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
PollyPravi@yandex.ru

Abstract. Analysis of the mechanisms underlying autism spectrum disorder (ASD) is an urgent task due to the ever-increasing prevalence of this condition. The study of critical periods of neuroontogenesis is of interest, since the manifestation of ASD is often associated with prenatal disorders of the brain development. One of the currently promising hypotheses postulates a connection between the pathogenesis of ASD and the dysfunction of neurotransmitters and neurotrophins. In this study, we investigated the expression of key dopamine receptors (*Drd1*, *Drd2*), brain-derived neurotrophic factor (*Bdnf*), its receptors (*Ntrkb2*, *Ngfr*) and the transcription factor *Creb1* that mediates BDNF action, as well as cerebral dopamine neurotrophic factor (*Cdnf*) during the critical periods of embryogenesis (e14 and e18) and postnatal development (p14, p28, p60) in the hippocampus and frontal cortex of BTBR mice with autism-like behavior compared to the neurotypical C57BL/6J strain. In BTBR embryos, on the 14th day of prenatal development, an increase in the expression of the *Ngfr* gene encoding the $p75^{NTR}$ receptor, which may lead to the activation of apoptosis, was found in the hippocampus and frontal cortex. A decrease in the expression of *Cdnf*, *Bdnf* and its receptor *Ntrkb2*, as well as dopamine receptors (*Drd1*, *Drd2*) was detected in BTBR mice in the postnatal period of ontogenesis mainly in the frontal cortex, while in the hippocampus of mature mice (p60), only a decrease in the *Drd2* mRNA level was revealed. The obtained results suggest that the decrease in the expression levels of CDNF, BDNF-TrkB and dopamine receptors in the frontal cortex in the postnatal period can lead to significant changes in both the morphology of neurons and dopamine neurotransmission in cortical brain structures. At the same time, the increase in $p75^{NTR}$ receptor gene expression observed on the 14th day of embryogenesis, crucial for hippocampus and frontal cortex development, may have direct relevance to the manifestation of early autism.

Key words: autism; BTBR and C57BL/6J mice; BDNF; CDNF; dopamine receptors; ontogenesis; hippocampus; frontal cortex.

For citation: Pravikova P.D., Arssan M.A., Zalivina E.A., Kondaurova E.M., Kulikova E.A., Belokopytova I.I., Naumenko V.S. Dopamine receptors and key elements of the neurotrophins (BDNF, CDNF) expression patterns during critical periods of ontogenesis in the brain structures of mice with autism-like behavior (BTBR) or its absence (C57BL/6J). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):407-415. DOI 10.18699/vjgb-24-46

Funding. The work was supported by the Russian Science Foundation (grant No. 22-15-00028).

Acknowledgements. The cost of animal housing was compensated by the basic research project No. FWNR-2022-0023.

Паттерны экспрессии рецепторов дофамина и основных элементов нейротрофических (BDNF, CDNF) систем в критические периоды онтогенеза в структурах мозга мышей с аутизм-подобным поведением (BTBR) или его отсутствием (C57BL/6J)

П.Д. Правикова, М.А. Арссан, Е.А. Заливина, Е.М. Кондаурова, Е.А. Куликова, И.И. Белокопытова, В.С. Науменко

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
PollyPravi@yandex.ru

Аннотация. Анализ механизмов расстройства аутистического спектра (РАС) является актуальной задачей в связи с широкой и постоянно растущей распространенностью этого состояния. Исследование критических периодов нейроонтогенеза представляет интерес, поскольку манифестацию РАС нередко связывают с внут-

риутробными нарушениями развития головного мозга. Одна из перспективных на сегодняшний день гипотез постулирует связь патогенеза РАС с дисфункцией нейротрансмиттерных и нейротрофических систем. В настоящей работе исследована экспрессия генов ключевых рецепторов дофамина (*Drd1*, *Drd2*), нейротрофического фактора мозга (*Bdnf*), его рецепторов (*Ntrkb2*, *Ngfr*) и опосредующего действие BDNF транскрипционного фактора *Creb1*, а также дофаминового нейротрофического фактора (*Cdnf*) в периоды эмбриогенеза (e14 и e18) и постнатального развития (p14, p28, p60) в гиппокампе и фронтальной коре мышей BTBR с аутистизм-подобным поведением по сравнению с нейротипичной линией C57BL/6 J. У эмбрионов BTBR на 14-й день пренатального развития в гиппокампе и во фронтальной коре установлено увеличение экспрессии гена *Ngfr*, кодирующего рецептор $p75^{NTR}$, трансдукция сигнала которого в эмбриогенезе приводит к активации апоптоза. Снижение экспрессии генов *Cdnf*, *Bdnf* и его рецептора *Ntrkb2*, а также дофаминовых рецепторов (*Drd1*, *Drd2*) у мышей BTBR обнаружено в постнатальный период преимущественно во фронтальной коре, при этом в гиппокампе у половозрелых особей (p60) зафиксировано падение уровня лишь мРНК *Drd2*. Полученные результаты позволяют предположить, что снижение в постнатальном периоде экспрессии генов *Cdnf*, *Bdnf* и *Ntrkb2*, а также дофаминовых рецепторов во фронтальной коре может приводить к существенным изменениям, характерным для РАС, как морфологии нейронов, так и дофаминовой нейротрансмиссии в корковых структурах мозга. Вместе с тем установленный рост экспрессии $p75^{NTR}$ в критический для развития гиппокампа и фронтальной коры 14-й день эмбриогенеза, возможно, является ключевым для формирования раннего аутизма.

Ключевые слова: аутизм; мыши BTBR и C57BL/6 J; BDNF; CDNF; рецепторы дофамина; онтогенез; гиппокамп; фронтальная кора.

Introduction

First described back in 1943 (Kanner, 1943), autism (now autism spectrum disorder, ASD) is defined as a group of conditions caused by disorders in prenatal and early postnatal neuroontogenesis that persist throughout a person's life. ASD is characterized by a decline in the ability to initiate and maintain social interactions and communication, as well as a series of restricted and repetitive inflexible behavior patterns. Data from the Centers for Disease Control and Prevention show a steady increase in the number of children diagnosed with ASD: in 2023, 1 in 36 children was diagnosed, while in 2010, the prevalence of ASD was 1 %. However, according to data from both the WHO and the Ministry of Health of the Russian Federation (Letter of the Ministry of Health No. 15-3/10/1-2140 dated 05/08/2013), the prevalence of ASD was about 1 % of the child population.

Currently, there is no unified concept of the pathogenesis of ASD, however, most hypotheses associate the development of this condition with early neurodevelopmental disorders leading to disturbances in mental functions (Hashem et al., 2020). In this regard, special attention when studying the ASD mechanisms is paid to early prenatal brain development (Courchesne et al., 2020). Investigations of induced stem cells obtained from people with ASD have supported the prenatal origin of this disorder (Adhya et al., 2021). A high rate of cell proliferation and a decrease in the degree of differentiation and maturation of GABAergic interneurons have been described (Mariani et al., 2015). Abnormal proliferation and excess prenatal neurogenesis in individuals with ASD appear to explain increases in both cortical neuron numbers (Courchesne et al., 2011) and overall brain mass (Sacco et al., 2015). In addition, the peak expression of most putative risk genes for ASD occurs during the prenatal period (Satterstrom et al., 2020) and is found in a number of brain regions, including cortical areas and the hippocampus (Krishnan et al., 2016; Courchesne et al., 2019). One of the morphological features of ASD is a decrease in volume and sometimes complete agenesis of the corpus callosum (Frazier, Hardan, 2009). At the same time, changes in the volume of the corpus callosum are obviously a consequence of prenatal developmental disorders, since in mammals its formation is completed at the last stage of embryogenesis (Richards et al., 2004).

The neurotransmitter dopamine (DA) is involved in the modulation of learning, reward, and emotional control, which are known to be impaired in autism (Hashem et al., 2020). In people with ASD, a number of polymorphisms are observed in the genes encoding the DA-transporter (DAT) (Dicarlo et al., 2019), DA-metabolic enzymes (Yoo et al., 2013) and DA-receptors (Hettinger et al., 2008; Staal et al., 2015). Moreover, DA neurons derived from pluripotent stem cells from patients with autism are characterized by morphological changes and abnormalities in Ca^{2+}-signal transduction (Nguyen et al., 2018). Based on these data, a connection between the pathogenesis of ASD and dysfunction of the brain DA system has been suggested (Pavăl, 2017).

Neurotrophic factors attract special attention because they play a key role in the regulation of neuronal growth and development, as well as in the control of neuroplasticity (Popova, Naumenko, 2019). Brain-derived neurotrophic factor (BDNF) is one of the most studied neurotrophins that controls synaptogenesis, triggers long-term potentiation, and is involved in memory formation (Castrén, Antila, 2017). An association has been shown between a decreased BDNF blood level in newborns and an increased risk of ASD development (Liu et al., 2021). At the same time, post-mortem studies of the brain of children with ASD have established an increase in the number of prefrontal neurons, which may be a consequence of impaired activity of the BDNF-signal transduction and lead to an excess of axonal connections (Anghelescu, Dettling, 2012). Cerebral dopamine neurotrophic factor (CDNF), first described in 2007, is a non-conventional growth factor and is predominantly localized in the striatum, substantia nigra, hippocampus, cortex and cerebellum (Lindholm et al., 2007, 2008). CDNF is currently being tested in clinical trials as a treatment for Parkinson's disease (PD) (Lindholm, Saarma, 2022), as it is able to slow down the degeneration of DA

neurons (Voutilainen et al., 2011). However, the relationship between neurotrophins and the DA system at different stages of ontogenesis in the context of the development of autism has not been investigated yet.

Based on the stated above, the aim of the current study was to identify the role of the DA system and neurotrophic factors in the development of autism by analyzing the expression patterns of dopamine receptors (*Drd1*, *Drd2*) and *Cdnf*, as well as *Bdnf*, its receptors (*Ntrkb2*, *Ngfr*), and transcription factor *Creb1* mediating BDNF effects in the brain structures of BTBR mice, which are known to be a model of autism, in comparison with neurotypical C57Bl/6 J mice at different periods of ontogenesis.

Materials and methods

Experimental animals. The BTBR inbred strain is a widely accepted idiopathic model of autism (Crawley, 2023) as it is characterized by social deficits as well as repetitive behavior (Bolivar et al., 2007; McFarlane et al., 2008). Experiments were conducted on male mice of pathogen-free (SPF) inbred strains BTBR T+tf/J (BTBR) and C57Bl/6 J at the Center for Genetic Resources of Laboratory Animals; Institute of Cytology and Genetics, supported by the Ministry of Science and Higher Education of the Russian Federation (unique identification number: RFMEFI62119X0023). The mice were housed under standard laboratory conditions with a 14-h light cycle, constant humidity (60 %), temperature (23 °C) and with access to balanced food and water *ad libitum*. All procedures performed with the involvement of laboratory animals were approved by the ethical standards of the Committee on Biological Ethics at the Institute of Cytology and Genetics SB RAS and complied with the ethical standards approved by the legal acts of the Russian Federation (Order of the Ministry of Health of the Russian Federation No. 267 of June 19, 2003), as well as protocols on the treatment of laboratory animals.

Experimental design. In autism, the memory consolidation regulated by the hippocampus is impaired, as well as the disturbances observed in executive function and social behavior, for the implementation of which the frontal cortex is mainly responsible. In addition, post-mortem studies of people with ASD revealed a reduced cell size and increased cell density in both the hippocampus and the frontal cortex (Kemper, Bauman, 1998; Courchesne et al., 2011). Based on this, the study of these brain structures in the context of the mechanisms of ASD development was of particular interest.

One of the most variable structures in ASD pathogenesis is the hippocampus, the formation of which begins on the 14th day of embryogenesis (Mangale et al., 2008), while on the 18th day of prenatal development it is already formed (Loones et al., 2000). In addition, on the 17th day of embryogenesis, the corpus callosum is completely formed (Richards et al., 2004); however, the agenesis of the corpus callosum is demonstrated both in BTBR mice (Bohlen et al., 2012) and, often, in people with ASD (Frazier, Hardan, 2009). One of the ASD criteria is hyper- or hyporeactivity to sensory input due to increased or decreased sensitivity to stimuli (DSM-5, ICD-11). Since rodents' eyes open at 12–13 days and their sensory perception becomes full (Rochefort et al., 2009), it was interesting to study a group of mice at 14 days of age. The juvenile period is an important postnatal stage of development in the study of autism-like behavior, as BTBR mice exhibit low levels of social interaction as early as 28 days after birth (McFarlane et al., 2008). Thus, the following periods of ontogenesis were selected: 14th or 18th day of embryogenesis, as well as 14th, 28th and 60th (reaching maturity) day of postnatal development.

Male mice (p14, p28, p60), as well as embryos of the BTBR and C57Bl/6 J strains on the 14th or 18th day of prenatal development, were removed from the experiment by decapitation, and their hippocampi and prefrontal cortices were removed on ice, frozen in liquid nitrogen and stored at –80 °C. For groups e14 and e18, a partial tail biopsy was also performed for subsequent genotyping of the Y chromosome (*Sry*). The sex of the p14 mice and older was determined by primary sexual characteristics at autopsy. The number of individuals in the experimental group of a certain ontogenetic day (e14, e18, p14, p28, p60) was 10 for each strain.

Obtaining embryos and determining their sex. To obtain embryos *in vivo*, sexually mature female mice of the BTBR and C57Bl/6 J strains, in a state of estrus, which was determined by analyzing vaginal smears, were mated overnight with males of the corresponding strains. The day of detection of sperm in the vaginal smear was considered the first day of pregnancy. For genotyping of the Y chromosome (*Sry*), genomic DNA was isolated from embryonic tail tissue by placing it in a lysis solution containing protease K for two hours at 50 °C, followed by extraction in saturated saline according to a previously described protocol (Aljanabi, Martinez, 1997). DNA samples were amplified with primers (see the Table) (Wambach et al., 2014), and PCR products were separated by electrophoresis on a 2 % agarose gel and visualized by ethidium bromide staining. Embryos with the presence of the Y chromosome were used in this work.

RT-qPCR. *Total RNA isolation.* Total RNA was isolated with TRIzol Reagent (Life Technologies, USA) as recommended by the manufacturer. Isolated RNA was diluted with water to the concentration of 0.125 μg/μl and stored at –70 °C. The presence of genomic DNA in the RNA preparations was determined as described in (Kulikov et al., 2005; Naumenko, Kulikov, 2006; Naumenko et al., 2008).

Reverse transcription and qPCR. Reverse transcription and real-time PCR were carried out according to the protocol previously described in detail (Kulikov et al., 2005; Naumenko, Kulikov, 2006; Naumenko et al., 2008). Two types of standards were used: external and internal. An internal standard (*Polr2a* mRNA) was used to monitor reverse transcription and as a basis for calculating the mRNA levels of the studied genes. Mouse DNA with a known concentration (external standard) was used as a PCR control and to determine the copy number of the assayed genes and *Polr2a* in the samples. Primers used for PCR amplification (see the Table) were developed based on the gene sequences deposited in the Ensembl database and were synthesized by BIOSSET (Russia).

Statistical analysis. The obtained results are presented as mean ± standard error of mean (m ± SEM). For pairwise comparison of means between mouse strains at a certain developmental day, Student's *t*-test for independent samples was used. The differences were considered significant at $p < 0.05$. Normal distribution of the data was verified by the Kolmogorov–Smirnov and Shapiro–Wilk tests. The outliers were identified and excluded by Dixon's Q test.

Sequences of the gene-specific oligonucleotide primers and their characteristics

Gene	Primer sequence	$T_{annealing}$, °C	PCR product length, bp
rPol2	F 5′-tgtgacaactccatacaatgc-3′ R 5′-ctctcttagtgaatttgcgtact-3′	60	194
Cdnf	F 5′- cggtggacctgtggaagatg-3′ R 5′- acatatttgggggccagctc-3′	60	130
Bdnf	F 5′-tagcaaaaagagaattggctg-3′ R 5′-tttcaggtcatggatatgtcc-3′	59	255
Ntrk2	F 5′-cattcactgtgagaggcaacc-3′ R 5′-atcagggtgtagtctccgttatt-3′	63	175
Ngfr	F 5′-acaacacccagcacccagga-3′ R 5′-cacaaccacagcagccaaga-3′	62	171
Creb1	F 5′-gctggctaacaatggtacggat-3′ R 5′-tggttgctgggcactagaat-3′	64	140
Drd1	F 5′-ggaaaccctgtcgaatgctctc-3′ R 5′-ccagccaaaccacacaaatacatcg-3′	59	222
Drd2	F 5′-tccgccacttcttgacatacattg-3′ R 5′-cccatccacagcctcctctaag-3′	64	203
Sry	F 5′-ttgtctagagagcatggagggccatgtcaa -3′ R 5′-ccactcctctgtgacactttagccctccga-3′	64	268

Results

In BTBR mice, a significant increase in the *Drd1* mRNA level was revealed in the hippocampus only on the 14th day of postnatal development compared to the C57Bl/6 J strain. A decrease in the expression of this gene was detected in the frontal cortex of BTBR mice on the 28th day of the postnatal period (Fig. 1*a*). For sexually mature individuals (p60), no interstrain differences in the level of *Drd1* expression were observed either in the hippocampus or in the frontal cortex. At the same time, *Drd1* gene expression in p60 in the hippocampus of both mouse strains decreased to embryonic values (e18) compared to other stages of postnatal development (p14, p28).

An increase in the *Drd2* mRNA level revealed in the frontal cortex of BTBR embryos on the 14th day of prenatal development was leveled out on the 18th day of embryogenesis (Fig. 1*b*). At the same time, in BTBR mice, an increase in the *Drd2* expression was observed in the frontal cortex on the p14 day, while no significant interstrain differences were found upon reaching sexual maturity (p60). Meanwhile, in the hippocampus, interstrain differences in the *Drd2* mRNA level were established only in mature mice (p60): in the BTBR strain, a decrease in *Drd2* expression was revealed in comparison with the C57Bl/6 J mice, which showed a dramatic increase in the *Drd2* mRNA level after the 28th day of the postnatal development.

In the frontal cortex of the BTBR mice, an increase in the *Creb1* mRNA level was revealed on the 14th day of embryogenesis; in contrast, at the age of 60 days in BTBR mice, a decrease in its expression was found (Fig. 2). This dynamics in the *Creb1* expression in the BTBR strain is consistent

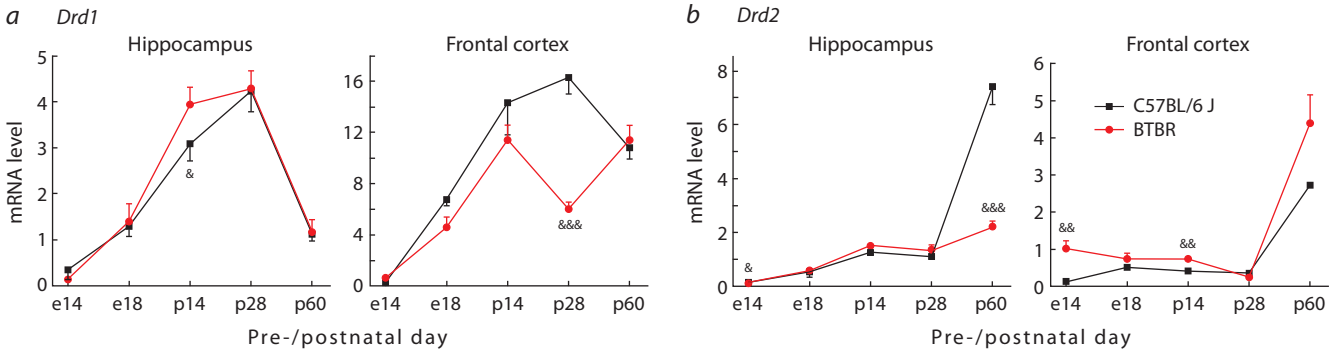


Fig. 1. *Drd1* (*a*) and *Drd2* (*b*) mRNA levels in the hippocampus and frontal cortex of BTBR and C57Bl/6 J mice at different periods of pre- and postnatal development.

Gene expression is presented as the number of cDNA copies per 100 copies of *Polr2a* cDNA. $n = 8–10$.

Significant difference: $^{\&}p < 0.05$; $^{\&\&}p < 0.01$; $^{\&\&\&}p < 0.001$ – interstrain comparison within a certain day of ontogenesis.

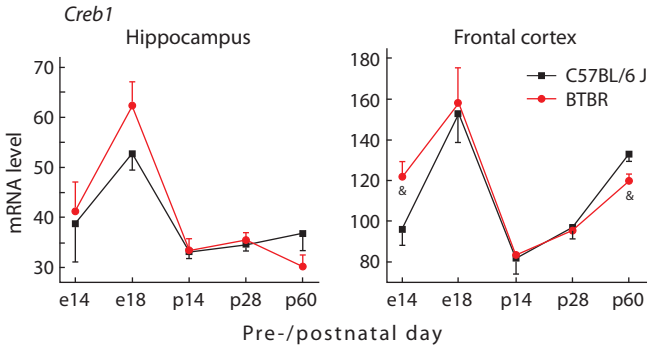


Fig. 2. mRNA level of *Creb1*, encoding the transcription factor CRE-binding protein, in the hippocampus and frontal cortex of BTBR and C57Bl/6 J mice at different periods of pre- and postnatal development.

Gene expression is presented as the number of cDNA copies per 100 copies of *Polr2a* cDNA. $n = 8–10$.

Significant difference: $^{\&}p < 0.05$ – interstrain comparison within a certain day of ontogenesis.

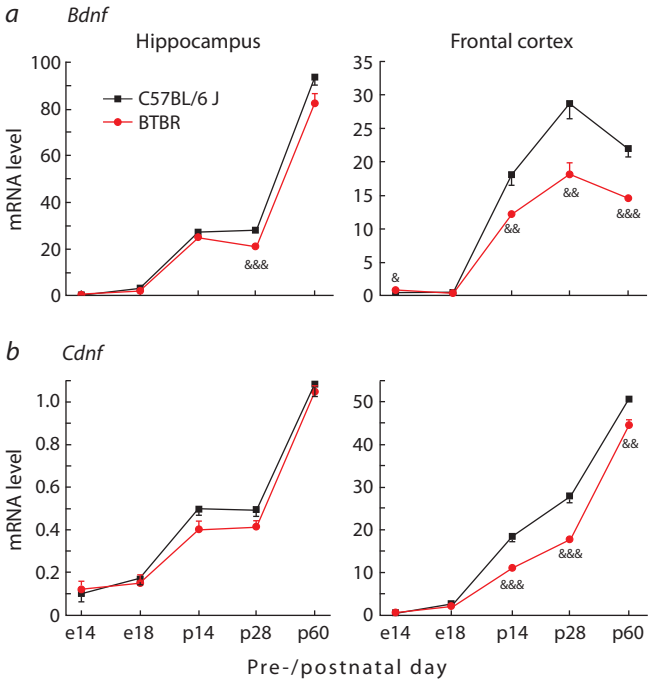


Fig. 3. *Bdnf* (*a*) and *Cdnf* (*b*) mRNA levels in the hippocampus and frontal cortex of BTBR and C57Bl/6 J mice at different periods of pre- and postnatal development.

Gene expression is presented as the number of cDNA copies per 100 copies of *Polr2a* cDNA. $n = 8–10$.

Significant difference: $^{\&\&}p < 0.01$; $^{\&\&\&}p < 0.001$ – interstrain comparison within a certain day of ontogenesis.

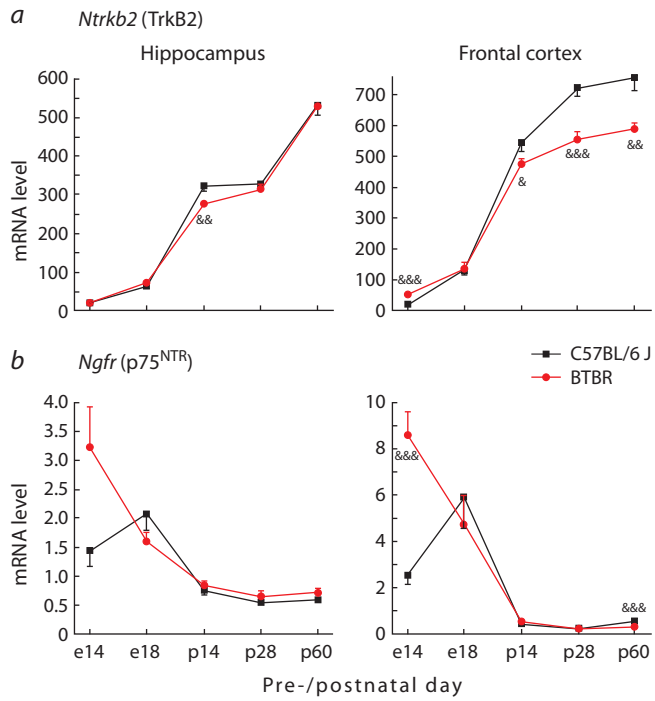


Fig. 4. *Ntrkb2* (*a*) and *Ngfr* (*b*) mRNA levels in the hippocampus and frontal cortex of BTBR and C57Bl/6 J mice at different periods of pre- and postnatal development.

Gene expression is presented as the number of cDNA copies per 100 copies of *Polr2a* cDNA. $n = 8–10$.

Significant difference: $^{\&}p < 0.05$; $^{\&\&}p < 0.01$; $^{\&\&\&}p < 0.001$ – interstrain comparison within a certain day of ontogenesis.

with changes in the cortical *Bdnf* mRNA level: on the 14th day of embryogenesis, an increase in *Bdnf* mRNA level was detected, while on the 60th day of postnatal development, a decrease in its expression was shown (Fig. 3*a*). At the same time, no interstrain differences in the *Creb1* expression were detected in the hippocampus during the investigated periods of ontogenesis. Along with this, in the hippocampus and frontal cortex, the peak of *Creb1* expression occurred on the 18th day of embryogenesis, while in the frontal cortex, the minimal value of the *Creb1* mRNA content was observed on the 14th day of postnatal development (Fig. 2).

In BTBR mice, in the frontal cortex during the studied periods of postnatal development (p14, p28, p60), a decrease in the *Bdnf* coding exon mRNA level was detected, while in the hippocampus, interstrain difference in its expression was revealed only on the 28th day of postnatal development (Fig. 3*a*). The detected changes in the *Bdnf* expression in the frontal cortex during the postnatal development of BTBR mice are consistent with the decrease in the *Ntrkb2* mRNA level (Fig. 4*a*), encoding the main BDNF receptor – tyrosine kinase receptor B (TrkB).

Similarly to the changes in the *Bdnf* mRNA level in the frontal cortex, a decrease in the *Cdnf* mRNA level on days 14, 28, and 60 of postnatal development was found in BTBR mice (Fig. 3*b*). At the same time, in the hippocampus, no interstrain differences in the *Cdnf* expression were observed (Fig. 3*b*), while, regardless of the mouse strain, a dramatic increase in the *Cdnf*, *Bdnf* and *Ntrkb2* expression was revealed after day p28.

Analysis of the *Ngfr* gene (encoding the $p75^{NTR}$ receptor mediating proBDNF action) expression dynamics showed its increase on the 14th day of embryogenesis in BTBR mice both in the hippocampus and frontal cortex (Fig. 4*b*). At the same time, in the frontal cortex on the 60th day of postnatal development, upon reaching puberty, the *Ngfr* mRNA level significantly decreased in BTBR mice. In BTBR mice, the peak expression of the *Ngfr* gene in the investigated brain

structures was observed on the 14th day of prenatal development, while in C57Bl/6 J mice, it was revealed on the 18th day of embryogenesis (Fig. 4*b*), which is in a good agreement with the increase of *Creb1* expression (Fig. 2).

Discussion

People suffering from ASD show changes in neuronal morphology (Minshew, Williams, 2007). In this regard, the analysis of neurotrophins is a promising task for studying the mechanisms of autism. It is known that the BDNF protein is synthesized as a precursor (proBDNF), which is then cleaved to the mature form (mBDNF) (Lessmann et al., 2003). mBDNF is responsible for increasing synaptic plasticity, whereas proBDNF, on the contrary, mediates its decrease (Koshimizu et al., 2009). proBDNF, the expression of which is elevated during the prenatal period (Yang et al., 2009), is a key neurotrophic factor regulating the development of the central nervous system through its influence on neurogenesis (Koshimizu et al., 2009). In addition, increased proBDNF levels were found in post-mortem sections of the fusiform gyrus of people with autism (Garcia et al., 2012), which may be the cause of a decrease in both neuronal differentiation and dendritic spines formation (Teng et al., 2005). BDNF action is mediated by two receptors, namely tyrosine kinase B receptor (TrkB) and $p75^{NTR}$ receptor. $p75^{NTR}$ receptor was previously suggested as a biomarker of ASD, since its mRNA level in peripheral blood was found to be increased in people with autism (Segura et al., 2015).

Here we found that in BTBR mice on the 14th day of embryogenesis in the hippocampus and frontal cortex, when neurogenesis should reach its maximum (Finlay, Darlington, 1995; Chen et al., 2017), the *Ngfr* gene (encoding $p75^{NTR}$ receptor) mRNA level is increased, while by the 18th day of prenatal development, these differences are already eliminated. Although the $p75^{NTR}$ receptor signal transduction pathways are extremely diverse (Lu et al., 2005), proBDNF binding to $p75^{NTR}$ is known to stimulate cell death (Teng et al., 2005). It can be suggested that in BTBR mice, on the 14th day of prenatal development, apoptotic activity in the hippocampus and frontal cortex is increased, which, together with the impaired synaptogenesis and a decrease in the ability of neurons to form functional connections, leads to behavioral deficits observed in autism (Wei et al., 2014). On the other hand, increased apoptosis can lead to agenesis of the corpus callosum, which is observed in BTBR mice (Bohlen et al., 2012), and often in humans with autism (Frazier, Hardan, 2009). At the same time, in the neurotypical C57Bl/6 J strain, the highest *Ngfr* mRNA level in the hippocampus and frontal cortex is observed on the 18th day of embryogenesis, which then dramatically decreases together with the decline of neurogenesis in these cell populations (Finlay, Darlington, 1995; Chen et al., 2017). An increase in the $p75^{NTR}$ receptor expression detected in the frontal cortex of BTBR mice on e14 is accompanied by an increase in the mRNA levels of *Bdnf* coding exon, transcription factor *Creb1*, and also the *Ntrkb2* gene encoding the mBDNF receptor TrkB. Apparently, increased expression of the key elements of the BDNF-TrkB signaling pathway is a compensatory response to the putative increase of apoptotic activity mediated by $p75^{NTR}$ receptor, since BDNF, when bound to TrkB, triggers protein synthesis, growth and maturation of neurons (Fenner, 2012). At the same time, in the frontal cortex of BTBR mice, a decrease in both *Bdnf* and its receptor *Ntrkb2* (TrkB) mRNA levels was observed from day p14 to sexual maturity. That may indicate a decrease in neuroprotective and neurotrophic properties in cortical structures of BTBR mice and lead to autism-like behavior. This, in accordance with our previous data, showed that BDNF overexpression in the hippocampus reduces anxiety and stereotypic behavior in BTBR mice, which are among the diagnostic criteria for ASD (Ilchibaeva et al., 2023).

Earlier it was found that *Cc2d1a*/Freud-1 knockdown in the hippocampus of BTBR mice did not affect spatial memory and phosphorylation of the CREB transcription factor (Belokopytova et al., 2022), although such an effect was found in C57Bl/6 J mice (Kondaurova et al., 2021). Based on these data, we suggested that the functional activity of the CREB transcription factor (stimulating BDNF expression) is impaired in BTBR mice, which may be the cause of the disturbances in the BDNF signaling cascade identified in the current study.

ASD is often accompanied by dysfunction in certain neurotransmitter systems (Rodnyy et al., 2023). In particular, increased levels of serotonin (Pourhamzeh et al., 2022), as well as disturbances in the brain DA system functioning, manifesting in changes in both DA metabolism (Yoo et al., 2013) and signal transduction (Staal et al., 2015; DiCarlo et al., 2019), have been established in ASD. It is known that DA can affect cell proliferation and differentiation of telencephalon cells during embryonic development: blockade of DA-type 1 receptor (D_1R) leads to a decrease in the rate of cell division, while stimulation of DA-type 2 receptor (D_2R), on the contrary, promotes its activation (Popolo et al., 2004). On the other hand, transduction of the DA signal upon D_2R activation reduces the migration of GABAergic interneurons in the telencephalon (Crandall et al., 2007), which can lead to disruption of inhibitory processes in cortical areas observed in BTBR mice (Cellot et al., 2016) and often in people with autism (Enticott et al., 2013). These data are in a good agreement with our results identifying an increase in the *Drd2* mRNA level in the frontal cortex and hippocampus of BTBR mice on the 14th day of prenatal development, which may likely contribute to disturbances in the formation and functioning of these brain structures.

There is a large amount of data on the relationship between the BDNF and DA systems. For example, BDNF has been proposed as a promising agent in the treatment of Parkinson's disease, given its stimulatory effect on both DA release (Neal et al., 2003) and the overall trophic effect on DA neurons (Palasz et al., 2020) upon activation of *Bdnf* expression (Küppers, Beyer, 2001). In this regard, unidirectional changes in the expression of both D_2R and BDNF-TrkB detected on the 14th day of embryogenesis in BTBR mice are consistent with the concept of the relationship between the DA and BDNF systems. At the same time, the decrease in the *Drd2* mRNA level observed in the hippocampus of BTBR mice at the age of 60 days together with the absence of changes in the expression of genes encoding the studied neurotrophic factors may be a consequence of a disruption in the CREB-dependent signaling pathway (Belokopytova et al., 2022) and lead to deficits in memory and learning. At the same time, the hippocampal and cortical *Drd1* expression did not change in the BTBR strain

during the studied periods of embryogenesis. However, in the postnatal period, interstrain differences in the *Drd1* mRNA levels were detected both in the hippocampus and frontal cortex. The *Drd1* expression level in the hippocampus of BTBR mice was increased on day p14, while in the frontal cortex, on the contrary, its decrease was found already in the juvenile period (p28). Despite the absence of interstrain differences in the *Drd1* expression level upon reaching sexual maturity (p60), the observed changes in the *Drd1* mRNA levels during critical periods after birth may be among the reasons for learning and memory impairments observed in BTBR mice already in juvenile age (McFarlane et al., 2008). Since the detected decrease in the *Drd1* mRNA level in the frontal cortex of BTBR mice on the 28th day after birth was accompanied by a decrease in the *Bdnf* and *Ntrkb2* (TrkB) expression, the participation of the BDNF-TrkB signaling pathway in the regulation of D_1R expression was suggested. This is in good agreement with our results showing that *Bdnf* overexpression in the hippocampus of BTBR mice leads to an increase in the *Drd1* gene expression along with a decrease in anxiety and stereotypy (Ilchibaeva et al., 2023).

The regulation of DA neurotransmission is also often associated with the recently discovered non-canonical neurotrophin CDNF that has neuroprotective properties in conditions associated with degeneration of DA neurons (Voutilainen et al., 2011). There are currently no data on the role of CDNF in the pathogenesis of ASD. Here we showed for the first time the decrease in the *Cdnf* expression in the frontal cortex of BTBR mice throughout the entire studied period of postnatal development, starting with eye opening at day p14 till the onset of sexual maturity at day p60. Considering that CDNF is characterized by anti-apoptotic and neurotrophic effects (Boĥok et al., 2018), it was suggested that there is an increased risk of cell death activation as well as reduction of cytoprotective properties in the frontal cortex of BTBR mice during the postnatal development, which may lead, among other things, to disturbances in DA neurotransmission and, hence, manifestation of autism-like behavior.

Conclusion

Thus, in the hippocampus and frontal cortex of BTBR mice, characterized by autism-like behavior, a significant dysregulation of the expression patterns of *Cdnf*, key DA receptors, *Bdnf* and its receptors, as well as the transcription factor CREB was shown. It was suggested that the identified disturbances in the expression of the studied genes on the 14th day of embryogenesis are critical for the formation of an autism-like phenotype. The decrease in the expression of *Cdnf*, as well as *Bdnf* and its receptor *Ntrkb2* in the frontal cortex during the studied periods of postnatal development apparently results in critical changes in the morphology of neurons in cortical brain regions. At the same time, the revealed decrease in *Drd2* gene expression in the postnatal period may be associated with learning and memory impairments observed in BTBR mice.

References

Adhya D., Swarup V., Nagy R., Dutan L., Shum C., Valencia-Alarcyn E.P., Jozwik K.M., Mendez M.A., Horder J., Loth E., Nowosiad P., Lee I., Skuse D., Flinter F.A., Murphy D., McAlonan G., Geschwind D.H., Price J., Carroll J., Srivastava D.P., Baron-Cohen S. Atypical neurogenesis in induced pluripotent stem cells from autistic individuals. *Biol. Psychiatry.* 2021;89(5):486-496. DOI 10.1016/j.biopsych.2020.06.014

Aljanabi S.M., Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 1997;25(22):4692-4693. DOI 10.1093/nar/25.22.4692

Anghelescu I., Dettling M. Neuron number in children with autism. *JAMA.* 2012;307(8):783. DOI 10.1001/jama.2012.191

Belokopytova I.I., Kondaurova E.M., Kulikova E.A., Ilchibaeva T.V., Naumenko V.S., Popova N.K. Effects of the *Cc2d1a*/Freud-1 knockdown in the hippocampus of BTBR mice on the autistic-like behavior, expression of serotonin 5-HT1A and D2 dopamine receptors, and CREB and NF-κB intracellular signaling. *Biochemistry* (*Mosc*.). 2022;87(10):1206-1218. DOI 10.1134/S0006297922100145

Bohlen M.O., Bailoo J.D., Jordan R.L., Wahlsten D. Hippocampal commissure defects in crosses of four inbred mouse strains with absent corpus callosum. *Genes Brain Behav.* 2012;11(7):757-766. DOI 10.1111/j.1601-183X.2012.00802.x

Boĥok V., Yu L.Y., Palgi J., Arumäe U. Antioxidative CXXC peptide motif from mesencephalic astrocyte-derived neurotrophic factor antagonizes programmed cell death. *Front. Cell Dev. Biol.* 2018;6:106. DOI 10.3389/fcell.2018.00106

Bolivar V.J., Walters S.R., Phoenix J.L. Assessing autism-like behavior in mice: variations in social interactions among inbred strains. *Behav. Brain Res.* 2007;176(1):21-26. DOI 10.1016/j.bbr.2006.09.007

Castrén E., Antila H. Neuronal plasticity and neurotrophic factors in drug responses. *Mol. Psychiatry.* 2017;22(8):1085-1095. DOI 10.1038/mp.2017.61

Cellot G., Maggi L., Di Castro M.A., Catalano M., Migliore R., Migliore M., Scattoni M.L., Calamandrei G., Cherubini E. Premature changes in neuronal excitability account for hippocampal network impairment and autistic-like behavior in neonatal BTBR T+tf/J mice. *Sci. Rep.* 2016;6:31696. DOI 10.1038/srep31696

Chen V.S., Morrison J.P., Southwell M.F., Foley J.F., Bolon B., Elmore S.A. Histology atlas of the developing prenatal and postnatal mouse central nervous system, with emphasis on prenatal days E7.5 to E18.5. *Toxicol. Pathol.* 2017;45(6):705-744. DOI 10.1177/0192623317728134

Courchesne E., Mouton P.R., Calhoun M.E., Semendeferi K., Ahrens-Barbeau C., Hallet M.J., Barnes C.C., Pierce K. Neuron number and size in prefrontal cortex of children with autism. *JAMA.* 2011;306(18):2001-2010. DOI 10.1001/jama.2011.1638

Courchesne E., Pramparo T., Gazestani V.H., Lombardo M.V., Pierce K., Lewis N.E. The ASD Living Biology: from cell proliferation to clinical phenotype. *Mol. Psychiatry.* 2019;24(1):88-107. DOI 10.1038/s41380-018-0056-y

Courchesne E., Gazestani V.H., Lewis N.E. Prenatal origins of ASD: The when, what, and how of ASD development. *Trends Neurosci.* 2020;43(5):326-342. DOI 10.1016/j.tins.2020.03.005

Crandall J.E., McCarthy D.M., Araki K.Y., Sims J.R., Ren J.Q., Bhide P.G. Dopamine receptor activation modulates GABA neuron migration from the basal forebrain to the cerebral cortex. *J. Neurosci*. 2007;27(14):3813-3822. DOI 10.1523/JNEUROSCI.5124-06.2007

Crawley J.N. Twenty years of discoveries emerging from mouse models of autism. *Neurosci. Biobehav. Rev.* 2023;146:105053. DOI 10.1016/j.neubiorev.2023.105053

DiCarlo G.E., Aguilar J.I., Matthies H.J., Harrison F.E., Bundschuh K.E., West A., Hashemi P., Herborg F., Rickhag M., Chen H., Gether U., Wallace M.T., Galli A. Autism-linked dopamine transporter mutation alters striatal dopamine neurotransmission and dopamine-dependent behaviors. *J. Clin. Invest.* 2019;129(8):3407-3419. DOI 10.1172/JCI127411

Enticott P.G., Kennedy H.A., Rinehart N.J., Tonge B.J., Bradshaw J.L., Fitzgerald P.B. GABAergic activity in autism spectrum disorders: an investigation of cortical inhibition via transcranial magnetic stimulation. *Neuropharmacology.* 2013;68:202-209. DOI 10.1016/j.neuropharm.2012.06.017

Fenner B.M. Truncated TrkB: beyond a dominant negative receptor. *Cytokine Growth Factor Rev.* 2012;23(1-2):15-24. DOI 10.1016/j.cytogfr.2012.01.002

Finlay B.L., Darlington R.B. Linked regularities in the development and evolution of mammalian brains. *Science.* 1995;268(5217):1578-1584. DOI 10.1126/science.7777856

Frazier T.W., Hardan A.Y. A meta-analysis of the corpus callosum in autism. *Biol. Psychiatry.* 2009;66(10):935-941. DOI 10.1016/j.biopsych.2009.07.022

Garcia K.L., Yu.G., Nicolini C., Michalski B., Garzon D.J., Chiu V.S., Tongiorgi E., Szatmari P., Fahnestock M. Altered balance of proteolytic isoforms of pro-brain-derived neurotrophic factor in autism. *J. Neuropathol. Exp. Neurol.* 2012;71(4):289-297. DOI 10.1097/NEN.0b013e31824b27e4

Hashem S., Nisar S., Bhat A.A., Yadav S.K., Azeem M.W., Bagga P., Fakhro K., Reddy R., Frenneaux M.P., Haris M. Genetics of structural and functional brain changes in autism spectrum disorder. *Transl. Psychiatry.* 2020;10(1):229. DOI 10.1038/s41398-020-00921-3

Hettinger J.A., Liu X., Schwartz C.E., Michaelis R.C., Holden J.J. A DRD1 haplotype is associated with risk for autism spectrum disorders in male-only affected sib-pair families. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 2008;147B(5):628-636. DOI 10.1002/ajmg.b.30655

Ilchibaeva T., Tsybko A., Lipnitskaya M., Eremin D., Milutinovich K., Naumenko V., Popova N. Brain-derived neurotrophic factor (BDNF) in mechanisms of autistic-like behavior in BTBR mice: crosstalk with the dopaminergic brain system. *Biomedicines.* 2023;11(5):1482. DOI 10.3390/biomedicines11051482

Kanner L. Autistic disturbances of affective contact. *Nervous Child.* 1943;2:217-250

Kemper T.L., Bauman M. Neuropathology of infantile autism. *J. Neuropathol. Exp. Neurol.* 1998;57(7):645-652. DOI 10.1097/00005072-199807000-00001

Kondaurova E.M., Plyusnina A.V., Ilchibaeva T.V., Eremin D.V., Rodnyy A.Y., Grygoreva Y.D., Naumenko V.S. Effects of a Cc2d1a/Freud-1 knockdown in the hippocampus on behavior, the serotonin system, and BDNF. *Int. J. Mol. Sci.* 2021;22(24):13319. DOI 10.3390/ijms222413319

Koshimizu H., Kiyosue K., Hara T., Hazama S., Suzuki S., Uegaki K., Nagappan G., Zaitsev E., Hirokawa T., Tatsu Y., Ogura A., Lu B., Kojima M. Multiple functions of precursor BDNF to CNS neurons: negative regulation of neurite growth, spine formation and cell survival. *Mol. Brain*. 2009;2:27. DOI 10.1186/1756-6606-2-27

Krishnan A., Zhang R., Yao V., Theesfeld C.L., Wong A.K., Tadych A., Volfovsky N., Packer A., Lash A., Troyanskaya O.G. Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. *Nat. Neurosci.* 2016;19(11):1454-1462. DOI 10.1038/nn.4353

Kulikov A.V., Naumenko V.S., Voronova I.P., Tikhonova M.A., Popova N.K. Quantitative RT-PCR assay of 5-HT1A and 5-HT2A serotonin receptor mRNAs using genomic DNA as an external standard. *J. Neurosci. Methods.* 2005;141(1):97-101. DOI 10.1016/j.jneumeth.2004.06.005

Küppers E., Beyer C. Dopamine regulates brain-derived neurotrophic factor (BDNF) expression in cultured embryonic mouse striatal cells. *Neuroreport.* 2001;12(6):1175-1179. DOI 10.1097/00001756-200105080-00025

Lessmann V., Gottmann K., Malcangio M. Neurotrophin secretion: current facts and future prospects. *Prog. Neurobiol.* 2003;69(5):341-374. DOI 10.1016/s0301-0082(03)00019-4

Lindholm P., Saarma M. Cerebral dopamine neurotrophic factor protects and repairs dopamine neurons by novel mechanism. *Mol. Psychiatry*. 2022;27(3):1310-1321. DOI 10.1038/s41380-021-01394-6

Lindholm P., Voutilainen M.H., Laurén J., Peränen J., Leppänen V.M., Andressoo J.O., Lindahl M., Janhunen S., Kalkkinen N., Timmusk T., Tuominen R.K., Saarma M. Novel neurotrophic factor CDNF protects and rescues midbrain dopamine neurons *in vivo*. *Nature.* 2007;448(7149):73-77. DOI 10.1038/nature05957

Lindholm P., Peränen J., Andressoo J.O., Kalkkinen N., Kokaia Z., Lindvall O., Timmusk T., Saarma M. MANF is widely expressed in mammalian tissues and differently regulated after ischemic and epileptic insults in rodent brain. *Mol. Cell. Neurosci.* 2008;39(3):356-371. DOI 10.1016/j.mcn.2008.07.016

Liu S.H., Shi X.J., Fan F.C., Cheng Y. Peripheral blood neurotrophic factor levels in children with autism spectrum disorder: a meta-analysis. *Sci. Rep.* 2021;11(1):15. DOI 10.1038/s41598-020-79080-w

Loones M.T., Chang Y., Morange M. The distribution of heat shock proteins in the nervous system of the unstressed mouse embryo suggests a role in neuronal and non-neuronal differentiation. *Cell Stress Chaperones.* 2000;5(4):291-305. DOI 10.1379/1466-1268(2000)005<0291:tdohsp>2.0.co;2

Lu B., Pang P.T., Woo N.H. The yin and yang of neurotrophin action. *Nat. Rev. Neurosci.* 2005;6(8):603-614. DOI 10.1038/nrn1726

Mangale V.S., Hirokawa K.E., Satyaki P.R., Gokulchandran N., Chikbire S., Subramanian L., Shetty A.S., Martynoga B., Paul J., Mai M.V., Li Y., Flanagan L.A., Tole S., Monuki E.S. Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. *Science.* 2008;319(5861):304-309. DOI 10.1126/science.1151695

Mariani J., Coppola G., Zhang P., Abyzov A., Provini L., Tomasini L., Amenduni M., Szekely A., Palejev D., Wilson M., Gerstein M., Grigorenko E.L., Chawarska K., Pelphrey K.A., Howe J.R., Vaccarino F.M. FOXG1-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell*. 2015;162(2):375-390. DOI 10.1016/j.cell.2015.06.034

McFarlane H.G., Kusek G.K., Yang M., Phoenix J.L., Bolivar V.J., Crawley J.N. Autism-like behavioral phenotypes in BTBR T+tf/J mice. *Genes Brain Behav.* 2008;7(2):152-163. DOI 10.1111/j.1601-183X.2007.00330.x

Minshew N.J., Williams D.L. The new neurobiology of autism: cortex, connectivity, and neuronal organization. *Arch. Neurol.* 2007;64(7):945-950. DOI 10.1001/archneur.64.7.945

Naumenko V.S., Kulikov A.V. Quantitative assay of 5-HT_{1A} receptor gene expression in the brain. *Mol. Biol.* (*Mosk.*). 2006;40(1):30-36. DOI 10.1134/S0026893306010067

Naumenko V.S., Osipova D.V., Kostina E.V., Kulikov A.V. Utilization of a two-standard system in real-time PCR for quantification of gene expression in the brain. *J. Neurosci. Methods*. 2008;170(2):197-203. DOI 10.1016/j.jneumeth.2008.01.008

Neal M., Cunningham J., Lever I., Pezet S., Malcangio M. Mechanism by which brain-derived neurotrophic factor increases dopamine release from the rabbit retina. *Invest. Ophthalmol. Vis. Sci.* 2003;44(2):791-798. DOI 10.1167/iovs.02-0557

Nguyen H.T.N., Kato H., Masuda K., Yamaza H., Hirofuji Y., Sato H., Pham T.T.M., Takayama F., Sakai Y., Ohga S., Taguchi T., Nonaka K. Impaired neurite development associated with mitochondrial dysfunction in dopaminergic neurons differentiated from exfoliated deciduous tooth-derived pulp stem cells of children with autism spectrum disorder. *Biochem. Biophys. Rep.* 2018;16:24-31. DOI 10.1016/j.bbrep.2018.09.004

Palasz E., Wysocka A., Gasiorowska A., Chalimoniuk M., Niewiadomski W., Niewiadomska G. BDNF as a promising therapeutic agent in Parkinson's disease. *Int. J. Mol. Sci.* 2020;21(3):1170. DOI 10.3390/ijms21031170

Pavăl D. A Dopamine hypothesis of autism spectrum disorder. *Dev. Neurosci.* 2017;39(5):355-360. DOI 10.1159/000478725

Popolo M., McCarthy D.M., Bhide P.G. Influence of dopamine on precursor cell proliferation and differentiation in the embryonic mouse telencephalon. *Dev. Neurosci.* 2004;26(2-4):229-244. DOI 10.1159/00008214

Popova N.K., Naumenko V.S. Neuronal and behavioral plasticity: the role of serotonin and BDNF systems tandem. *Expert Opin. Ther. Targets.* 2019;23(3):227-239. DOI 10.1080/14728222.2019.1572747

Pourhamzeh M., Moravej F.G., Arabi M., Shahriari E., Mehrabi S., Ward R., Ahadi R., Joghataei M.T. The roles of serotonin in neuro-

psychiatric disorders. *Cell. Mol. Neurobiol.* 2022;42(6):1671-1692. DOI 10.1007/s10571-021-01064-9

Richards L.J., Plachez C., Ren T. Mechanisms regulating the development of the corpus callosum and its agenesis in mouse and human. *Clin. Genet.* 2004;66(4):276-289. DOI 10.1111/j.1399-0004.2004.00354.x

Rochefort N.L., Garaschuk O., Milos R.I., Narushima M., Marandi N., Pichler B., Kovalchuk Y., Konnerth A. Sparsification of neuronal activity in the visual cortex at eye-opening. *Proc. Natl. Acad. Sci. USA.* 2009;106(35):15049-15054. DOI 10.1073/pnas.0907660106

Rodnyy A.Y., Kondaurova E.M., Tsybko A.S., Popova N.K., Kudlay D.A., Naumenko V.S. The brain serotonin system in autism. *Rev. Neurosci.* 2023;35(1):1-20. DOI 10.1515/revneuro-2023-0055

Sacco R., Gabriele S., Persico A.M. Head circumference and brain size in autism spectrum disorder: A systematic review and meta-analysis. *Psychiatry Res.* 2015;234(2):239-251. DOI 10.1016/j.pscychresns.2015.08.016

Satterstrom F.K., Kosmicki J.A., Wang J., Breen M.S., De Rubeis S., An J.Y., Peng M., … Guerrero E.E., Dias C.; Autism Sequencing Consortium; iPSYCH-Broad Consortium; Betancur C., Cook E.H., Gallagher L., Gill M., Sutcliffe J.S., Thurm A., Zwick M.E., Børglum A.D., State M.W., Cicek A.E., Talkowski M.E., Cutler D.J., Devlin B., Sanders S.J., Roeder K., Daly M.J., Buxbaum J.D. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell.* 2020;180(3):568-584.e23. DOI 10.1016/j.cell.2019.12.036

Segura M., Pedreño C., Obiols J., Taurines R., Pamias M., Grünblatt E., Gella A. Neurotrophin blood-based gene expression and social cognition analysis in patients with autism spectrum disorder. *Neurogenetics.* 2015;16(2):123-131. DOI 10.1007/s10048-014-0434-9

Staal W.G., Langen M., van Dijk S., Mensen V.T., Durston S. *DRD3* gene and striatum in autism spectrum disorder. *Br. J. Psychiatry.* 2015;206(5):431-432. DOI 10.1192/bjp.bp.114.148973

Teng H.K., Teng K.K., Lee R., Wright S., Tevar S., Almeida R.D., Kermani P., Torkin R., Chen Z.Y., Lee F.S., Kraemer R.T., Nykjaer A., Hempstead B.L. ProBDNF induces neuronal apoptosis via activation of a receptor complex of p75NTR and sortilin. *J. Neurosci.* 2005;25(22):5455-5463. DOI 10.1523/JNEUROSCI.5123-04.2005

Voutilainen M.H., Bäck S., Peränen J., Lindholm P., Raasmaja A., Männistö P.T., Saarma M., Tuominen R.K. Chronic infusion of CDNF prevents 6-OHDA-induced deficits in a rat model of Parkinson's disease. *Exp. Neurol.* 2011;228(1):99-108. DOI 10.1016/j.expneurol.2010.12.013

Wambach C.M., Patel S.N., Kahn D.A. Maternal and fetal factors that contribute to the localization of T regulatory cells during pregnancy. *Am. J. Reprod. Immunol*. 2014;71(5):391-400. DOI 10.1111/aji.12223

Wei H., Alberts I., Li X. The apoptotic perspective of autism. *Int. J. Dev. Neurosci.* 2014;36:13-18. DOI 10.1016/j.ijdevneu.2014.04.004

Yang J., Siao C.J., Nagappan G., Marinic T., Jing D., McGrath K., Chen Z.Y., Mark W., Tessarollo L., Lee F.S., Lu B., Hempstead B.L. Neuronal release of proBDNF. *Nat. Neurosci.* 2009;12(2):113-115. DOI 10.1038/nn.2244

Yoo H.J., Cho I.H., Park M., Yang S.Y., Kim S.A. Association of the catechol-o-methyltransferase gene polymorphisms with Korean autism spectrum disorders. *J. Korean Med. Sci.* 2013;28(9):1403-1406. DOI 10.3346/jkms.2013.28.9.1403

Conflict of interest. The authors declare no conflict of interest.
Received January 31, 2024. Revised March 11, 2024. Accepted March 12, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):416-423

DOI 10.18699/vjgb-24-47

History, status and genetic characteristics of native cattle breeds from the Republic of Kazakhstan

A.K. Khamzina [1], A.A. Yurchenko [2], N.S. Yudin [2], P.Sh. Ibragimov [1], Y.S. Ussenbekov [1], D.M. Larkin [3]

[1] Kazakh National Agrarian Research University, Almaty, Kazakhstan
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Royal Veterinary College, University of London, London, United Kingdom
dmlarkin@gmail.com

Abstract. This work provides a comprehensive review of the history, status, and genetic characteristics of cattle breeds in Kazakhstan. The current breeding status is analysed, including information on popular breeds such as Kazakh white-headed, Auliekol, Alatau, Aulieata, and Kalmyk, their production and economic significance. An overview of genetic studies using DNA fingerprinting, microsatellites, and SNPs aimed at identifying unique characteristics, genetic diversity, and genes under selection, as well as markers of economically important and productive traits of Kazakh cattle breeds, is also provided. The study examined the genetic structure of the Kazakh white-headed and Alatau breeds based on whole-genome SNP genotyping. Unique genetic components characterizing Kazakhstan cattle breeds were described, and comparisons were made with genetic data from other breeds. Structural analysis showed that the Kazakh white-headed breed contains genetic components of the Hereford, Kalmyk, and Altai cattle. The Alatau breed has a composite structure, containing components of the Brown Swiss, Braunvieh, Kalmyk, and Holstein breeds. The results not only reveal the genetic diversity and characteristics of cattle breeds in Kazakhstan and the historical development and current state of animal husbandry in the country, but also emphasize the importance of further research to identify adaptive and unique genetic markers affecting economically important traits of local breeds.

Key words: cattle; breeds; history; Kazakhstan; genetic characteristics; single nucleotide polymorphism.

For citation: Khamzina A.K., Yurchenko A.A., Yudin N.S., Ibragimov P.Sh., Ussenbekov Y.S., Larkin D.M. History, status and genetic characteristics of native cattle breeds from the Republic of Kazakhstan. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):416-423. DOI 10.18699/vjgb-24-47

Funding. The work was performed in the framework of a grant funding project for scientific and (or) scientific and technical projects for 2023–2025 (Ministry of Science and Higher Education of the Republic of Kazakhstan) AP19674808 – Creation of genetic passports and study of genetics of local Kazakhstani cattle breeds using genome resequencing.

История, современное состояние и генетическая характеристика локальных пород крупного рогатого скота Республики Казахстан

А.К. Хамзина [1], А.А. Юрченко [2], Н.С. Юдин [2], П.Ш. Ибрагимов [1], Е.С. Усенбеков [1], Д.М. Ларкин [3]

[1] Казахский национальный аграрный исследовательский университет, Алматы, Казахстан
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Королевский ветеринарный колледж, Университет Лондона, Лондон, Великобритания
dmlarkin@gmail.com

Аннотация. В этой работе мы представляем обзор истории происхождения, текущего состояния и генетических особенностей пород крупного рогатого скота (КРС) Казахстана. Отражена информация о современном состоянии разведения, в том числе о пяти наиболее популярных локальных породах – казахской белоголовой, аулиекольской, алатауской, аулиеатинской и калмыцкой, в рамках их производственной и экономической значимости. Приведен обзор генетических исследований по этим породам, направленных на выявление их уникальных характеристик, генетического разнообразия и генов, находящихся под давлением отбора, а также ДНК-маркеров экономически важных и продуктивных признаков казахстанских пород КРС. Представлены также оригинальные данные по особенностям генетической структуры казахской белоголовой и алатауской пород на основе результатов полногеномного генотипирования однонуклеотидных полиморфизмов. Описаны уникальные генетические компоненты, характеризующие эти породы КРС Казахстана, и проведено их сравнение с генетической структурой популяций этих же и других пород из Российской Федерации. Структурный анализ показал, что казахская белоголовая порода содержит генетические компоненты герефордской и калмыцкой по-

род, а также алтайского скота. Алатауская порода имеет композитную структуру и содержит компоненты бурой швицкой, браунви, калмыцкой и голштинской пород. Результаты демонстрируют генетическое разнообразие и особенности пород КРС Казахстана, развитие и текущее состояние скотоводства в стране, а также подчеркивают необходимость более детальных исследований для выявления генетических маркеров адаптации и экономически важных признаков местных пород.
Ключевые слова: крупный рогатый скот; породы; история; Казахстан; генетическая характеристика; однонуклеотидный полиморфизм.

Introduction

For over 10,000 years, cattle have been an important element of agriculture and food production (Argynbaev, 1969; Dakhshleyger, 1980). The first mention of cattle breeding on the territory of Kazakhstan dates back to the Botai culture of the Bronze Age (III–II centuries BC). The study of bone remains indicates that these herds included mainly horses, but the remains of small and large livestock were also found (Adilova, Ilyassov, 2018). In the subsequent period (from the 15th to the 17th centuries), cattle breeding in limited quantities was noted in the Kazakh Khanate (Ratchnevsky, 1993; Allsen, 2001). Before the second half of the 19th century and early 20th century, Kazakhs practised, for the most part, a nomadic form of agriculture (Frizen, 2022). Nomadic cattle farming was extensive, using vast grazing areas rather than intensive farming methods in a limited area. They used pastures, where livestock lived throughout the year or almost all year round on natural pasture. This also determined the composition of the herd, which could only include animals able to pasture during the winter (Diarov, 1963; Argynbayev, 1969). To the most extent, these were horses and sheep, which made up the majority of the nomadic herd (Diarov, 1963; Tolybekov, 1971). In the 19th century, due to socio-economic changes, new forms of economy began to appear, such as semi-nomadic cattle breeding and agriculture.

A distinctive feature of semi-nomadic cattle breeding was that it was combined with agriculture (Tolybekov, 1971). Haymaking and farming are associated with an increase in the share of cattle in the Kazakh nomadic economy. Cattle became the main traction force in this type of farming. Another circumstance that contributed to cattle breeding in Kazakhstan was the emergence of a market for the sale of meat (Tolybekov, 1971).

Kazakhstan, a vast Central Asian country known for its diverse landscapes from steppes to mountains, has developed cattle breeds that meet specific human needs and are adapted to the local environment. Kazakh cattle breeds were formed in such a way as to live in the harsh, often extreme conditions of this country, while at the same time having high productivity (Diarov, 1963).

Modern local cattle breeds are characterized by their adaptability, sustainability, and ability to provide the population with necessary resources, such as meat, milk and hides (Kazkenova, Ainakanova, 2016). These breeds are the most important asset of Kazakhstan and the whole world. Therefore, it is necessary to study their unique genetics for subsequent improvement, as well as to create new commercial breeds that can maintain their outstanding properties in the harsh, sharply continental steppe climate of Kazakhstan and other countries.

In this work, we will review the literature describing modern cattle breeds in Kazakhstan, their commercial properties and genetic characteristics, and also present our data on the genetic structure of populations of two breeds: the Kazakh white-headed and Alatau based on data from whole-genome genotyping of samples of these breeds from Kazakhstan and the Russian Federation and their comparisons with other breeds.

Current status, distribution area and description of breeds

According to the Bureau of National Statistics of the Republic of Kazakhstan, as of March 1, 2023, the total number of livestock in the region is over 10 million heads (Fig. 1). This figure is higher than in previous years and indicates a positive growth trend of ~4 % per year (www.stat.gov.kz).

There are 23 cattle breeds in the country, registered in the information and analytical system of Kazakhstan (www.plem.kz), including four breeds that resulted from crossing imported breeds with local livestock and are well adapted to the harsh climatic conditions of the region (Alatau, Aulieata, Kazakh white-headed and Auliekol) (Diarov, 1963; Torekhanov et al., 2006). These breeds have unique features, such as adaptation to harsh climatic conditions (extreme temperatures) and limited access to feed (Torekhanov et al., 2011), resistance to local diseases and parasites (Sattarova et al., 2023), high meat and dairy productivity in the country's conditions (Torekhanov et al., 2011). To create these breeds, breeds imported to Kazakhstan from other countries were used to improve the economic characteristics of local livestock or adapt to new conditions of keeping and growing (Kazhgaliyev et al., 2016; Zhumanov, Baimukanov, 2020; Ulimbashev et al., 2023).

The main imported breeds that are now successfully bred in Kazakhstan include Kalmyk, Angus, Hereford, Holstein, Kholmogory, Limousin, Santa Gertrudis and others. According to the Republican Chamber of Dairy and Combined Cattle Breeds of Kazakhstan (www.qazaqsut.kz), which includes Alatau, Aulieata, Holstein, Black pied and other breeds of this productivity, the number of stud farms for 2023 is 628 farms, and the number of commercial ones is 1,271 farms. Most dairy and combined productivity cattle are Simmental and Holstein breeds.

Alatau cattle breed – meat and dairy productivity breed (Fig. 2*a*). Research on breeding the Alatau breed was carried out in 1930–1950 in the Kirghiz SSR and the southern regions of the Kazakh SSR by crossing local cattle with animals of the Kostroma and Brown Swiss breeds (Nysanbaev, 2004). The breed is adapted to living in high mountain areas, its colour is mostly brown, of different shades. As of the beginning of 2024, the population of breeding cattle of this breed

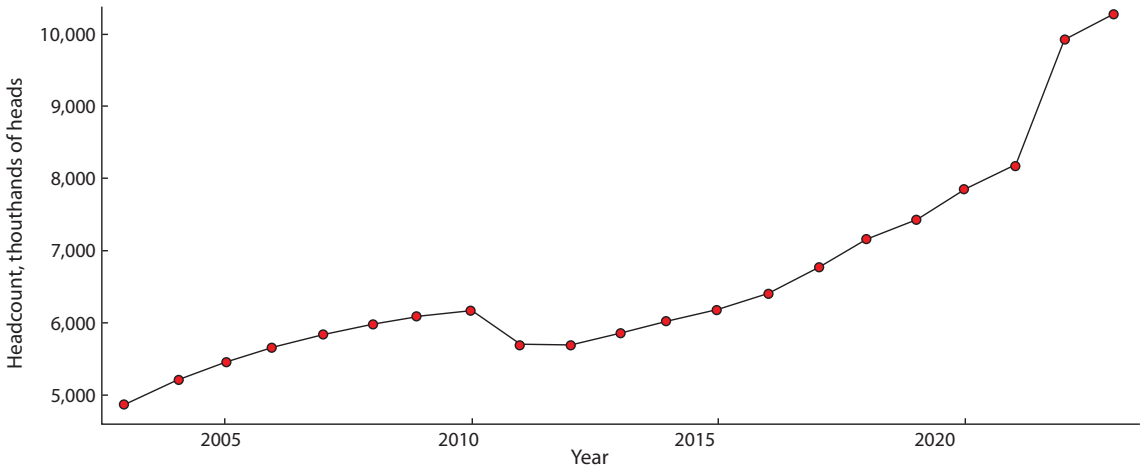


Fig. 1. Total headcount of cattle in the Republic of Kazakhstan from 2003 to 2023 (www.stat.gov.kz).

Fig. 2. Alatau breed (*a*); Aulieata breed (*b*); Kazakh white-headed breed (*c*); Auliekol breed (*d*); Kalmyk breed (*e*).

is about 7 thousand heads, which is ~2.8 percent higher than the previous year (www.qazaqsut.kz). This breed is mainly bred in the Almaty and Turkestan regions of the Republic (www.gov.kz).

The Alatau cattle breed from Kazakhstan has been the subject of several studies aimed at improving its breeding and rearing. Thus, A.D. Baimukanov and co-authors (Baimukanov et al., 2021) focused on the effective breeding of the Kazakh population, while S.K. Abugaliev and co-authors (Abugaliev et al., 2020) studied the growth and development of heifer calves under various rearing technologies.

The Aulieata dairy production breed (Fig. 2*b*) was bred in the Kyrgyz and Kazakh SSR by crossing local cattle with Dutch cattle and subsequent inbreeding of crossbreeds (Nysanbaev, 2004). It was first tested in 1952. A distinctive feature of the breed is its adaptability to breeding in hot climates

and resistance to blood-parasitic diseases. The colour of the animals is predominantly black-and-white, but light grey is also found. The Aulieata breed fattens well. The animals are characterized by a purely milky body type, a well-developed udder, and correctly positioned limbs (Nysanbaev, 2004). As of 2024, the number of pedigree cattle of the Aulieata breed in Kazakhstan is about 1 thousand heads (www.qazaqsut.kz), with the main breeding happening in the south of the country (www.gov.kz).

The Kazakh white-headed meat production breed was developed in the USSR in 1930–1940 (Fig. 2*c*), and it was officially tested in 1950. The selection was carried out by complex reproductive crossing of a breeding stock of local Kazakh and partly Kalmyk cattle with Hereford bulls, as a result of which the cattle acquired the best qualities of all these animals: high adaptive ability, strong constitution, early maturity and high meat yield (Porter, 2016). The colour is red, of varying intensity, with a white head, chest, belly, lower limbs and tail brush. There are animals with white markings on the withers and rump; the front part is better developed than the back part; the hair is thick and short in summer, and long and slightly curly in winter (Dmitriev, Ernst, 1989; Nysanbaev, 2004). The total number of pedigree cattle of the Kazakh white-headed breed in 2022 is about 500 thousand heads, including about 200 thousand cows (www.gov.kz). The breed is bred countrywide, but the largest population is in the East Kazakhstan region. The Kazakh white-headed cattle breed makes a significant contribution to beef production in Kazakhstan (Bozymov, 2018).

The Auliekol breed was created by a complex reproductive crossing of three specialized meat breeds: Kazakh white-headed, Charolais and Aberdeen Angus (Fig. 2*d*). Per the international classification, the breed belongs to large breeds of beef cattle. It was registered in 1992 (Nysanbaev, 2004). The breed is located mainly in the Kostanay region; it was also imported to the farms of the Pavlodar, North Kazakhstan, Almaty, and Karaganda regions. As of 2022, the number of pedigree cattle of the Auliekol breed is about 70 thousand heads, of which approximately 33 thousand are cows (www.gov.kz). The specialized Auliekol meat breed is characterized by good early maturity, high yield and quality of meat, high growth energy, and adaptability to local conditions. The colour of the animals is light grey, 70 % of the livestock are polled. Animals have a strong constitution. In winter, they grow thick hair and are well adapted to harsh natural and climatic conditions of a sharply continental climate (Nysanbaev, 2004). In summer, animals quickly gain weight, easily tolerate heat, and in winter they are tolerant to frost when outdoors.

The Kalmyk meat production breed has been bred in Kazakhstan since the 17th century (Fig. 2*e*). It was introduced by nomadic Kalmyk tribes more than 350 years ago from the western part of Mongolia and China (Bichurin, 1991; www.qalmaq.kz). The final formation of the Kalmyk breed took place in the conditions of a nomadic economy with year-round grazing of animals. Cows of the Kalmyk breed are, in general, medium in size and compact in build (Narmaev, 1963). The colour of the animals is red, with white markings on the head, belly or limbs. In winter, cows of the Kalmyk breed grow thick hair. As of 2022, the total number of breeding cattle of the Kalmyk breed is about 23 thousand heads, including about 15 thousand cows. The Kalmyk breed is mainly bred in the Zhambyl and Turkestan regions of Kazakhstan (www.gov.kz).

The Kalmyk cattle breed, which belongs to the group of Turano-Mongolian breeds (Yurchenko et al., 2018a), has high adaptive abilities and similar production and reproductive characteristics to the Mongolian breed (Fedotova et al., 2020). The productivity of Kalmyk bulls varies depending on breeding methods, while bulls of the Kalmyk breed of Buryat selection have a higher live weight compared to bulls of Kalmyk and Rostov selections (Lumbunov, Garmaev, 2021).

Genetic characteristics of cattle breeds in Kazakhstan

Molecular genetic studies of Kazakhstan cattle breeds have so far been carried out using DNA fingerprinting, microsatellites, and SNP markers. These DNA markers are highly informative and variable for studying genetic diversity. However, in most cases, the analysis includes a limited number of markers, which does not provide a comprehensive study of the animal genome.

Population structure. Analysis of the genotypes of three cattle breeds in Kazakhstan (Terletsky et al., 2019), Alatau, Kazakh white-headed and Auliekol, was carried out by DNA fingerprinting using DNA probes, which revealed the highest degree of genetic similarity in animals of the Auliekol breed (BS = 0.64), then in the Alatau breed (BS = 0.54), and the smallest, in the Kazakh white-headed breed (BS = 0.51). The genetic distance between the Kazakh white-headed and Auliekol breeds was the smallest (D = 0.025), which confirms their known genetic relationship. The Alatau breed showed the highest distance from the Kazakh white-headed and Auliekol breeds (D = 0.055 and D = 0.060, respectively). Heterozygosity (H) values are higher in the Kazakh white-headed breed (0.54), which exceeds the value of the Auliekol breed (0.38), confirming the higher genetic variability of the former breed (Terletsky et al., 2019).

Analysis of 12 microsatellite loci confirmed the relationship of the Kazakh white-headed breed with the Hereford breed, which is associated with the use of Hereford bulls for its creation (Shamshidin et al., 2019; Abdelmanova et al., 2021). This is confirmed by data from genome-wide genotyping of 154 thousand SNP markers, where animals of the Kazakh white-headed breed of Russian selection formed a cluster both in principal component analysis (PCA) and in structural and phylogenetic analyses, with the Hereford breed (Yurchenko et al., 2018b; Yudin, Larkin, 2019; Beishova et al., 2022a). On the other hand, the Kazakh white-headed breed has a high level of genetic diversity and has retained a significant fraction of Turano-Mongolian genetic components, which most likely originate from local Kazakh cattle and Kalmyk breeds.

Clustering of SNP markers revealed the genetic relationship of the Alatau breed with the Kostroma, Brown Swiss and Braunvieh breeds, which confirms the known history of the formation of the Alatau and Kostroma breeds (Yudin, Larkin,

2019). Of the microsatellite alleles found in museum Kalmyk cattle samples, more than 80 % were also present in modern representatives of the breed (Abdelmanova et al., 2021).

As a result of genome-wide genotyping of SNP markers, a genetic relationship was revealed between the Kalmyk breed and the Serbian Busha breed (Iso-Touru et al., 2016). In turn, the Auliekol breed showed heterogeneity using SNP genotyping of 154 thousand markers, forming its own cluster in PCA and structural (ADMIXTURE) analyses (Beishova et al., 2022a).

The distribution and frequency of regions of homozygosity (ROH) in the genomes of the Kazakh white-headed and Auliekol cattle breeds were studied as well (Beishova et al., 2022b). In this study, it was shown that the Kazakh white-headed breed had a higher number of ROHs (55.976) compared to the Auliekol breed (13.137). Calculation of the average ROH length showed differences between the values of the Kazakh white-headed (211.59 ± 92.98 Mb) and Auliekol (99.62 ± 46.48 Mb) breeds.

Genes under selective pressure. When analysing genetic signatures of selection in the Kazakh white-headed breed, regions of the *KIT, KITLG* and *EDN3* genes were identified, associated with white, roan coat colour and the "white head" phenotype, respectively (Yudin, Larkin, 2019). Analysis of haplotype frequencies from genome-wide genotyping data showed that the Kazakh white-headed breed exhibits signals on chromosome 6, in the *LCORL-NCAPG* gene region, which has been associated with a number of growth traits in cattle (average daily weight gain, muscle development, and carcass traits). The selection was also found in the interval on chromosome 14 containing the *DGAT1* gene, which contributes to milk fat content.

The *FKBP2* gene, which has been associated with milk protein yield and content, was found to be under selection in the Kazakh white-headed breed. In the Kalmyk breed, the areas under selection were in the region of the *HMGA2* gene, which is associated with growth in cattle, and the *TRPV5* gene, associated with hypocalcemia and postpartum paresis in cattle (Yurchenko et al., 2018b). In the Kalmyk breed, as well as in other Russian breeds, it was found that the *RAD52* gene was subject to selection pressure. This gene is associated with DNA repair and is involved in antiviral defence processes (Yudin, Larkin, 2019).

Genetic markers of economically important traits. Analysis of the association of genotypes for the calpain (*CAPN1*) and somatotropic hormone (*GH*) genes with productivity traits showed that Kazakh white-headed animals homozygous for the *CAPN1* (CC) locus and homozygous for the *GH* (VV) locus are significantly superior to animals without the C and V alleles based on such characteristics as milk productivity, average daily body weight gain, pre-slaughter body weight, slaughter weight, carcass weight, pulp weight, chemical composition and histological characteristics of meat (Plakhtukova et al., 2020). Genetic markers such as blood group antigens A1, A2, D', W, V, and Z have been identified in the Kalmyk breed, which may have potential implications for selection and breeding (Chimidova et al., 2022).

A study of cows of the Aulieata breed in Southern Kazakhstan in comparison with other breeds showed a high occurrence of the kappa-casein gene (*κ-Cn*, *CSN3*) in animals with genotypes AB and BB, as well as a more frequent occurrence of the B allele, which is important for cheese making. Phylogenetic analysis showed that animals of the Aulieata breed are closest to the German black-pied cattle and are included in a common cluster with them. Although the black-pied alleles are rare in the Aulieata breed, they are positively correlated with the level of milk yield over the 305-day lactation period (Alentayev, 2010).

Population genetics analysis of the Kazakh white-headed and Alatau breeds

To carry out this analysis, blood samples of the Alatau breed (40 individuals) were used from Kakpatas LLP in Zhambyl region, 53 blood samples of the Kazakh white-headed breed from the Agro Baltabay peasant farm in Almaty region, 25 hair follicle samples were obtained from the Elimay peasant farm, East Kazakhstan region. Genotyping of DNA samples of the Kazakh white-headed and Alatau breeds was carried out using the BovineSNP50 v.3 array (Illumina, USA) following the manufacturer's protocol at Miratorg-Genetika LLC. The results of the genotyping of these two breeds of Kazakh selection were combined with genotyping data of Altai cattle and closely related breeds from Russia (Yurchenko et al., 2018a) using the PLINK v. 1.9 program (Purcell et al., 2007). Structural analysis of pooled genotyping data from 389 individuals was performed using the fastSTRUCTURE program (Raj et al., 2014).

Analysis of the genetic structure (Fig. 3) of the populations of the Kazakh white-headed and Alatau cattle breeds of Kazakhstan selection in the context of these breeds from Russia and related breeds show the division of breed groups into two main populations at K = 2.

The first group includes the Hereford breed, and the second group includes the Brown Swiss breed. The remaining breeds have either predominantly Hereford components (Kazakh white-headed of Russian and Kazakhstan selections), or predominantly Brown Swiss components (Altai cattle, Kholmogory, Black pied, Holstein, Kalmyk, Alatau of Kazakhstan and Russian selections, Braunvieh, Kostroma).

At K = 3, a component of the Brown Swiss and Kostroma breeds appears, which distinguishes a group of similar breeds (Brown Swiss, Kostroma, Braunvieh, and Alatau breeds of Kazakhstan and Russian selections). A cluster of dairy breeds becomes clear: Kholmogory, Black pied and Holstein. Altai cattle and Kalmyk breeds appear to be hybrid populations. At K = 4, the Kalmyk breed forms a separate cluster, its unique component can be traced in the Kazakh white-headed breed, Altai cattle, as well as the Alatau breed and Braunvieh. At K > 5 this component disappears in the Braunvieh breed. At K = 5, the Kholmogory breed is separated from the general cluster with Holsteins and the Black pied breed. At K = 6, the structure of the Alatau breed appears, which is composite and has components of the Brown Swiss (Kostroma), Braunvieh, Kalmyk, and Holstein (Black pied) breeds.

The proximity to the Kostroma and Swiss breeds is likely explained by the origin of the Alatau breed. Animals of the Alatau breed of Kazakhstan selection have a slightly more pronounced component of Holstein cattle and the Kalmyk

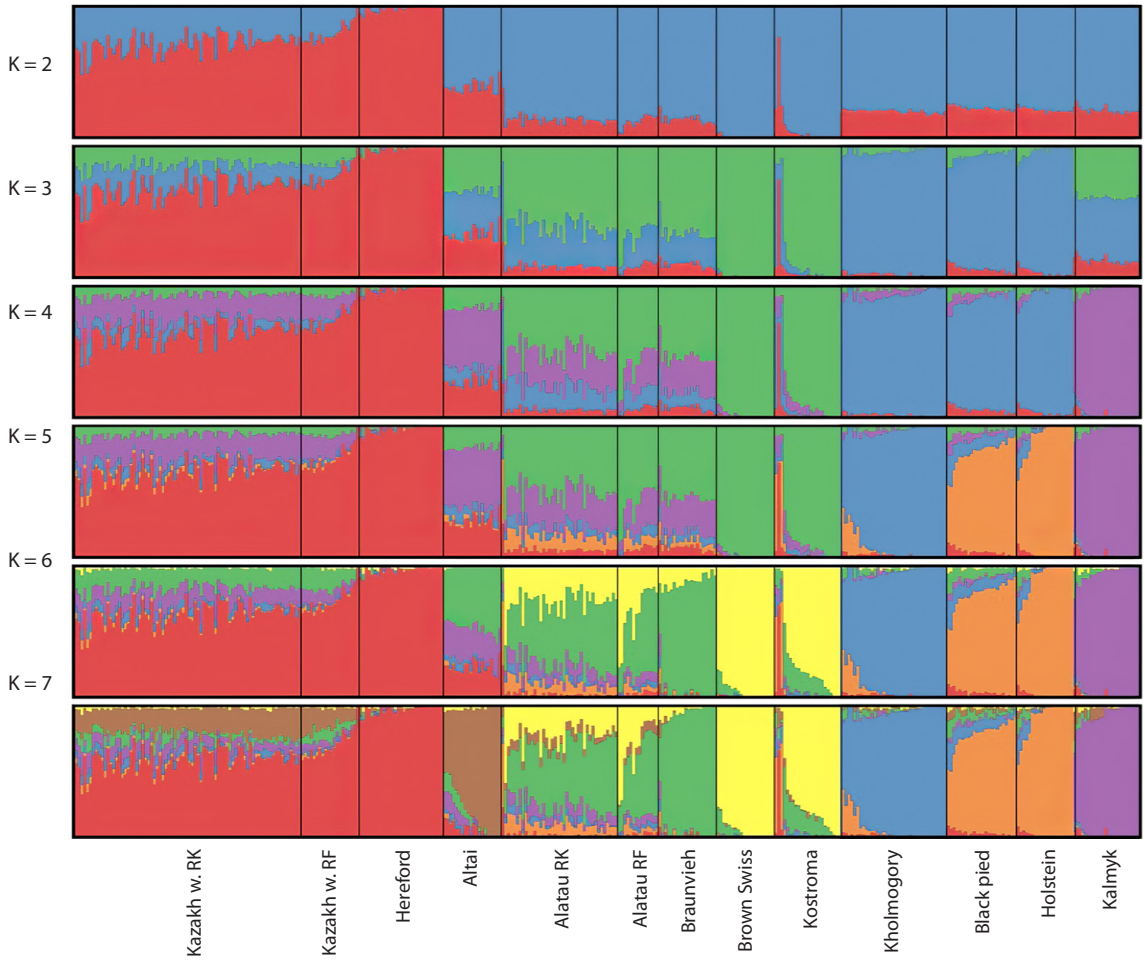


Fig. 3. Genetic structure of cattle breeds from Kazakhstan and other Eurasian breeds.

Kazakh w. RK – Kazakh white-headed breed from the Republic of Kazakhstan; Kazakh w. RF – Kazakh white-headed breed from the Russian Federation; Altai – Altai cattle; Alatau RK – Alatau breed from the Republic of Kazakhstan; Alatau RF – Alatau breed from the Russian Federation. The results of population clustering using the fastSTRUCTURE program from K = 2 to K = 7 are shown.

breed compared to the Russian population. The number of animals of Kazakhstan selection of this breed has increased from ~500 heads to 7 thousand over the past 10 years. Thus, the observed differences may be explained by bottleneck effects and genetic drift. At K = 7, Altai cattle form a separate cluster, a component of which is present in the Kazakh white-headed breed. Thus, the Kazakh white-headed breed has a pronounced component of the Hereford breed, Kalmyk and Altai cattle. Altai cattle are probably close in genetics to the original Kazakh cattle used to produce the Kazakh white-headed breed. On average, the Kazakh white-headed breed of Kazakhstan selection has a smaller component of Hereford and a larger component of Altai and Kalmyk cattle compared to the Kazakh white-headed breed of Russian selection.

In the last decade, active work has been carried out in Kazakhstan to preserve local livestock breeds, including the Kazakh white-headed breed (www.aqbas.kz). One of the goals of this program is to gradually reduce the use of imported breeds in the breeding of local ones. It is possible that this strategy reduced the fraction of Hereford genetics in the Kazakh white-headed population bred in Kazakhstan compared to the population from Russia. Overall, analysis of the genetic structure of these cattle breeds highlights the importance of conserving and maintaining their genetic diversity to ensure resilience and adaptability to changing environmental conditions and livestock production needs.

Conclusion

Throughout the long history of livestock farming in Kazakhstan, unique breeds have been developed and adapted to its climatic and environmental conditions, which play a crucial role in the country's livestock sector. Molecular genetic studies show their closeness not only to European breeds but also to the group of Turano-Mongolian breeds. Recent work on DNA fingerprinting, microsatellites and SNP markers shows that Kazakhstan's cattle need to be studied in more detail to identify adaptive and unique genetic markers for economically important traits of local breeds. The most promising approach may be whole-genome sequencing of the main cattle breeds of Kazakhstan and their comparison with the genomes of breeds from around the world. The emphasis on preserving the genetic diversity of Kazakhstan's cattle breeds is consistent with global efforts to maintain the biodiversity of local domestic animal populations.

References

Abdelmanova A.S., Kharzinova V.R., Volkova V.V., Dotsev A.V., Sermyagin A.A., Boronetskaya O.I., Chinarov R.Y., Brem G., Zinovieva N.A. Analysis of genetic diversity of Kazakh white-headed cattle breed by microsatellites compare with ancestral breeds. In: Proceedings of the 1st International Electronic Conference on Biological Diversity, Ecology and Evolution, 15–31 March 2021. Basel, Switzerland: MDPI, 2021;9487. DOI 10.3390/BDEE2021-09487

Abugaliev S.K., Bupebaeva L.K., Baybatyrova M.A., Matkerimov N.K., Matkerimova K.G. The growth and development of heifers of Alatau breed at different technologies of rearing. *Glavnyj Zootekhnik = Head of Animal Breeding*. 2020;(6):43-49. DOI 10.33920/sel-03-2006-06 (in Russian)

Adilova A.K., Ilyassov Sh.A. Development of cattle breeding in Kazakhstan in the Bronze Age. *Bull. Karaganda Univ*. 2018;1(89): 8-13

Alentayev A. Increase of the Efficiency of Aulieata Cattle Breeding in Southern Kazakhstan by Using the World Gene Pool. Shymkent, 2010 (in Russian)

Allsen T.T. Culture and Conquest in Mongol Eurasia. Cambridge: Cambridge University Press, 2001. DOI 10.1017/CBO9780511497445

Argynbayev Kh. Ethnographic Essays on Kazakh Cattle Breeding. Almaty, 1969 (in Kazakh)

Baimukanov A.D., Artykovich Y.Y., Alexandrovich D.V., Amkhatovich M.T., Ablgazinovich A.K. Efficient breeding in Kazakhstan Alatau cattle breed population. *Am. J. Anim. Vet. Sci*. 2021;16(4): 318-326. DOI 10.3844/ajavsp.2021.318.326

Beishova I., Dossybayev K., Shamshidin A., Belaya A., Bissembayev A., Khamzin K., Kovalchuk A., Nametov A. Population analysis and genetic structure of two Kazakh cattle breeds using 150K SNP. *HAYATI J. Biosci.* 2022a;29(3):301-309. DOI 10.4308/hjb.29.3.301-309

Beishova I., Dossybayev K., Shamshidin A., Belaya A., Bissembayev A., Khamzin K., Kovalchuk A., Nametov A. Distribution of homozygosity regions in the genome of Kazakh cattle breeds. *Diversity*. 2022b;14(4):279. DOI 10.3390/d14040279

Bichurin N.Ya. Historical Overview of the Oirats, or Kalmyks, from the 15th Century to the Present. Elista, 1991 (in Russian)

Bozymov K. Zonal type of Kazakh white-headed breed of cattle in West Kazakhstan. Saarbrucken: Lambert Academic Publishing, 2018

Bureau of National Statistics of Kazakhstan. Official Site. Available at: www.stat.gov.kz

Chimidova N.V., Moiseikina L., Ubushieva A.V., Khakhlinov A.I., Kedeeva O. Genetic structure of population of the Kalmyk breed cattle. *E3S Web Conf*. 2022;363:03025. DOI 10.1051/e3sconf/202236303025

Dakhshleyger G. Economy of the Kazakhs at the Turn of the 19th-20th Centuries: Materials for the Historical and Ethnographic Atlas. Alma-Ata, 1980 (in Russian)

Diarov K. Livestock Farming in Kazakhstan. Alma-Ata, 1963 (in Russian)

Dmitriev N.G., Ernst L.K. (Eds.). Animal Genetic Resources of the USSR. Rome: FAO of the United Nations, 1989

Fedotova G.V., Slozhenkina M.I., Tsitsige, Natyrov A.K., Erendzhenova M.V. Comparative analysis of economic and biological features of Kalmyk and Mongolian cattle breeds. *IOP Conf. Ser. Earth Environ. Sci*. 2020;548(8):082076. DOI 10.1088/1755-1315/548/8/082076

Frizen D.Ya. On the problem of the development of cattle breeding and agriculture among the Kazakhs in the Western Kazakhstan at the turn of the XIX-XX centuries. *Vestnik Tyumenskogo Gosudarstvennogo Universiteta. Gumanitarnye Issledovaniya*. *Humanitates = Tyumen State University Herald. Humanities Research. Humanitates*. 2022;8(1):241-252. DOI 10.21684/2411-197x-2022-8-1-241-252 (in Russian)

Iso-Touru T., Tapio M., Vilkki J., Kiseleva T., Ammosov I., Ivanova Z., Popov R., Ozerov M., Kantanen J. Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe. *Anim. Genet*. 2016;47(6):647-657. DOI 10.1111/age.12473

Kazhgaliyev N.Z., Shauyenov S.K., Omarkozhauly N., Shaikenova K.H., Shurkin A.I. Adaptability and productive qualities of imported beef cattle under the conditions of the Northern region of Kazakhstan. *Biosci. Biotechnol. Res. Asia*. 2016;13(1):531-538. DOI 10.13005/bbra/2065

Kazkenova A., Ainakanova B. Current state of dairy cattle breeding in Mongolia and the Republic of Kazakhstan. *Mongolian Journal of Agricultural Sciences*. 2016;14(1):125-129. DOI 10.5564/mjas.v14i1.592 (in Russian)

Lumbunov S.G., Garmaev B.D. Beef productivity of differently bred steers of the Kalmyk breed. *Glavnyj Zootekhnik = Head of Animal Breeding*. 2021;(11):31-46. DOI 10.33920/sel-03-2111-04 (in Russian)

Nysanbaev A.N. Kazakhstan. National Encyclopedia. Almaty, 2004 (in Russian)

Plakhtukova V., Selionova M.I., Gladkikh M.Y., Kuznetsova O.V. Association of CAPN1 and GH genes with productivity traits in Kazakh Whiteheaded cattle. *E3S Web Conf*. 2020;224:04027. DOI 10.1051/e3sconf/202022404027

Porter V., Alderson L., Hall S.J., Sponenberg D.P. Mason's world encyclopedia of livestock breeds and breeding. Volume 1 and Volume 2. CABI, 2016. DOI 10.1079/9781845934668.0000

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 2007;81(3):559-575. DOI 10.1086/519795

Raj A., Stephens M., Pritchard J.K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 2014;197(2):573-589. DOI 10.1534/genetics.114.164350

Ratchnevsky P. Genghis Khan: His Life and Legacy. Hoboken, New Jersey: Wiley-Blackwell, 1993

Sattarova R., Shynybaev K., Bakiyeva F., Strochkov V., Boranbayeva K., Zhanserkenova O., Kassymbekova S., Ibadullayeva A., Khamzina A. Metagenomic analysis and identification of epizootic strains of the causative agent of infectious bovine keratoconjunctivitis in Kazakhstan. *Int. J. Vet. Sci*. 2023;12(6):822-831. DOI 10.47278/journal.ijvs/2023.071

Shamshidin A.S., Kharzhau A., Baimukanov D.A., Sermyagin A.A. Molecular genetic profile of Kazakhstan populations of cattle breeds. *Bull. Natl. Acad. Sci. Republic Kazakhstan*. 2019;6(382):154-162. DOI 10.32014/2019.2518-1467.157

Terletsky V.P., Tyshchenko V.I., Usenbekov E.S. Genetic features of Kazakh cattle breeds. *Effectivnoye Zhivotnovodstvo = Efficient Livestock Farming*. 2019;(1):64-66. DOI 10.24411/9999-007A-2019-10015 (in Russian)

The Electronic Government of the Republic of Kazakhstan. Ministry of Agriculture. Official site. Available at: www.gov.kz

The Republican Chamber of Dairy and Combined Breeds of Cattle. Official site. Available at: www.qazaqsut.kz

The Republican Chamber of Kazakh White-Headed Cattle Breed. Official site. Available at: www.aqbas.kz

The Republican Chamber of the Kalmyk Breed. Official site. Available at: www.qalmaq.kz

Tolybekov S. The Nomadic Society of Kazakhs in the 17th–early 20th Centuries. Alma-Ata, 1971 (in Russian)

Torekhanov A., Karimov J., Dalenov Sh., Naimanov D., Zhazilbekov N. Animal Husbandry. Almaty, 2006 (in Kazakh)

Torekhanov A., Musabaev B., Tadzhiev K., Karymsakov T. Scientific Achievements in Animal Husbandry. Almaty, 2011 (in Russian)

Ulimbashev M., Gosteva E., Krasnova O., Konik N., Tletseruk I. Condition of meat breed resources of cattle in Russia (review). *Genetika i Razvedenie Zhivotnykh = Genetics and Breeding of Animals*. 2023; (1):52-64. DOI 10.31043/2410-2733-2023-1-52-64 (in Russian)

Yudin N.S., Larkin D.M. Whole genome studies of origin, selection and adaptation of the Russian cattle breeds. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2019; 23(5):559-568. DOI 10.18699/VJ19.525 (in Russian)

Yurchenko A., Yudin N., Aitnazarov R., Plyusnina A., Brukhin V., Soloshenko V., Lhasaranov B., Popov R., Paronyan I.A., Plemyashov K.V., Larkin D.M. Genome-wide genotyping uncovers genetic profiles and history of the Russian cattle breeds. *Heredity*. 2018a; 120(2):125-137. DOI 10.1038/s41437-017-0024-3

Yurchenko A.A., Daetwyler H.D., Yudin N., Schnabel R.D., Vander Jagt C.J., Soloshenko V., Lhasaranov B., Popov R., Taylor J.F., Larkin D.M. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Sci. Rep*. 2018b;8(1):12984. DOI 10.1038/s41598-018-31304-w

Zhumanov K., Baimukanov A. Dairy productivity of cows of the Holstein black-and-white cattle of the Kazakhstan population. *Rep. Natl. Acad. Sci. Republic Kazakhstan*. 2020;(6):109-114

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

Received March 21, 2024. Revised April 16, 2024. Accepted April 16, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):424-432

DOI 10.18699/vjgb-24-48

Polymorphic variants of the *hOGG1*, *APEX1*, *XPD*, *SOD2*, and *CAT* genes involved in DNA repair processes and antioxidant defense and their association with breast cancer risk

A.A. Timofeeva [1] ✉, V.I. Minina [1, 2], A.V. Torgunakova [1, 2], O.A. Soboleva [1], R.A. Titov [1, 2], Ya.A. Zakharova [1, 2], M.L. Bakanova [1], A.N. Glushkov [1]

[1] Federal Research Center of Coal and Coal Chemistry of the Siberian Branch of the Russian Academy of Sciences, Kemerovo, Russia
[2] Kemerovo State University, Kemerovo, Russia
✉ annateam86@gmail.com

Abstract. Breast cancer is one of the leading causes of mortality among women. The most frequently encountered tumors are luminal tumors. Associations of polymorphisms in the *hOGG1* (rs1052133), *APEX1* (rs1130409), *XPD* (rs13181), *SOD2* (rs4880), and *CAT* (rs1001179) genes were studied in 313 nonsmoking postmenopausal patients with luminal B subtype breast cancer. The control group consisted of 233 healthy nonsmoking postmenopausal women. Statistically significant associations of the *XPD* and *APEX1* gene polymorphisms with the risk of developing luminal B Her2-negative subtype of breast cancer were observed in a log-additive inheritance model, while the *CAT* gene polymorphism showed an association in a dominant inheritance model (OR = 1.41; CI 95 %: 1.08–1.85; Padj.= 0.011; OR = 1.39; CI 95 %: 1.07–1.81; Padj = 0.013 и OR = 1.70; CI 95 %: 1.19–2.43; Padj = 0.004, respectively). In the group of elderly women (aged 60–74 years), an association of the *CAT* gene polymorphism with the risk of developing luminal B subtype of breast cancer was found in a log-additive inheritance model (OR = 1.87; 95 % CI: 1.22–2.85; Padj = 0.0024). Using MDR analysis, the most optimal statistically significant 3-locus model of gene-gene interactions in the development of luminal B Her2-negative subtype breast cancer was found. MDR analysis also showed a close interaction and mutual enhancement of effects between the *APEX1* and *SOD2* loci and the independence of the effects of these loci from the *CAT* locus in the formation of luminal B subtype breast cancer.
Key words: breast cancer; luminal B subtype; *hOGG1*; *APEX1*; *XPD*; *SOD2*; *CAT*.

For citation: Timofeeva A.A., Minina V.I., Torgunakova A.V., Soboleva O.A., Titov R.A., Zakharova Ya.A., Bakanova M.L., Glushkov A.N. Polymorphic variants of the *hOGG1*, *APEX1*, *XPD*, *SOD2*, and *CAT* genes involved in DNA repair processes and antioxidant defense and their association with breast cancer risk. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):424-432. DOI 10.18699/vjgb-24-48

Funding. The work was carried out within the framework of the state assignment AAAA-A21-121011590009-9 "Immunohormonal Interactions in Breast Cancer" and with the use of grant funds for the establishment of a youth laboratory (Resolution of the Government of the Kemerovo Region No. 632 dated September 19, 2022).

Анализ ассоциаций полиморфных вариантов генов *hOGG1*, *APEX1*, *XPD*, *SOD2* и *CAT*, участвующих в процессах репарации ДНК и антиоксидантной защите, с риском развития рака молочной железы

А.А. Тимофеева [1] ✉, В.И. Минина [1, 2], А.В. Торгунакова [1, 2], О.А. Соболева [1], Р.А. Титов [1, 2], Я.А. Захарова [1, 2], М.Л. Баканова [1], А.Н. Глушков [1]

[1] Федеральный исследовательский центр угля и углехимии Сибирского отделения Российской академии наук, Кемерово, Россия
[2] Кемеровский государственный университет, Кемерово, Россия
✉ annateam86@gmail.com

Аннотация. Онкологические заболевания молочной железы – одна из ведущих причин смертности у женщин. Рак молочной железы относится к числу распространенных мультифакториальных полигенных заболеваний, реализующихся в результате сочетанного взаимодействия генетических и средовых факторов. Наиболее часто встречаются люминальные опухоли. Люминальный подтип В рака молочной железы характеризуется худшим прогнозом и ранними рецидивами. Для изучения генетических факторов риска развития злокачественных но-

вообразований молочной железы необходимо определить полиморфные варианты генов, играющих важную роль в канцерогенезе, к числу которых относятся гены репарации ДНК и системы антиоксидантной защиты. Изучены ассоциации полиморфизмов генов *hOGG1* (rs1052133), *APEX1* (rs1130409), *XPD* (rs13181), *SOD2* (rs4880) и *CAT* (rs1001179) у 313 некурящих пациенток в постменопаузе с диагнозом люминального подтипа В Her2-негативного рака молочной железы. В контрольную группу вошли 233 здоровые некурящие женщины в постменопаузе. Зарегистрированы с поправкой на возраст статистически значимые ассоциации полиморфных вариантов генов *XPD* (rs13181) и *APEX1* (rs1130409) с риском развития люминального подтипа В Her2-негативного рака молочной железы в лог-аддитивной модели наследования, гена *CAT* (rs1001179) – в доминантной модели OR = 1.41; CI 95 %: 1.08–1.85; Padj = 0.011; OR = 1.39; CI 95 %: 1.07–1.81; Padj = 0.013 и OR = 1.70; CI 95 %: 1.19–2.43; Padj = 0.004 соответственно). В группе женщин пожилого возраста (60–74 года) выявлена ассоциация вариантов гена *CAT* (rs1001179) с риском развития рака молочной железы в лог-аддитивной модели наследования (OR = 1.87; CI 95 %: 1.22–2.85; Padj = 0.0024). С помощью MDR-анализа найдена наиболее оптимальная статистически значимая 3-локусная модель межгенных взаимодействий при развитии онкозаболеваний молочной железы люминального подтипа В. MDR-анализ показал также тесное взаимодействие и взаимное усиление эффектов между локусами *APEX1* и *SOD2* и независимость эффектов данных локусов от эффекта локуса *CAT* при формировании люминального подтипа В рака молочной железы.

Ключевые слова: рак молочной железы; люминальный подтип В; *hOGG1*; *APEX1*; *XPD*; *SOD2*; *CAT*.

Introduction

Malignant transformations of the breast are the most wide spread oncological pathologies, by amount of deaths they take second place in world statistics (Siegel, 2021). Age, excess weight, heritage can be referred to as risk factors for oncopathology of breasts. Genetical, reproductive and hormonal factors can make a significant contribution to breast cancer. According to literature data, hormonal (luminal) malignancies are the most widespread (Ignatiadis, Sotiriou, 2013). Luminal B subtype of breast cancer, as opposed to luminal A subtype, is characterized by poor prognosis, early recurrence and high frequency of metastases in lymph nodes (Nishimura et al., 2010).

Breast cancer (BC) is a complex disorder with a high level of heterogeneity. The most well-studied markers of hereditary risk of BC are mutations in genes like *BRCA1/2*, *PALB2*, *TP53.* They influence the risk increase for BC more than twofold in comparison with the whole population. BC that is linked with germinal mutations in *BRCA1* has a triple negative phenotype (70–85 %), while ER-positive cases can be detected in carriers of mutations in the *BRCA2*, *ATM*, *CHEK2* and *PALB2* genes (Breast Cancer Association Consortium, 2021).

Meanwhile the majority of BC cases are sporadic (only from 5 to 10 % cases of BC are hereditary forms). There is a need for significant prognostic markers for sporadic forms of BC that can allow us to determine the group of risk to decrease mortality and morbidity.

Genome-wide association studies (GWASs) allowed to register over 170 loci of susceptibility for malignant breast transformation development, among them the biggest contribution can be made by single nucleotide polymorphisms (Michailidou et al., 2017; Ferreira et al., 2019). In Caucasian women, via GWAS, 32 loci associated with BC risk were identified. Five loci showed associations ($P < 0.05$) in the opposite direction between luminal and non-luminal subtypes of BC. *In silico* studies demonstrated that these five loci consist of cell-specific enhancers that differ in normal, luminal and basal cells of breasts (Zhang H. et al., 2020). A large number of variants detected by similar studies as a rule are located in regulatory non-coding regions, especially in distal enhancers and transcription factor binding sites (Pan et al., 2021).

Variants of DNA repair genes among different biomarkers are of greatest interest. DNA aberrations such as oxidative and reductive nitrogen bases, adducts and mutations induced by methylation agents can be recovered by enzymes of base excision repair (BER).

The *hOGG1* gene encodes a key enzyme of the BER pathway, bifunctional DNA-glycosilase/β-lyase, which excludes residues of 8-oxoguanine. The most well-studied and useful *hOGG1* polymorphic variant is rs1052133, which causes substitution of serine with cysteine in region 326 of the protein, decreasing the ability for repair activity (Niu et al., 2012). In a study using a BC cell line (HCC1937), it was shown that these cells are able to accumulate high levels of 8-oxoguanine in comparison with to normal glandular tissue (Nyaga et al., 2006).

Another gene of the BER pathway is *APEX1*, which encodes apurinic/apirimidinic endonuclease that can delete DNA sites with no nitrogen bases. *APEX1* rs1130409 polymorphic variant is linked to transversion of thymine to guanine in the 5th exon and causes substitution of asparagine acid with glutamine acid (Asp148Glu). It is associated with the ability of this enzyme to interact with other components of BER, thus decreasing the effectiveness of repair (Hadi et al., 2000).

Nucleotide excision repair (NER) plays a crucial role in stabilization of genome structure due to its ability to recover a high spectrum of DNA mutations (Sugasawa, 2010). One of the key components of this pathway is the *XPD* gene that encodes helicase, which participates in DNA unwinding and recognition of adducts and thymine dimers (Fontana et al., 2008). Substitution of adenine with cytosine in region 2251 of the gene (rs13181) promotes replacement of lysine by glutamine in region 751 of the protein, thus changing its configuration and causing interaction with helicase activator (Romaniuk et al., 2014).

Oxidative stress is one the most important factors in cancerogenesis caused by active forms of oxygen production that can affect DNA and initiate lipid peroxidation and modification of protein molecules (Caporaso, 2003; Tas et al., 2005). Effectiveness of autoxidation system performance is ensured by individual genetic properties. Catalase (CAT) and super-

oxide dismutase (SOD2) refer to proteins that can protect cells against oxidative stress (Ambrosone, 2000).

CAT is a key enzyme involved in neutralization of active oxygen forms via breakdown of hydrogen peroxide to water and oxygen (Ambrosone, 2000). Allele variants of this gene are associated with reduction of catalytic activity of this enzyme. rs1001179 is a well-studied polymorphic variant in the promoter region of the gene that can influence gene expression and cause a decrease in enzyme activity (Forsberg et al., 2001; Bastaki et al., 2006). A hypothesis about a link between estrogen exposure and catalase activity was made. It was shown that exposition of normal epithelial cells of human breasts to estradiol decreases the activity of cellular catalase (Forsberg et al., 2001).

Manganese-dependent superoxide dismutase works in the antioxidative system and is expressed in mitochondria. Transition of cytosine to thymine in the 47th region of the gene (rs4880) causes alanine-to-valine substitution in the 16th region of the protein and alteration of the secondary structure of the signal peptide. Destabilization of its alpha-helix domain decreases import of the protein from the cytoplasm to the mitochondria matrix causing enzyme absence. For *T* variant, mRNA instability is typical (Sutton et al., 2005). Association of this single nucleotide polymorphism with *SOD2* overexpression and accumulation of genotoxic oxygen peroxide has already been described (Ji et al., 2012).

Based on the above, the aim of this study was the analysis of association of loci *hOGG1* (rs1052133), *APEX1* (rs1130409), *XPD* (rs13181), *SOD2* (rs4880) and *CAT* (rs1001179) with BC development risk in women with luminal B Her2-negative subtype.

Materials and methods

Overall, 2,150 women with breast cancer that are Kemerovo region residents were observed. Inclusion criteria of patients in the study were as follows: Caucasian, female, age over 40, postmenopausal, previously diagnosed with luminal B Her2-negative BC, absence of family forms of oncopathology. Exclusion criteria were: smoking, oncopathology forms in anamnesis, relatives with oncopathology.

313 non-smoking women were selected from the whole sample of patients (median age 60.88±0.35), 42.04 % had the I stage of disease, 42.04 % had the II stage, 13.38 and 2.55 % patients were diagnosed with the III and IV stages of BC, respectively. Metastases in lymph nodes and/or in distal organs were observed in 51 women. All patients were observed by medicals of Kuzbass Clinical Oncological Dispensary using a whole complex of diagnostics methods, after that it became possible to make a certain pathomorphological diagnosis for each woman. Classification of subtypes was based on expressional parameters of estrogen (ER) and progesterone (PR) receptors and also those of receptor tyrosine kinase (Her2) and level of proliferative activity of Ki-67 (Goldhirsch et al., 2013).

233 Kemerovo region residents were included into the control group without any symptoms of oncological disorders (median age 58.44±0.34). Inclusion criteria in the control group were: Caucasian, female, age over 40 years, postmenopausal. Exclusion criteria were: smoking, oncological cases in anamnesis, relatives with oncopathology. Age characteristics of the observed groups (according to the WHO recommendations of 2016) are presented in Table 1.

Table 1. Characteristics of the comparison groups

Age, years	Patients, *N* (%)	Controls, *N* (%)
45–59	119 (36.39)	137 (58.80)
60–74	194 (59.33)	96 (41.20)

This study was approved by the ethics committee of the Federal Research Center of Coal and Coal Chemistry of SB RAS according to the statements of the Helsinki declaration (ratified in 2000). Collection of data and samples of peripheral blood was conducted after receiving voluntary informed consent from patients and healthy individuals.

DNA was purified from peripheral blood via the standard method of phenol-chloroform extraction (Sambrook et al., 1989). Variants of the *hOGG1* (rs1052133), *APEX1* (rs1130409), *XPD* (rs13181), and *CAT* (rs1001179), *SOD2* (rs4880) genes were genotyped by real-time PCR using TaqMan primers from SibDNA kits (SibDNA, Novosibirsk, Russia). Amplification and detection of the results were performed using the CFX96 amplificator (BioRad, USA).

SNPStats (http://bioinfo.iconcologia.net/SNPstats) and STATISTICA 10.0 (StatSoft Inc., Tulsa, Oklahoma, USA) programs were used for statistical processing of the obtained results. Analysis of rare allele frequency, accordance to Hardy–Weinberg equilibrium were provided by available online sources (https://gene-calc.pl/hardy-weinberg-page and http://www.quantpsy.org/chisq/chisq.htm, respectively). Statistically significant results were accepted with $p < 0.05$. For minimization of type I statistical error, multiple comparisons problem was used. Using age parameter, we performed a logistic regression analysis with odds ratio (OR) calculation (with 95 % confidence interval). The most convenient statistical model with the lowest value was selected using Akaike Information Criteria (AIC). With Multifactor Dimensionality Reduction (MDR), which allows to evaluate all possible models of SNP combinations, we investigated intergenic interactions. Contribution of each gene and/or their interactions were evaluated by *H*-parameter (caused by entropy) and represented as a percentage (%) (Moore et al., 2006). To perform this analysis, the program package of MDR 3.2.0 was used (Computational Genetics Laboratory, Philadelphia, Pennsylvania, USA).

Results

Investigation of the *hOGG1*, *APEX1*, *XPD*, *SOD2* and *CAT* genes polymorphic variants was conducted in cohorts of non-smoking women with luminal B subtype of BC and healthy women of similar age (Table 2).

Distribution of alleles and genotypes in the studied groups corresponds to Hardy–Weinberg equilibrium and to parameters observed in Caucasian population (http://www.ensembl.org/Homo_sapiens). No statically significant differences were detected between different groups of patients (malignancy

Table 2. Distribution of DNA repair and antioxidant system genes polymorphic variants in the study groups

Loci	Genotypes and alleles	BC, *N* (%)	Controls, *N* (%)	*P* (*df*)**
XPD c.2251A>C, *p.K751Q* (rs13181)	*AA*	125 (39.94)	118 (50.64)	0.06 (2)/0.05 (1)
	AC	152 (48.56)	95 (40.77)	
	CC	36 (11.50)	20 (8.58)	
	A	201 (64.22)	166 (71.03)	
	C	118(35.78)	67 (28.97)	
	p^{HWE}*	0.39	0.87	
APEX1 c.444T>G, *p.D148E* (rs1130409)	*TT*	107 (34.19)	96 (41.20)	0.10 (2)/0.16 (1)
	TG	157 (50.16)	114 (48.93)	
	GG	49 (15.65)	23 (9.87)	
	T	186 (59.27)	153 (65.67)	
	G	127(40.73)	80 (34.33)	
	p^{HWE}*	0.56	0.24	
hOGG1 c.977C>G, *p.S326C* (rs1052133)	*CC*	185 (59.11)	142 (60.94)	0.28 (2)/0.97(1)
	CG	118 (37.70)	77 (33.05)	
	GG	10 (3.19)	14 (6.01)	
	C	244 (77.96)	181 (77.47)	
	G	69 (22.04)	52 (22.53)	
	p^{HWE}*	0.10	0.45	
CAT g.4760 C>G (rs1001179)	*CC*	168 (53.67)	151 (64.81)	0.045 (2)/0.07 (1)
	CG	119 (38.02)	69 (29.61)	
	GG	26 (8.31)	13 (5.58)	
	C	228 (72.68)	186 (79.62)	
	G	85 (27.32)	47 (20.38)	
	p^{HWE}*	0.48	0.22	
SOD2 c.47T>C, *p.A16V* (rs4880)	*TT*	84 (26.84)	65 (27.90)	0.24 (2)/0.41 (1)
	TC	147 (46.96)	122 (52.36)	
	CC	82 (26.20)	46 (19.74)	
	T	157 (50.32)	126 (54.08)	
	C	156 (49.68)	107 (45.92)	
	p^{HWE}*	0.30	0.71	

* Accordance to Hardy–Weinberg equilibrium (HWE); ** level of significance after comparison of alleles and genotypes frequency in the study groups.

stage, its localization, metastases development). Significant differences between genotypes and alleles distribution in DNA repair and antioxidant system genes, taking into account the Bonferroni correction were not detected in study groups.

Analysis of different hereditary models with correction for age allowed to detect association between the risk of luminal B Her2-negative BC development and *XPD* (rs13181) and *APEX1* (rs1130409) in the log-additive model, and *CAT* (rs1001179) in the dominant model (OR = 1.41; CI 95 %: 1.08–1.85; Padj = 0.011; OR = 1.39; CI 95 %: 1.07–1.81; Padj = 0.013 and OR = 1.70; CI 95 %: 1.19–2.43; Padj = 0.004 respectively).

Distribution of genotypes and alleles of the studied genes in different age groups of patients with BC and healthy women is presented in Table 3.

Analysis of different hereditary models allowed to reveal links between polymorphic variants of the *CAT* (rs1001179) gene with the risk of luminal B Her2-negative BC develop-

Table 3. Distribution of different variants of DNA repair and antioxidative system genes in different study groups

Age	Loci	Genotypes and alleles	BC, *N* (%)	Controls, *N* (%)	*P* (*df*)*
45–59	*XPD c.2251A>C, p.K751Q* (rs13181)	*AA/AC/CC*	49 (41.18)/52 (43.70)/18 (15.12)	68 (49.64)/56 (40.87)/13 (9.49)	0.35
		A	75 (63.03)	96 (70.08)	0.29
		C	44 (36.97)	41 (29.02)	
	APEX1 c.444T>G, p.D148E (rs1130409)	*TT/TG/GG*	41 (34.45)/58 (48.74)/20 (16.81)	53 (38.69)/67 (48.91)/17 (12.40)	0.68
		T	70 (58.82)	87 (63.15)	0.52
		G	49 (41.18)	50 (36.85)	
	hOGG1 c.977C>G, p.S326C (rs1052133)	*CC/CG/GG*	75 (63.03)/40 (33.61)/4 (3.36)	88 (64.23)/42 (30.66)/7 (5.11)	0.89
		C	95 (79.84)	109 (79.56)	0.92
		G	24 (20.16)	28 (20.44)	
	CAT g.4760 C>G (rs1001179)	*CC/CG/GG*	64 (53.78)/50 (42.02)/ 5 (4.20)	83 (60.58)/45 (32.85)/9 (6.57)	0.39
		C	89 (74.79)	106 (77.01)	0.74
		G	30 (25.21)	31 (22.99)	
	SOD2 c.47T>C, p.A16V (rs4880)	*TT/TC/CC*	29 (24.37)/61 (51.26)/29 (24.37)	30 (21.90)/69 (50.36)/38 (27.74)	0.89
		T	59 (50.00)	64 (47.08)	0.74
		C	60 (50.00)	73 (52.92)	
60–74	*XPD c.2251A>C, p.K751Q* (rs13181)	*TT/TG/GG*	76 (39.18)/100 (51.55)/ 18 (9.27)	50 (52.08)/39 (40.63)/7 (7.29)	0.16
		T	126 (64.96)	70 (72.40)	0.22
		G	68 (35.04)	26 (28.60)	
	APEX1 c.444T>G, p.D148E (rs1130409)	*TT/TG/GG*	66 (34.02)/99 (51.03)/ 29 (14.95)	43 (44.79)/47 (48.96)/ 6 (6.25)	0.08
		T	116 (59.54)	67 (69.27)	0.13
		G	78 (40.46)	29 (30.73)	
	hOGG1 c.977C>G, p.S326C (rs1052133)	*CC/CG/GG*	110 (56.70)/78 (40.21)/ 6 (3.09)	54 (56.25)/35 (36.46)/ 7 (7.29)	0.40
		C	149 (76.81)	72 (74.48)	0.85
		G	45 (23.19)	24 (25.52)	
	CAT g.4760C>G (rs1001179)	*CC/CG/GG*	104 (53.61)/69 (35.57)/ 21 (10.82)	68 (70.83)/24 (25.00)/ 4 (4.16)	0.02
		C	139 (71.40)	80 (83.33)	0.04
		G	55 (28.60)	16 (16.67)	
		Log-additive model (OR = 1.87, CI 95 % 1.22–2.85, Padj = 0.0024)			
	SOD2 c.47T>C, p.A16V (rs4880)	*TT/TC/CC*	56 (28.87)/85 (43.81)/53 (27.32)	13 (13.54)/55 (57.29)/28 (29.17)	0.02
		T	98 (50.77)	40 (42.19)	0.20
		C	96 (49.23)	56 (57.81)	

* Level of significance in comparison of alleles and genotypes distribution between different study groups.

ment in elder patients (60–74 years) in the log-additive model (OR = 1.87; CI 95 %: 1.22–2,85; Padj = 0.0024).

Via the MDR method, the most optimal 3-loci model of intergenic interactions with a high level of precision, minimal rate of error for BC risk prediction and maximal level of reproducibility evaluation was obtained (Table 4).

Analysis of the model in contingency tables, which represent all possible variants for the 3-loci model, revealed 12 protective and 15 risk combinations for luminal B Her2-negative BC development (Fig. 1).

The MDR analysis showed a simultaneous strengthening of effects between two loci, *APEX1* (rs1130409) (*H* = 0.07 %)

Table 4. Significant intergenic interactions during BC development

Loci	Tr.Bal.Acc.	Test.Bal.Acc.	Sign Test (*p*)	Se.	Sp.	CVC	Pre.
CAT (rs1001179) * *APEX1* (rs1130409) * *SOD2* (rs4880)	0.616	0.557	< 0.0001	0.473	0.752	10/10	0.799

Note. Tr.Bal.Acc. – training balanced accuracy; Test.Bal.Acc. – testing balanced accuracy; Sign Test (*p*) – test for significance; Se. – sensitivity; Sp. – specificity; CVC – repeatability of the result; Pre. – precision of the model

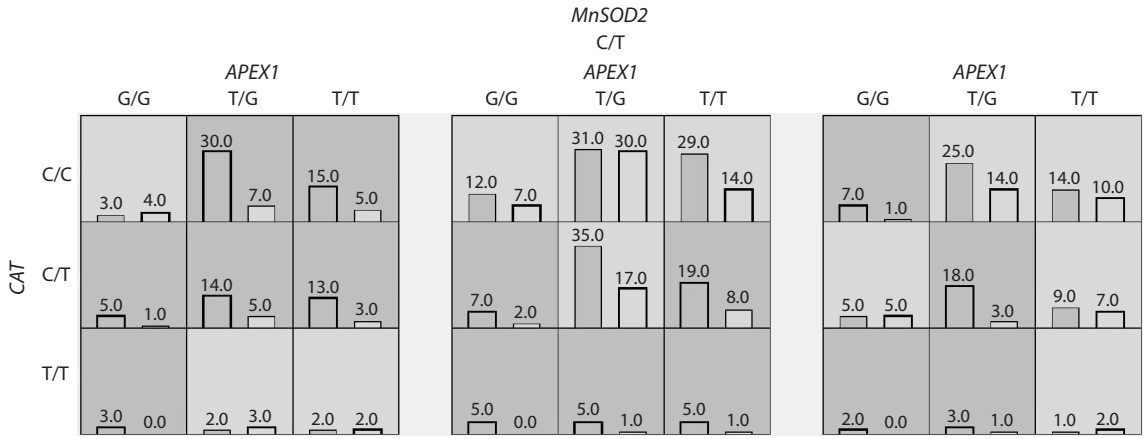


Fig. 1. Combination of genotypes for the 3-loci model of *CAT* (rs1001179), *APEX1* (rs1130409) and *SOD2* (rs4880) that can predispose to the risk of luminal B Her2-negative BC development.

Dark grey cells – genotypes of increased risk, light grey cells – genotypes of decreased risk (left columns in the cells – patients with BC, right columns – healthy women).

and *SOD2* (rs4880) (H = 0.55 %), and also independence of their effects from *CAT* (rs1001179) (H = 0.44 %) during formation of luminal B Her2-negative BC (Fig. 2).

Fig. 2. Dendrogram of intergenic interactions during formation of luminal B Her2-negative BC.

Red – synergy of effects, brown – independent interaction.

Discussion

Sensitivity of an organism to air pollutants depends on the correct work of many enzyme systems, which include DNA repair and the antioxidative system. The level of breast tissues exposition to exo- and endogenous estrogens (providing DNA adducts formation) makes a big contribution to disease pathogenesis (Martucci, Fishman, 1993; Hanawalt, 2002). Estrogens are involved in regulation of antioxidative enzymes and can initiate oxidative mutations in DNA due to formation of active forms of oxygen during metabolic reactions (Tjønneland et al., 2004; Bergman et al., 2005; Silva et al., 2006; Liou, Storz,2010).

In one of the articles, influence of obesity on BC risk in female carriers of at least one minor allele of myeloperoxidase gene or DNA repair genes like *GMT*, *MSH2*, *XPG* and *XRCC1* was detected (McCullough et al., 2015). In another study, it was shown that genes involved in oxidative stress and DNA repair can increase survival of women affected by breast oncopathologies (Rodrigues et al., 2012). At the same time there were no scientific works aimed at synergetic influence of DNA excision repair genes with genes of the antioxidative system on BC risk.

DNA aberrations that are formed due to active forms of oxygen can be recovered via the BER and NER pathways. Results obtained in our work concerning *APEX1* (rs1130409) association with BC risk are consistent with literature data (Mitra et al., 2008; Smith et al., 2008; Kim et al., 2013). Additionally, a link between the *444T* allele and estrogen-positive BC development was revealed in Chinese women (Wang T. et al., 2018). Besides the repair function, this enzyme can perform oxidative-reductive activity of transcriptional factors (Kelley et al., 2012; Wang Z. et al., 2014). Redox activity of the protein contributes to synergy between the *APEX1* (rs1130409) and *SOD2 47* (rs4880) loci during BC formation.

The *hOGG1* gene is another key component of the BER pathway. In our study, no links were found between *hOGG1* (rs1052133) and BC risk. Similar results were demonstrated in the meta-analysis by M. Kamali et al. (2017), where association of *977G* with BC wasn't revealed in Caucasian as well as in Asian women (Kamali et al., 2017). At the same time, in a scientific work performed among Polish patients, *hOGG1 977GG* genotype contributed to the risk of BC development (Romanowicz et al., 2017).

Results of scientific studies that are aimed at *XPD* (rs13181) association with oncological disorders of the breast are not obvious. In works that were conducted using material of Ca-

nadian, Brazilian and Chinese women no significant results were obtained (Dufloth et al., 2005; Zhang L. et al., 2005; Onay et al., 2006). Observation of Indian patients allowed to reveal association of the *2251C* allele with enhanced risk of BC (Samson et al., 2011). Later, a meta-analysis was conducted that showed an increased risk of BC in *2251C* allele carriers in Caucasian and mixed populations (Yan et al., 2014). Similar results were demonstrated using Polish patients (Smolarz et al., 2019).

Manganese-dependent superoxide dismutase is one of the most important enzymes of the antioxidative system. Besides its own essential function (antioxidative activity), SOD2 protein has binding sites with different factors of transcription that are useful for its activation and are also involved in defense of cells against oxidative stress (Alateyah et al., 2022). Results of molecular and genetical studies of *SOD2* (rs4880) association with BC risk are quite controversial. In our study, no influence of this polymorphic variant on the risk of malignant transformation development in breasts was detected. Similar results were obtained in the works conducted among Polish and Greek women (Jablonska et al., 2015; Kakkoura et al., 2016). In Mexican female patients, an association between the *47T* allele of the *SOD2* gene and luminal A subtype formation was detected, but not with luminal B (Gallegos-Arreola et al., 2022). In Iraqi and Taiwan, an association between this allele and increased BC risk was also detected (Tsai et al., 2012; Jabir, Hoidy, 2018).

Results of studies aimed to link the *CAT* (rs1001179) gene polymorphism with BC risk are still controversial. In some scientific works among American patients, an association between a decreased risk of BC and the *-262 CC* genotype was revealed in comparison with *T* allele carriers (Ahn et al., 2004, 2005). In our study, we got similar results. Ambiguous data were obtained by Y. Li et al. (2009), who registered a small decrease in BC risk in postmenopausal women with the *CAT -262 CC* genotype that were consumed a huge number of fruits and vegetables (over two portions a day). Among women with a small rate of fruits and vegetables consumption, *CAT -262 CC* was linked with an increased risk of BC (Li et al., 2009).

Conclusion

The combined influence of DNA repair and antioxidative system genes variants on breast cancer risk was demonstrated. This work was conducted using material of postmenopausal women; to better understand the influence of individual genetical features on breast cancer development, it is also advisable to include younger women in experimental study.

To clarify the ability of the system of risk prognosis for BC risk evaluation, it is necessary to increase the number of studied patients to perform an additional study.

References

Ahn J., Gammon M.D., Santella R.M., Gaudet M.M., Britton J.A., Teitelbaum S.L., Terry M.B., Neugut A.I., Josephy P.D., Ambrosone C.B. Myeloperoxidase genotype, fruit and vegetable consumption, and breast cancer risk. *Cancer Res*. 2004;64(20):7634-7639. DOI 10.1158/0008-5472.CAN-04-1843

Ahn J., Gammon M.D., Santella R.M., Gaudet M.M., Britton J.A., Teitelbaum S.L., Terry M.B., Nowell S., Davis W., Garza C., Neugut A.I., Ambrosone C.B. Associations between breast cancer risk and the catalase genotype, fruit and vegetable consumption, and supplement use. *Am. J. Epidemiol*. 2005;162(10):943-952. DOI 10.1093/aje/kwi306

Alateyah N., Gupta I., Rusyniak R.S., Ouhtit A. *SOD2*, a potential transcriptional target underpinning CD44-promoted breast cancer progression. *Molecules*. 2022;27(3):811. DOI 10.3390/molecules27030811

Ambrosone C.B. Oxidants and antioxidants in breast cancer. *Antioxid. Redox Signal*. 2000;2(4):903-917. DOI 10.1089/ars.2000.2.4-903

Bastaki M., Huen K., Manzanillo P., Chande N., Chen C., Balmes J.R., Tager I.B., Holland N. Genotype-activity relationship for Mn-superoxide dismutase, glutathione peroxidase 1 and catalase in humans. *Pharmacogenet. Genomics*. 2006;16(4):279-286. DOI 10.1097/01.fpc.0000199498.08725.9c

Bergman M., Ahnström M., Palmebäck Wegman P., Wingren S. Polymorphism in the manganese superoxide dismutase (MnSOD) gene and risk of breast cancer in young women. *J. Cancer Res. Clin. Oncol*. 2005;131(7):439-444. DOI 10.1007/s00432-004-0663-7

Breast Cancer Association Consortium. Breast cancer risk genes – association analysis in more than 113,000 women. *N. Engl. J. Med.* 2021;384(5):428-439. DOI 10.1056/NEJMoa1913948

Calculation for the Chi-Square Test [Electronic resource]. URL: http://www.quantpsy.org/chisq/chisq.htm (accessed: 06.07.2023)

Caporaso N. The molecular epidemiology of oxidative damage to DNA and cancer. *J. Natl. Cancer Inst.* 2003;95(17):1263-1265. DOI 10.1093/jnci/djg065

Dufloth R.M., Costa S., Schmitt F., Zeferino L.C. DNA repair gene polymorphisms and susceptibility to familial breast cancer in a group of patients from Campinas, Brazil. *Genet. Mol. Res.* 2005;4(4):771-782

Ensembl [Electronic resource]. URL: http://www.ensembl.org/Homo_sapiens (accessed: 06.07.2023)

Ferreira M.A., Gamazon E.R., Al-Ejeh F., Aittomäki K., Andrulis I.L., Anton-Culver H., Arason A., Arndt V., Aronson K.J., Arun B.K., ... Yang X.R., Yannoukakos D., Ziogas A., Kraft P., Antoniou A.C., Zheng W., Easton D.F., Milne R.L., Beesley J., Chenevix-Trench G. Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. *Nat. Commun.* 2019;10(1):1741. DOI 10.1038/s41467-018-08053-5

Fontana L., Bosviel R., Delort L., Guy L., Chalabi N., Kwiatkowski F., Satih S., Rabiau N., Boiteux J.P., Chamoux A., Bignon Y.J., Bernard-Gallon D.J. DNA repair gene *ERCC2*, *XPC*, *XRCC1*, *XRCC3* polymorphisms and associations with bladder cancer risk in a French cohort. *Anticancer Res*. 2008;28(3B):1853-1856

Forsberg L., Lyrenäs L., de Faire U., Morgenstern R. A common functional C-T substitution polymorphism in the promoter region of the human catalase gene influences transcription factor binding, reporter gene transcription and is correlated to blood catalase levels. *Free Radic. Biol. Med*. 2001;30(5):500-505. DOI 10.1016/s0891-5849(00)00487-1

Gallegos-Arreola M.P., Ramírez-Patiño R., Sánchez-López J.Y., Zúñiga-González G.M., Figuera L.E., Delgado-Saucedo J.I., Gómez-Meda B.C., Rosales-Reynoso M.A., Puebla-Pérez A.M., Lemus-Varela M.L., Garibaldi-Ríos A.F., Marín-Domínguez N.A., Pacheco-Verduzco D.P., Mohamed-Flores E.A. SOD2 gene variants (rs4880 and rs5746136) and their association with breast cancer risk. *Curr. Issues Mol. Biol*. 2022;44(11):5221-5233. DOI 10.3390/cimb44110355

Goldhirsch A., Winer E.P., Coates A.S., Gelber R.D., Piccart-Gebhart M., Thürlimann B., Senn Y.-J. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann. Oncol*. 2013;24(9):2206-2223. DOI 10.1093/annonc/mdt303

Hadi M.Z., Coleman M.A., Fidelis K., Mohrenweiser H.W., Wilson D.M. 3rd. Functional characterization of Ape1 variants identified in the human population. *Nucleic Acids Res.* 2000;28(20):3871-3879. DOI 10.1093/nar/28.20.3871

Hanawalt P.C. Subpathways of nucleotide excision repair and their regulation. *Oncogene.* 2002;21(58):8949-8956. DOI 10.1038/sj.onc.1206096

Hardy-Weinberg equilibrium [Electronic resource]. URL: https://gene-calc.pl/hardy-weinberg-page (accessed: 06.07.2023)

Ignatiadis M., Sotiriou C. Luminal breast cancer: from biology to treatment. *Nat. Rev. Clin. Oncol.* 2013;10(9):494-506. DOI 10.1038/nrclinonc.2013.124

Jabir F.A., Hoidy W.H. Pharmacogenetics as personalized medicine: association investigation of SOD2 rs4880, CYP2C19 rs4244285, and FCGR2A rs1801274 polymorphisms in a breast cancer population in Iraqi women. *Clin. Breast Cancer.* 2018;18(5):e863-e868. DOI 10.1016/j.clbc.2018.01.009

Jablonska E., Gromadzinska J., Peplonska B., Fendler W., Reszka E., Krol M.B., Wieczorek E., Bukowska A., Gresner P., Galicki M., Zambrano Quispe O., Morawiec Z., Wasowicz W. Lipid peroxidation and glutathione peroxidase activity relationship in breast cancer depends on functional polymorphism of *GPX1*. *BMC Cancer*. 2015;15:657. DOI 10.1186/s12885-015-1680-4

Ji M., Tang J., Zhao J., Xu B., Qin J., Lu J. Polymorphisms in genes involved in drug detoxification and clinical outcomes of anthracycline-based neoadjuvant chemotherapy in Chinese Han breast cancer patients. *Cancer Biol. Ther*. 2012;13(5):264-271. DOI 10.4161/cbt.18920

Kakkoura M.G., Demetriou C.A., Loizidou M.A., Loucaides G., Neophytou I., Malas S., Kyriacou K., Hadjisavvas A. MnSOD and CAT polymorphisms modulate the effect of the Mediterranean diet on breast cancer risk among Greek-Cypriot women. *Eur. J. Nutr.* 2016;55(4):1535-1544. DOI 10.1007/s00394-015-0971-5

Kamali M., Kargar S., Heiranizadeh N., Zare M., Kargar Sh., Zare Shehneh M., Neamatzadeh H. Lack of any association between the Hogg1 Ser326Cys polymorphism and breast cancer risk: a systematic review and meta-analysis of 18 studies. *Asian Pac. J. Cancer Prev*. 2017;18(1):245-251. DOI 10.22034/APJCP.2017.18.1.245

Kelley M.R., Georgiadis M.M., Fishel M.L. APE1/Ref-1 role in redox signaling: translational applications of targeting the redox function of the DNA repair/redox protein APE1/Ref-1. *Curr. Mol. Pharmacol*. 2012;5(1):36-53. DOI 10.2174/1874467211205010036

Kim K.Y., Han W., Noh D.Y., Kang D., Kwack K. Impact of genetic polymorphisms in base excision repair genes on the risk of breast cancer in a Korean population. *Gene.* 2013;532(2):192-196. DOI 10.1016/j.gene.2013.09.069

Li Y., Ambrosone C.B., McCullough M.J., Ahn J., Stevens V.L., Thun M.J., Hong C.C. Oxidative stress-related genotypes, fruit and vegetable consumption and breast cancer risk. *Carcinogenesis.* 2009;30(5):777-784. DOI 10.1093/carcin/bgp053

Liou G.Y., Storz P. Reactive oxygen species in cancer. *Free Radic. Res*. 2010;44(5):479-496. DOI 10.3109/10715761003667554

Martucci C.P., Fishman J. P450 enzymes of estrogen metabolism. *Pharmacol. Ther.* 1993;57(2-3):237-257. DOI 10.1016/0163-7258(93)90057-k

McCullough L.E., Eng S.M., Bradshaw P.T., Cleveland R.J., Steck S.E., Terry M.B., Shen J., Crew K.D., Rossner P. Jr., Ahn J., Ambrosone C.B., Teitelbaum S.L., Neugut A.I., Santella R.M., Gammon M.D. Genetic polymorphisms in DNA repair and oxidative stress pathways may modify the association between body size and postmenopausal breast cancer. *Ann. Epidemiol.* 2015;25(4):263-269. DOI 10.1016/j.annepidem.2015.01.009

Michailidou K., Lindström S., Dennis J., Beesley J., Hui S., Kar S., Lemaçon A., Soucy P., Glubb D., Rostamianfar A., … García-Closas M., Schmidt M.K., Chanock S.J., Dunning A.M., Edwards S.L., Bader G.D., Chenevix-Trench G., Simard J., Kraft P., Easton D.F. Association analysis identifies 65 new breast cancer risk loci. *Nature.* 2017;551(7678):92-94. DOI 10.1038/nature24284

Mitra A.K., Singh N., Singh A., Garg V.K., Agarwal A., Sharma M., Chaturvedi R., Rath S.K. Association of polymorphisms in base excision repair genes with the risk of breast cancer: a case-control study in North Indian women. *Oncol. Res*. 2008;17(3):127-135. DOI 10.3727/096504008785055567

Moore J.H., Gilbert J.C., Tsai C.T., Chiang F.T., Holden T., Barney N., White B.C. A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J. Theor. Biol*. 2006;241(2):252-261. DOI 10.1016/j.jtbi.2005.11.036

Nishimura R., Osako T., Okumura Y., Hayashi M., Toyozumi Y., Arima N. Ki-67 as a prognostic marker according to breast cancer subtype and a predictor of recurrence time in primary breast cancer. *Exp. Ther. Med.* 2010;1(5):747-754. DOI 10.3892/etm.2010.133

Niu Y., Li F., Tang B., Shi Y., Yu P. Association of hOGG1 Ser326Cys polymorphism with gastric cancer risk: a meta-analysis. *Mol. Biol. Rep*. 2012;39(6):6563-6568. DOI 10.1007/s11033-012-1485-3

Nyaga S.G., Lohani A., Jaruga P., Trzeciak A.R., Dizdaroglu M., Evans M.K. Reduced repair of 8-hydroxyguanine in the human breast cancer cell line, HCC1937. *BMC Cancer*. 2006;6:297. DOI 10.1186/1471-2407-6-297

Onay V.U., Briollais L., Knight J.A., Shi E., Wang Y., Wells S., Li H., Rajendram I., Andrulis I.L., Ozcelik H. SNP-SNP interactions in breast cancer susceptibility. *BMC Cancer*. 2006;6:114. DOI 10.1186/1471-2407-6-114

Pan Q., Liu Y.J., Bai X.F., Han X.L., Jiang Y., Ai B., Shi S.S., Wang F., Xu M.C., Wang Y.Z., Zhao J., Chen J.X., Zhang J., Li X.C., Zhu J., Zhang G.R., Wang Q.Y., Li C.Q. VARAdb: a comprehensive variation annotation database for human. *Nucleic Acids Res*. 2021;49(D1):D1431-D1444. DOI 10.1093/nar/gkaa922

Rodrigues P., Furriol J., Bermejo B., Chaves F.J., Lluch A., Eroles P. Identification of candidate polymorphisms on stress oxidative and DNA damage repair genes related with clinical outcome in breast cancer patients. *Int. J. Mol. Sci.* 2012;13(12):16500-16513. DOI 10.3390/ijms131216500

Romanowicz H., Pyziak Ł., Jabłoński F., Bryś M., Forma E., Smolarz B. Analysis of DNA repair genes polymorphisms in breast cancer. *Pathol. Oncol. Res*. 2017;23(1):117-123. DOI 10.1007/s12253-016-0110-5

Romaniuk O.P., Nikitchenko N.V., Savina N.V., Kuzhir T.D., Goncharova R.I. The polymorphism of DNA repair genes *XPD*, *XRCC1*, *OGG1*, and *ERCC6*, life expectancy, and the inclination to smoke. *Russ. J. Genet*. 2014;50(8):860-869. DOI 10.1134/S1022795414080067

Sambrook J., Fritsch E.R., Maniatis T. Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press, 1989

Samson M., Singh S.S., Rama R., Sridevi V., Rajkumar T. XPD Lys751Gln increases the risk of breast cancer. *Oncol. Lett*. 2011;2(1):155-159. DOI 10.3892/ol.2010.220

Siegel R.L., Miller K.D., Fuchs H.E., Jemal A. Cancer statistics, 2021. *CA Cancer J. Clin*. 2021;71(1):7-33. DOI 10.3322/caac.21654

Silva S.N., Cabral M.N., Bezerra de Castro G., Pires M., Azevedo A.P., Manita I., Pina J.E., Rueff J., Gaspar J. Breast cancer risk and polymorphisms in genes involved in metabolism of estrogens (CYP17, HSD17beta1, COMT and MnSOD): possible protective role of MnSOD gene polymorphism Val/Ala and Ala/Ala in women that never breast fed. *Oncol. Rep.* 2006;16(4):781-788

Smith T.R., Levine E.A., Freimanis R.I., Akman S.A., Allen G.O., Hoang K.N., Liu-Mares W., Hu J.J. Polygenic model of DNA repair genetic polymorphisms in human breast cancer risk. *Carcinogenesis.* 2008;29(11):2132-2138. DOI 10.1093/carcin/bgn193

Smolarz B., Michalska M.M., Samulak D., Romanowicz H., Wójcik L. Polymorphism of DNA repair genes in breast cancer. *Oncotarget.* 2019;10(4):527-535. DOI 10.18632/oncotarget.26568

SNPStats [Electronic resource]. URL: http://bioinfo.iconcologia.net/SNPstats (accessed: 06.07.2023)

Sugasawa K. Regulation of damage recognition in mammalian global genomic nucleotide excision repair. *Mutat. Res.* 2010;685(1-2):29-37. DOI 10.1016/j.mrfmmm.2009.08.004

Sutton A., Imbert A., Igoudjil A., Descatoire V., Cazanave S., Pessayre D., Degoul F. The manganese superoxide dismutase Ala16Val dimorphism modulates both mitochondrial import and mRNA stability. *Pharmacogenet. Genomics.* 2005;15(5):311-319. DOI 10.1097/01213011-200505000-00006

Tas F., Hansel H., Belce A., Ilvan S., Argon A., Camlica H., Topuz E. Oxidative stress in breast cancer. *Med. Oncol.* 2005;22(1):11-15. DOI 10.1385/MO:22:1:011

Tjønneland A., Christensen J., Thomsen B.L., Olsen A., Overvad K., Ewertz M., Mellemkjaer L. Hormone replacement therapy in relation to breast carcinoma incidence rate ratios: a prospective Danish cohort study. *Cancer*. 2004; 100(11):2328-2337. DOI 10.1002/cncr.20250

Tsai S.M., Wu S.H., Hou M.F., Chen Y.L., Ma H., Tsai L.Y. Oxidative stress-related enzyme gene polymorphisms and susceptibility to breast cancer in non-smoking, non-alcohol-consuming Taiwanese women: a case-control study. *Ann. Clin. Biochem.* 2012;49(Pt. 2):152-158. DOI 10.1258/acb.2011.011098

Wang T., Wang H., Yang S., Guo H., Zhang B., Guo H., Wang L., Zhu G., Zhang Y., Zhou H., Zhang X., Li H., Su H. Association of *APEX1* and *OGG1* gene polymorphisms with breast cancer risk among Han women in the Gansu Province of China. *BMC Med. Genet.* 2018;19(1):67. DOI 10.1186/s12881-018-0578-9

Wang Z., Ayoub E., Mazouzi A., Grin I., Ishchenko A.A., Fan J., Yang X., Harihar T., Saparbaev M., Ramotar D. Functional variants of human APE1 rescue the DNA repair defects of the yeast AP endonuclease/3′-diesterase-deficient strain. *DNA Repair.* 2014;22:53-66. DOI 10.1016/j.dnarep.2014.07.010

World Health Organization. [Electronic resource]. URL: https://www.who.int/ru (accessed: 03.10.2023)

Yan Y., Liang H., Light M., Li T., Deng Y., Li M., Li S., Qin X. *XPD* Asp312Asn and Lys751Gln polymorphisms and breast cancer susceptibility: a meta-analysis. *Tumour Biol.* 2014;35(3):1907-1915. DOI 10.1007/s13277-013-1256-3

Zhang H., Ahearn T.U., Lecarpentier J., Barnes D., Beesley J., Qi G., Jiang X., O'Mara T.A., Zhao N., Bolla M.K., … Kraft P., Simard J., Pharoah P.D.P., Michailidou K., Antoniou A.C., Schmidt M.K., Chenevix-Trench G., Easton D.F., Chatterjee N., García-Closas M. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* 2020;52(6):572-581. DOI 10.1038/s41588-020-0609-2

Zhang L., Zhang Z., Yan W. Single nucleotide polymorphisms for DNA repair genes in breast cancer patients. *Clin. Chim. Acta.* 2005;359(1-2):150-155. DOI 10.1016/j.cccn.2005.03.047

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

Received August 3, 2023. Revised January 12, 2024. Accepted February 26, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):433-442

DOI 10.18699/vjgb-24-49

Polymorphism of angiogenesis regulation factor genes (*VEGF/VEGFR*), and extracellular matrix remodeling genes (*MMP/TIMP*), and the levels of their products in extracellular tissues of patients with primary and secondary lymphedema

V.I. Konenkov, V.V. Nimaev, A.V. Shevchenko, V.F. Prokofiev

Research Institute of Clinical and Experimental Lymphology – Branch of the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
shalla64@mail.ru

Abstract. Cells of various organs and systems perform their functions and intercellular interactions not in an inert environment, but in the microenvironment of tissue fluids. Violations of the normal drainage of tissue fluids accompany lymphedema. An important mechanism of angiogenesis and vasculogenesis regulation in tissue fluids is the production and reception of vascular endothelial growth factors in combination with the regulation of matrix metalloproteinases. The aim of the work was to perform: a comparative analysis of some polymorphisms of vascular endothelial growth factor and their receptors and the genes encoding matrix metalloproteinases in two forms of lymphedema; an analysis of the relationship of these genes' polymorphisms with the levels of vascular endothelial growth factor and matrix metalloproteinases and their inhibitors in serum and affected tissues. Polymorphism of *VEGF* (rs699947, rs3025039), *KDR* (rs10020464, rs11133360), *NRP2* (rs849530, rs849563, rs16837641), matrix metalloproteinases *MMP2* (rs2438650), *MMP3* (rs3025058), *MMP9* (rs3918242), *Timp1* (rs6609533) and their combinations were analyzed by the Restriction Fragment Length Polymorphism method and TaqMan RT-PCR. The serum and tissue fluid levels were determined using the ELISA test system. Changes in the frequency distribution of *MMP2* genotypes in primary and *MMP3* in secondary lymphedema are shown. Significant frequency differences in *NRP2* genotypes were revealed by comparing primary and secondary lymphedema. Features of the distribution of complex genotypes in primary and secondary lymphedema were revealed. The correlation analysis revealed the interdependence of the concentrations of the MMP, TIMP and VEGF products and differences in the structure of the correlation matrices of patients with both forms of lymphedema. It was shown that, in primary lymphedema, genotypes associated with low MMP2 and TIMP2 in serum and tissue fluid are detected, while in secondary lymphedema, other associations of the production levels with combined genetic traits are observed.

Key words: primary lymphedema; secondary lymphedema; *VEGF*; *MMP*; *TIMP*; *KDR*; *NRP*.

For citation: Konenkov V.I., Nimaev V.V., Shevchenko A.V., Prokofiev V.F. Polymorphism of angiogenesis regulation factor genes (*VEGF/VEGFR*), and extracellular matrix remodeling genes (*MMP/TIMP*), and the levels of their products in extracellular tissues of patients with primary and secondary lymphedema. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):433-442. DOI 10.18699/vjgb-24-49

Funding. The study had no sponsor support.

Полиморфизм генов факторов регуляции ангиогенеза (*VEGF/VEGFR*), генов ремоделирования внеклеточного матрикса (*MMP/TIMP*) и уровни соответствующих белков во внеклеточных тканях пациентов с первичной и вторичной лимфедемой

В.И. Коненков, В.В. Нимаев, А.В. Шевченко, В.Ф. Прокофьев

Научно-исследовательский институт клинической и экспериментальной лимфологии – филиал Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
shalla64@mail.ru

Аннотация. Клетки различных органов и систем осуществляют свои функции и межклеточные взаимодействия не в инертной среде, а в микроокружении тканевых жидкостей. Нарушения нормального дренажа тканевых жидкостей сопровождают лимфедему. Важный механизм регуляции ангиогенеза и васкулогенеза в подкожной клетчатке – выработка и рецепция факторов роста эндотелия сосудов в сочетании с регуляцией матриксных металлопротеиназ. Цель настоящего исследования: сравнительный анализ полиморфизма генов фактора роста эндотелия сосудов и его рецепторов вместе с генами матриксных металлопротеиназ при двух формах лимфедемы, анализ взаимосвязи полиморфизма этих генов с уровнем фактора роста эндотелия сосудов и матриксных металлопротеиназ и их ингибиторов в сыворотке крови и пораженных тканях. Полиморфизм *VEGF* (rs699947, rs3025039), рецепторов к нему *KDR* (rs10020464, rs11133360), *NRP2* (rs849530, rs849563, rs16837641), матриксных металлопротеиназ *MMP2* (rs2438650), *MMP3* (rs3025058), *MMP9* (rs3918242), ингибитора металлопротеиназ *Timp1* (rs6609533) и их комбинаций проанализирован методами анализа длин рестрикционных фрагментов и TaqMan RT-PCR. Уровень белков в сыворотке и тканевой жидкости определяли с использованием тест-систем ELISA. Показаны изменения частот распределения генотипов *MMP2* при первичной и *MMP3* при вторичной лимфедеме. Высокодостоверные различия частот генотипов *NRP2* обнаружены при сравнении первичной и вторичной лимфедемы. Выявлены особенности распределения «комплексных» генотипов при первичной и вторичной лимфедеме. Корреляционный анализ показал взаимозависимость концентрации исследуемых белковых продуктов MMP, TIMP и VEGF и выраженные различия в структуре корреляционных матриц пациентов с обеими формами лимфедемы. Продемонстрировано, что при первичной лимфедеме выявляются генотипы, ассоциированные с низкими значениями MMP2 и TIMP2 в сыворотке крови и тканевой жидкости, а при вторичной лимфедеме – иные связи концентраций исследуемых белков с комбинированными генетическими признаками.
Ключевые слова: лимфедема первичная; лимфедема вторичная; *VEGF*; *MMP*; *TIMP*; *KDR*; *NRP*.

Introduction

In recent years, the interest of researchers in the state of the extracellular matrix and the vascular bed of the circulatory and lymphatic systems immersed in it has been constantly growing. The number of scientific publications has been increasing more than 5–6 times a year over the past 40 years. This is due to the understanding that cells of various organs and systems, with their complex internal metabolism, carry out the most important functions and intercellular interactions not in an inert environment, but in a constant microenvironment of tissue fluids carrying a huge number of regulatory factors of the most diverse secreted and membrane-associated nature.

Violations of the normal drainage of tissue fluids lead to tissue hypoxia and a variety of edematous syndromes accompanying various pathological changes, ranging from inflammation to tumor growth. A striking example of impaired drainage of tissue fluid is lymphedema (Miller, 2020). It is represented by both a predominantly genetically determined "primary" and a "secondary" form associated with post-mastectomy consequences or chronic venous insufficiency (Poveshchenko et al., 2010; Quirion, 2010). According to some estimates, between 140 and 200 million people worldwide suffer from lymphedema (Forte et al., 2019). Despite such a wide spread of this disease and numerous studies in this area, the main treatment method is comprehensive physical decongestant therapy and lifelong supportive use of compression knitwear (Vignes, 2017; Executive Committee…, 2020). One of the factors of such, relatively speaking, pathogenetic therapy is, in our opinion, the lack of clear ideas about vascular disorders in the functioning of the blood and lymphatic channels and their interaction with the extracellular matrix, leading to obstruction of the physiological outflow of tissue fluid in the affected regions.

The most important mechanism for the regulation of angiogenesis and vasculogenesis in subcutaneous tissue is the production and reception of the VEGF vascular endothelial growth factor system, represented by VEGF-A, VEGF-B, VEGF-C, VEGF-D, PGF and the VEGFR-1 (Flt-1), VEGFR-2 (Flk-1/KDR), VEGFR-3 (Flt-4) receptor families to them (Vaahtomeri et al., 2017; Rauniyar et al., 2018). Their interaction ensures the growth, remodeling and functioning of the circulatory and lymphatic systems. The genes of these proteins are polymorphic, which affects the level of their expression, affinity and functional activity (Luo et al., 2019; Yap et al., 2019).

It has been established that the VEGF-A ligand binds and transmits signals through the VEGFR-1 and VEGFR-2 receptors, whereas VEGF-B transmits signals exclusively through VEGFR-1, and VEGF-C and VEGF-D have high affinity for VEGFR-3. In addition, there are two neuropilin receptors, which are transmembrane glycoproteins that function on the VEGF-VEGFR axis. Neuropilin-1 (NRP-1), a non-kinase coreceptor for VEGFR-2, functions to enhance the binding and signaling of certain isoforms of VEGF-A, and NRP-2 is a non-kinase coreceptor for VEGFR-3 (Mucka et al., 2016; Stevens, Oltean, 2019; Gao Y. et al., 2020). Understanding the genetic mechanisms underlying endothelial apoptosis and lymphangiogenesis will shed light on the role of disruption of these processes in the development of chronic inflammation and transformation of connective tissue in lymphedema (Saik et al., 2019).

The family of matrix metalloproteinases (MMP) is directly related to the processes of angiogenesis and the activity of regulatory growth factors of the vascular endothelium. The activity of these tissue enzymes is controlled by their tissue inhibitor system (TIMP) (Cabral-Pacheco et al., 2020). The genes encoding them also are widely polymorphic, and protein products are expressed on lymphatic endothelial cells and degrade the collagen of the vascular endothelial lining (Detry et al., 2012). Previously, numerous data were obtained on the effect of regulatory factor gene polymorphism on cell expression and production (Watson et al., 2000; Gao X. et al., 2019).

The aim of this study is a comparative analysis of polymorphism of genes of vascular endothelial growth factor and its

receptors together with genes of matrix metalloproteinases in two forms of lymphedema, analysis of the relationship of polymorphism of these genes with the level of vascular endothelial growth factor and matrix metalloproteinases and their inhibitors in blood serum and affected tissues.

Materials and methods

Patients. The study included patients with the confirmed diagnosis of lymphedema of the extremities. The recruitment of patients was carried out on the basis of Research Institute of Clinical and Experimental Lymphology – Branch of the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences (RICEL) in the period from January 2017 to November 2018. The sample of 174 patients (18–81 years old, median age 52 years) was divided into primary (72 people) and secondary limb lymphedema (102 people). The division of groups into primary and secondary lymphedema is based on etiological signs according to clinical recommendations (Executive Committee…, 2020).

The criteria for primary lymphedema of the extremities were considered to be the development of clinical manifestations without connection with such etiological factors as removal of lymph nodes, radiation therapy, trauma or surgical intervention in the projection of lymphatic collectors. The appearance of clinically significant edema due to a single episode of erysipelas, which is considered a provoking factor against the background of insufficiency of the functional reserve of the lymphatic region. The sample of patients with secondary lymphedema included patients with lesions of both upper and lower extremities. The majority of patients with secondary lymphedema of the upper extremities underwent complex treatment of breast cancer (66 patients – 97.1 %), in 2 patients the cause of edema of the upper limb was the combined treatment of lymphosarcoma of all groups of peripheral lymph nodes and mediastinum IIIA st, lymphogranulomatosis with lesions of the cervical and axillary lymph nodes. The appearance of edema after repeated recurrence of erysipelas was attributed to secondary post-inflammatory lymphedema.

All patients included in the study had no progression or recurrence of a malignant tumor, and belonged to the 3rd clinical group of dispensary observation. According to the classification of the International Society of Lymphologists, stage 2 of the disease prevailed. The third stage of the disease was represented by 7 % of the primary lymphedema group and 8.9 % of the secondary lymphedema group. Informed consent to participate in the study was signed by each participant of the study. The protocol of the clinical research was approved by the RICEL Local Ethics Committee (Primary Protocol No. 127 dated 01/13/2017). Blood serum and interstitial fluid were collected in the morning, on an empty stomach, from patients admitted to the RICEL clinic for a course of complex decongestant therapy, before it began. Patients with current or recent erysipelas were not included in the study. The clinical characteristics of the patients are presented in Table 1.

The comparison group consisted of 339 people of comparable gender and age, residents of the Novosibirsk region. Relatives were not included in either the patient groups or the comparison group.

Genotyping. Genotyping of *VEGF-2578* (rs699947) and *MMP9-1562* (rs3918242) was performed by the Restriction Fragment Length Polymorphism (RFLP) method. The structure of the primers, restriction endonucleases and the product size are shown in Table 2.

Gene polymorphism of *VEGF+936* (rs3025039), *NRP2 13581* (rs849530), *NRP2 68279* (rs849563), *NRP2 92646* (rs16837641), *KDR 17693* (rs10020464), *KDR 14011*

Table 1. Clinical characteristics of patients with primary and secondary lymphedema

Clinical parameters	Secondary lymphedema (n = 102)	Primary lymphedema (n = 72)	P
Age (*Me*)	58 [51; 65]	40 [30; 59]	$P < 0.001$
Women	97 (96.1 %)	53 (74.6 %)	$P < 0.001$
Men	4 (3.9 %)	18 (25.4 %)	$P < 0.001$
Body mass index (*Me*)	32.9 [28.6; 38.6]	27 [24.1; 39.3]	$P = 0.04$
Duration of the disease (months) (*Me*)	56.5 [20.5; 132.3]	192 [84; 300]	
Age of onset of the disease (*Me*)	51.5 [44.59]	25 [15.5; 33.5]	$P < 0.001$
Erysipelas in anamnesis	51 (51 %)	32 (45.1 %)	$P = 0.54$
Hypertension	47 (46.5 %)	13 (18.3 %)	$P < 0.001$
Type 2 SD	19 (18.6 %)	11 (15.5 %)	$P = 0.685$
Hypothyroidism	16 (15.7 %)	4 (5.6 %)	$P = 0.053$
Coronary heart disease	7 (6.9 %)	3 (4.2 %)	$P = 0.529$
Osteochondrosis	14 (13.7 %)	2 (2.8 %)	$P = 0.016$
Chronic venous insufficiency	2 (2 %)	1 (1.4 %)	$P = 1.000$

Note. *Me* is the median; the interquartile range is indicated in square brackets; percentages are indicated in parentheses.

Table 2. Genotyping by the Restriction Fragment Length Polymorphism method

Polymorphic position	The structure of the primers	Restriction endonuclease	Hydrolysis products, bp	
			The "wild" allele	Minor allele
MMP9-1562 C→T	5′ GCCTGGCACATAGTAGGCCC 3′ 5′ CTTCCTAGCCAGCCGGCATC 3′	*Sph I*	435	247; 188
VEGFA-2578 C→A	5′ GGGCCTTAGGACACCATACC 3′ 5′TGCCCCAGGGAACAAAGT 3′	*Bgl II*	267	208; 60

(rs11133360), *MMP2-1306* (rs2438650), *MMP3-1171* (rs3025058), *Timp1* (rs rs6609533) was analyzed by Real-Time PCR using commercial test systems of TaqMan probe method (Syntol, Russia) on the DT-96 thermocycler (DNA Technology, Russia) according to the instructions.

Enzyme-linked immunosorbent assay (ELISA). Quantitative determination of vascular endothelial growth factor, metalloproteinases and their inhibitors was carried out by ELISA kits (ng/mL) according to the instructions: Human VEGF (Quantikine ELISA, R&D Systems, USA) MMP-3 (AESKU Diagnostics, Germany), TIMP-3 (Brand Owner CLOUD-CLONE CORP., USA), MMP-2 and TIMP-2 (Quantikine ELISA, R&D Systems, USA).

Statistical analysis. Statistical processing was carried out using specialized IBM SPSS Statistics 23 (USA) and the software package for volumetric processing of bioinformatics, including multidimensional genetic analysis. In the statistical analysis of the results of the genetic study, the frequency of occurrence of alleles, genotypes and their polylocus combinations, the odds ratio (OR), and the 95 % confidence interval for OR (OR's 95 % CI) were calculated.

The distribution of genotypes was checked with the Hardy–Weinberg equilibrium. The significance level of differences in the frequency in the compared groups was determined by the two-sided criterion of the exact Fisher method for four-field tables, with Bonferroni correction. Considering that the distribution of most of the studied quantitative features was different from normal, nonparametric statistical methods were used. The intergroup differences were assessed using the Mann–Whitney U-test and the Kraskel–Wallis ANOVA. Intra-group differences were assessed using the Wilcoxon sign rank criterion for related samples. Spearman's rank correlation method was used to analyze the strength and direction of the correlation between pairs of features.

The description of quantitative variables is presented in the form of median (Me) and interquartile range (the interval between the 25th [Q 0.25] and 75th [Q 0.75] percentiles). The hypothesis of the normal distribution of quantitative parameters was tested using the Shapiro–Wilk criterion and the Kolmogorov–Smirnov criterion with the Lilliefors correction. The mathematical processing of the relationship of genetic traits with quantitative laboratory parameters was carried out in accordance with the methodological approaches of quantile analysis. With this approach, ranges above p75 (upper quartile, Q3) and lower ranges below p25 (lower quartile, Q1) were taken as parameters of an increased concentration of indicators. The critical level of significance when testing statistical hypotheses was assumed to be 0.05.

Results

The analysis of the distribution of the analyzed complex genetic traits among patients with primary lymphedema revealed a number of pronounced differences from the conditionally "normal" distribution established by us in the study of a significant group of healthy individuals without signs of lymphatic edema of the extremities. When analyzing the degree of differences between individual variants of the studied genetic parameters, they were established only for *MMP2-1306*, associated with *C* predominance among the patient ($p = 0.029$). Along with this, when analyzing the frequency of occurrence of combined genetic traits, including polymorphic variants of both the *MMP* and *VEGF* genes, significantly more pronounced differences were revealed when comparing groups of patients with primary lymphedema and healthy individuals (Table 3).

Basically, these differences were associated with an increase in almost all the analyzed signs in the patient with a level of reliability of differences in the range from 0.048 to 0.001 according to the two-sided Fisher exact test. At the same time, such combinations of genotypes as *VEGF-2578 CA*: *VEGF+936 TT*: *MMP9-1562 CT* and *VEGF-2578 CA*: *VEGF+936 TT*: *MMP2-1306 CC*: *MMP9-1562 CT* were not found among representatives of the group of healthy individuals and were identified exclusively among patients with primary lymphedema. With a certain degree of probability, they can be attributed to the "genetic markers" of an individual's constitutional predisposition to the development of primary lymphedema. Further research is needed to test this hypothesis.

When conducting a similar analysis with secondary lymphedema, in the development of which anatomical factors that occur during surgical damage to the lymphatic and circulatory pathways of the outflow of tissue fluid are of greater importance, we get a different picture from the previous group. In these patients, associative links between the development of the disease and variants of the *MMP2* gene are no longer found, but deviations from the "normal" distribution of the *MMP3* and *MMP9* genes in the analyzed sites are revealed. Thus, the predominance of *MMP3-1171 5A5A* and *MMP9-1562 CT* genotypes was revealed. The frequency of the combination of these homozygous genotypes is also twice as high among these patients relative to the control group.

The inclusion of *VEGF* in the analysis increases the degree of differences between the patients and the controls. Thus, the frequency of a combined genetic trait represented by a combination of homozygous *VEGF+936 CC*: *MMP3-1171 5A5A*: *MMP9-1562 CC* among the patients with secondary lymphe-

Table 3. Frequency of distribution of individual and complex *MMP* and *VEGF* genotypes among patients with primary and secondary lymphedema

Combinations of gene polymorphisms	Genotypes	Patients with lymphedema	Comparison group		OR*	OR's 95 % CI	P(tmF_2)
			n (%)	*N*			
	Primary lymphedema (*N* = 72 patients)						
VEGF-2578:VEGF+936:MMP9-1562	*CA-TT-CT*	2(2.78)	0(0.00)	339	14.37	1.47–140.12	0.0303
VEGF-2578:VEGF+936:MMP2-1306:MMP9-1562	*CA-TT-CC-CT*	2(2.78)	0(0.00)	288	12.21	1.25–119.15	0.0396
VEGF+936:MMP9-1562	*TT-CT*	3(4.17)	2(0.58)	346	7.48	1.23–45.60	0.0379
VEGF-2578:VEGF+936:MMP2-1306:MMP3-1171	*CA-CC-CC-6A6A*	10(13.89)	2(2.35)	85	6.69	1.42–31.65	0.0127
VEGF-2578:VEGF+936:MMP2-1306:MMP3-1171:MMP9-1562	*CA-CC-CC-6A6A-CC*	8(11.11)	2(2.35)	85	5.19	1.06–25.27	0.0445
VEGF+936:MMP3-1171:MMP9-1562	*CC-5A5A-CC*	15(20.83)	6(6.82)	88	3.60	1.32–9.83	0.0104
VEGF-2578:VEGF+936:MMP2-1306:MMP9-1562	***CA-CC-CC-CC***	**22(30.56)**	**39(13.54)**	288	**2.81**	**1.53–5.14**	**0.0013**
MMP3-1171:MMP9-1562	*5A5A-CC*	16(22.22)	9(10.23)	88	2.51	1.03–6.08	0.0488
VEGF-2578:VEGF+936:MMP2-1306	***CA-CC-CC***	27(37.50)	60(20.62)	291	**2.31**	**1.33–4.02**	**0.0051**
VEGF+936:MMP2-1306:MMP9-1562	***CC-CC-CC***	**31(43.06)**	**75(25.51)**	294	**2.21**	**1.29–3.77**	**0.0055**
VEGF+936:MMP2-1306	*CC-CC*	41(56.94)	120(40.40)	297	1.95	1.16–3.28	0.0121
MMP2-1306	*CC*	51(70.83)	182(57.05)	319	1.83	1.05–3.18	0.0338
VEGF-2578:VEGF+936:MMP9-1562	*CA-CC-CC*	27(37.50)	84(24.78)	339	1.82	1.06–3.12	0.0399
VEGF-2578:MMP2-1306:MMP9-1562	*CA-CC-CC*	24(33.33)	66(21.64)	305	1.81	1.03–3.17	0.0453
MMP2-1306:MMP9-1562	*CC-CC*	37(51.39)	117(37.62)	311	1.75	1.05–2.94	0.0338
VEGF-2578:VEGF+936	*CA-CC*	35(48.61)	123(35.76)	344	1.70	1.02–2.84	0.0457
	Secondary lymphedema (*N* = 102 patients)						
VEGF-2578:VEGF+936:MMP2-1306:MMP9-1562	*CC-CC-TT-CC*	4(3.92)	1(0.35)	288	11.71	1.29–106.07	0.0178
VEGF-2578:VEGF+936:MMP9-1562	*AA-CT-CT*	4(3.92)	2(0.59)	339	6.88	1.24–38.11	0.0277
VEGF-2578:VEGF+936:MMP2-1306	*CC-CC-TT*	4(3.92)	2(0.69)	291	5.90	1.06–32.70	0.0416
VEGF-2578:MMP2-1306:MMP9-1562	*CC-TT-CC*	5(4.90)	4(1.31)	305	3.88	1.02–14.73	0.0475
VEGF+936:MMP3-1171:MMP9-1562	*CC-5A5A-CC*	20(19.80)	6(6.82)	88	3.37	1.29–8.84	0.0110
MMP3-1171:MMP9-1562	*5A5A-CC*	23(22.77)	9(10.23)	88	2.59	1.13–5.95	0.0314
VEGF+936:MMP2-1306:MMP3-1171	*CC-CC-5A5A*	22(21.78)	9(10.23)	88	2.44	1.06–5.64	0.0478
VEGF+936:MMP3-1171	*CC-5A5A*	31(30.69)	15(17.05)	88	2.16	1.07–4.33	0.0408
VEGF-2578:VEGF+936:MMP9-1562	*CA-CC-CT*	15(14.71)	26(7.67)	339	2.08	1.05–4.09	0.0496
MMP3-1171	*5A5A*	35(34.65)	18(20.45)	88	2.06	1.07–99.00	0.0353
MMP9-1562	*CT*	37(36.27)	97(25.00)	388	1.71	1.07–2.72	0.0252
VEGF+936:MMP2-1306	*CC-CC*	53(51.96)	120(40.40)	176	1.60	1.01–2.51	0.0490
VEGF+936:MMP9-1562	*CT-CC*	9(8.82)	59(17.05)	297	0.47	0.22–0.99	0.0421
MMP2-1306:MMP3-1171:MMP9-1562	*TC-5A6A-CC*	7(6.93)	15(17.05)	88	0.36	0.14–0.94	0.0401
VEGF+936:MMP2-1306:MMP3-1171	*CC-TC-6A6A*	1(0.99)	8(9.09)	88	0.10	0.01–0.82	0.0131

Note. The comparison group for primary and secondary lymphedema *N* = 339, OR – odds ratio; OR's 95 % CI – 95 % confidence interval for OR; P(tmF_2) – level of statistical significance (*p*) of differences according to the exact Fisher test (two–sided).
* The data in the table are sorted in descending OR, significant differences $p < 0.01$ are highlighted in bold.

Table 4. Comparative analysis of the distribution of individual and complex genotypes between groups with primary and secondary lymphedema

Combinations of gene polymorphisms	Genotypes	Primary lymphedema			Secondary lymphedema			OR*	OR's 95 % CI	$P(tmF_2)$	P_cor
		n	*N*	%	*n*	*N*	%				
NRP2 13581:NRP2 92646:KDR 14011	*TT-GG-TC*	9	72	12.50	0	102	0.00	16.09	2.01–128.72	0.0003	0.0065
MMP3-1171:NRP2 13581:NRP2 92646	*6A6A-TT-GG*	9	72	12.50	1	101	0.99	14.29	1.77–115.49	0.0018	0.0432
MMP2-1306:MMP3-1171:NRP2 13581:NRP2 92646	*CC-6A6A-TT-GG*	8	72	11.11	0	101	0.00	14.12	1.75–114.10	0.0007	0.0354
MMP2-1306:NRP2 13581:NRP2 92646:KDR 17693	*CC-TT-GG-CC*	8	72	11.11	1	102	0.98	12.63	1.54–103.34	0.0040	0.1760
MMP2-1306:NRP2 13581:NRP2 92646:KDR 14011	*CC-TT-GG-TC*	7	72	9.72	0	102	0.00	12.48	1.53–102.13	0.0017	0.0851
NRP2 68279:NRP2 13581	*CC-TT*	7	72	9.72	0	102	0.00	12.48	1.53–102.13	0.0017	0.0156
MMP3-1171:MMP9-1562:NRP2 13581:NRP2 92646	*6A6A-CC-TT-GG*	7	72	9.72	0	101	0.00	12.36	1.51–101.15	0.0018	0.0833
MMP2-1306:NRP2 68279:NRP2 13581	*CC-CC-TT*	6	72	8.33	0	102	0.00	10.76	1.29–89.46	0.0044	0.0973
MMP9-1562:NRP2 13581:NRP2 92646:KDR 14011	*CC-TT-GG-TC*	6	72	8.33	0	102	0.00	10.76	1.29–89.46	0.0044	0.2079
VEGF+936:NRP2 13581:KDR 14011	*CT-TT-TC*	6	72	8.33	0	102	0.00	10.76	1.29–89.46	0.0044	0.0884
VEGF+936:NRP2 68279:NRP2 13581	*CC-CC-TT*	6	72	8.33	0	102	0.00	10.76	1.29–89.46	0.0044	0.0884
MMP3-1171:NRP2 13581:NRP2 92646:KDR 14011	*6A6A-TT-GG-TC*	6	72	8.33	0	101	0.00	10.66	1.28–88.59	0.0046	0.2748
MMP3-1171:NRP2 13581:NRP2 92646:KDR 17693	*6A6A-TT-GG-CC*	6	72	8.33	0	101	0.00	10.66	1.28–88.59	0.0046	0.2428
MMP9-1562:KDR 14011:FOXC2-512	*CC-TC-CT*	10	42	23.81	2	55	3.64	8.28	1.71–40.22	0.0040	0.0725
MMP2-1306:MMP3-1171:NRP2 68279:NRP2 13581	*CC-6A6A-AA-TT*	10	72	13.89	2	101	1.98	7.98	1.69–37.65	0.0042	0.2078
MMP2-1306:NRP2 13581:NRP2 92646	*CC-TT-GG*	12	72	16.67	3	102	2.94	6.60	1.79–24.34	0.0020	0.0426
MMP2-1306:MMP3-1171:NRP2 13581	*CC-6A6A-TT*	15	72	20.83	4	101	3.96	6.38	2.02–20.16	0.0008	0.0195
MMP3-1171:NRP2 13581	*6A6A-TT*	17	72	23.61	5	101	4.95	5.93	2.08–16.97	0.0004	0.0034
MMP3-1171:NRP2 68279:NRP2 13581	*6A6A-AA-TT*	11	72	15.28	3	101	2.97	5.89	1.58–21.96	0.0045	0.1069
VEGF+936:MMP3-1171:NRP2 13581	*CC-6A6A-TT*	11	72	15.28	3	101	2.97	5.89	1.58–21.96	0.0045	0.0935
MMP9-1562:NRP2 13581	*CC-TT*	25	72	34.72	14	102	13.73	3.34	1.59–7.04	0.0016	0.0124
NRP2 13581	*TT*	34	72	47.22	26	102	25.49	2.62	1.38–4.97	0.0036	0.0109
NRP2 13581	*GT*	28	72	38.89	62	102	60.78	0.41	0.22–0.76	0.0055	0.0167
TIMP1-536:NRP2 13581:KDR 14011	*CC-GT-TC*	0	15	0.00	13	32	40.63	0.09	0.01–0.75	0.0039	0.0386
VEGF+936:TIMP1-536:NRP2 13581:KDR 14011	*CC-CC-GT-TC*	0	15	0.00	13	32	40.63	0.09	0.01–0.75	0.0039	0.0463
VEGF-2578:VEGF+936:NRP2 68279:KDR 17693	*CC-CC-AA-CC*	1	72	1.39	14	102	13.73	0.09	0.01–0.69	0.0046	0.2022

Note. OR – odds ratio; OR's 95 % CI – 95 % confidence interval for OR; $P(tmF_2)$ – level of statistical significance (p) of differences according to the exact Fisher test (two–sided); P_cor – adjusted value of $P(tmF_2)$ (taking into account the Bonferroni correction).
* The data in the table are sorted in descending order of the OR value.

dema exceeds the frequency of a similar indicator in the control group by more than 3 times (OR = 3.37; $p = 0.0110$). An even more "broad" combined genetic trait, including a combination of homozygous *VEGF-2578 CC*: *VEGF+936 CC*: *MMP2-1306 TT*: *MMP9-1562 CC*, is more often detected among patients with secondary lymphedema (OR = 11.71; $p = 0.0178$). Along with this, the frequency of another genetic trait represented by a combination of *VEGF+936 CC*: *MMP2-1306 TC*: *MMP3-1171 6A6A* in the patient group is almost 10 times lower: from 9.09 % in the control group to 0.99 % in the patient group (OR = 0.10; $p = 0.0131$). The presence of this combination in the human genome can to some extent be considered a protective factor.

In order to analyze the differences between the structural parameters of angiogenesis genes in more detail, we conducted a study of the distribution of combined signs in both groups of patients with primary and secondary lymphedema. In this report, we have included data on polymorphisms of the *KDR* genes in two polymorphic positions, the *NRP* gene in three polymorphic positions and the *TIMPI* gene. For a clearer representation of the data on the differences obtained, the data with $p < 0.005$ are presented (Table 4).

When evaluating the results, attention is drawn to the presence of combined genetic signs in both groups of patients, which alternatively are not detected in the compared samples. Thus, in the group with primary lymphedema, *TIMP1-536*

CC: *NRP2 13581 GT*: *KDR 14011 TC* and *VEGF+936 CC*: *TIMP1-536 CC*: *NRP2 13581 GT: KDR 14011 TC* are completely absent. In both cases, these combinations are quite widely represented (more than 40 %) in the group with secondary lymphedema ($p = 0.0039$). In contrast to these data, in the group of patients with secondary lymphedema, there are completely no combined genetic signs represented by combinations *NRP2 13581 TT*: *NRP2 92646 GG*: *KDR 14011 TC*; *MMP2-1306 CC*: *MMP3-1171 6A6A*: *NRP2 13581 TT*: *NRP2 92646 GG*; *MMP2-1306CC*: *NRP2 13581 TT*: *NRP2 92646 GG*: *KDR 14011 TC*; *MMP3-1171 6A6A*: *MMP9-1562 CC*: *NRP2 13581 TT*: *NRP2 92646 GG* and a number of others. The significance level of the differences is 0.0003–0.005.

Taking into account the previously obtained numerous data on the effect of regulatory factor gene polymorphism on expression and production, we conducted a study of the MMP 1, 2, 3, 9 proteins level; their tissue inhibitors TIMP 1, 2, 3 level and VEGF level, which did not reveal significant differences between groups of patients with primary and secondary lymphedema according to the median bilateral Mann–Whitney U criterion. At the same time, a continuous correlation analysis revealed not only the interdependence of the analyzed protein MMP, TIMP and VEGF levels, but also pronounced differences in the structure of the correlation matrices of patients with both forms of lymphedema.

Thus, in primary lymphedema, the most significant correlations are revealed between MMP2 and the tissue inhibitor TIMP2 levels (OR = 0.703; $p < 0.01$), whereas the VEGF serum level is inversely correlated with the MMP3 serum level (OR = –0.629; $p < 0.05$). In secondary lymphedema, the most significant interdependencies are revealed between the MMP2 and TIMP2 extracellular fluid levels (OR = 0.727; $p < 0.01$). The VEGF serum level is inversely correlated with this growth factor and MMP9 extracellular fluid level (data are not presented in the table). Other direct and inverse correlations between the signs are also revealed, which probably indicates the functioning of the unified system of humoral factors involved in the processes of angiogenesis and lymphangiogenesis.

Taking into account the presented data on a pronounced associative relationship between the analyzed complex genetic traits and various forms of lymphedema, we conducted an additional analysis of the dependence of high or low MMP 1, 2, 3, 9 of their tissue inhibitors TIMP 1, 2, 3 and VEGF proteins levels on the presence of various combined genotypes in patients of both groups (Table 5).

The conducted quantile analysis showed that in primary lymphedema, genotypes associated exclusively with low MMP2 and TIMP2 level are detected both in the serum and in the extracellular fluid of patients. In the group of patients with secondary lymphedema, other multidirectional associations of the proteins levels with combined genetic traits are shown, which are absent in the group of patients with primary lymphedema.

Discussion

Primary lymphedema occurs as a result of an isolated or developing congenital anomaly of the lymphatic system as part of the syndrome and is associated with dysplasia, hypoplasia or hyperplasia of components of the lymphatic system. The lower extremities are affected in primary lymphedema in the vast majority of cases (Gordon et al., 2021). Secondary lymphedema develops as a complication of another disease or intervention as a result of a violation of the anatomical integrity or obliteration of lymphatic collectors, removal or lesion of lymph nodes, followed by impaired lymph outflow, lymph stasis in lymph vessels and increased endolymphatic pressure (Executive Committee…, 2020). Therefore, it is important to understand the genetic characteristics of these disorders.

In primary lymphedema, we found an increase in the frequency of the *MMP2-1306C* allelic variant of the gene and the homozygous *CC* variant in a single genotype and in other 10 out of 17 combined genetic traits, the frequency of which is higher in this form of the disease with an obvious genetic predisposition (Poveshchenko et al., 2010; Shevchenko et al., 2020).

MMP2 is one of the zinc-dependent endopeptidases that were first discovered as proteases targeting and cleaving extracellular proteins. However, the intracellular significance of MMP has also been discussed over the past 20 years, and research on a new aspect of their functions has been expanding (Bassiouni et al., 2021). Polymorphism of the *MMP2-1306* gene plays a significant role in carcinogenesis, in particular, the *C* variant is associated with a protective role in the development of prostate cancer (Zhang et al., 2017), its frequency is higher among patients with bronchial asthma (Chen et al., 2020). There are a number of reports on changes in the frequency of variants of this polymorphic gene in the promoter region and with other diseases; however, we present data on its association with the development of primary lymphedema for the first time.

Also, for the first time, we present data on changes in the frequency of the *MMP3-1171 5A/6A* gene in lymphedema. Thus, among patients with secondary lymphedema, *5A5A* in the composition of combined signs is characteristic, the frequency of which is higher in this form of the disease, along with *6A6A* in the composition of signs, the frequency of which is lower in secondary lymphedema. There is no such pattern in primary lymphedema. The discriminating role of the homozygous *MMP3-1171 6A6A* is clearly manifested in the analysis of data comparing the distribution of combined genetic traits between groups of patients with primary and secondary forms of lymphedema. We believe that the identified phenomenon requires further study and more detailed clinical analysis.

According to the data presented in Table 3, *MMP3-1171 6A6A*, both as a single trait and as part of a number of combined genetic traits, is associated with low MMP2 and TIMP2 serum levels in patients with primary lymphedema. The homozygous *MMP2-1306 CC* in the composition of combined genetic traits is also associated with low MMP2 and TIMP2 serum levels in primary lymphedema.

In patients with secondary lymphedema, *MMP2 C* is already associated with a high MMP1 and MMP3 serum level, with high VEGF and TIMP3 levels in extracellular fluid. The dependence of any level of the analyzed quantitative signs with the *MMP3-1171* gene polymorphism in secondary lymphedema, unlike its primary form, was not established.

It can be concluded that genetic factors associated with the family of *VEGF* genes and their *VEGFR* receptors involved

Table 5. Relationship between gene polymorphisms and laboratory parameters levels in serum and tissues of patients with primary and secondary lymphedema

Combinations of gene polymorphisms	Genotypes	Laboratory indicators	The level of the laboratory indicator						OR	OR's 95 % CI	P(tmF$_2$)
			high			low					
			n	*N*	%	*n*	*N*	%			
Primary lymphedema (*N* = 44 patients)											
MMP3-1171	*6A6A*	MMP2-serum	2	7	28.57	7	8	87.50	0.06	0.00–0.82	0.0406
MMP3-1171	*6A6A*	TIMP2-serum	0	7	0.00	7	8	87.50	0.03	0.00–0.42	0.0014
MMP2-1306:KDR 14011	*CC-TC*	MMP2-serum	0	7	0.00	5	8	62.50	0.08	0.01–0.95	0.0256
MMP3-1171:KDR 14011	*6A6A-TC*	MMP2-serum	0	7	0.00	6	8	75.00	0.05	0.00–0.64	0.0070
MMP2-1306:MMP3-1171	*CC-6A6A*	TIMP2-serum	0	7	0.00	5	8	62.50	0.08	0.01–0.95	0.0256
MMP3-1171:KDR 17693	*6A6A-CC*	TIMP2-serum	0	7	0.00	5	8	62.50	0.08	0.01–0.95	0.0256
MMP3-1171:KDR 14011	*6A6A-TC*	TIMP2-serum	0	7	0.00	5	8	62.50	0.08	0.01–0.95	0.0256
Secondary lymphedema (*N* = 66 patients)											
KDR 17693	*CT*	VEGF-serum	4	5	80.00	0	6	0.00	17.50	1.22–250.37	0.0152
NRP2 68279	*AA*	TIMP3-ex	11	11	100.0	3	10	30.00	24.00	2.25–255.95	0.0010
VEGF2578:MMP9-1562	*CA-CC*	MMP1-serum	5	7	71.43	1	9	11.11	20.00	1.42–282.46	0.0350
MMP2-1306:MMP9-1562	*TC-CC*	MMP1-serum	4	7	57.14	0	9	0.00	12.50	1.09–143.44	0.0192
MMP9-1562:KDR 17693	*CC-CC*	MMP2-serum	6	11	54.55	0	10	0.00	12.83	1.26–130.52	0.0124
MMP9-1562:KDR 17693	*CC-CC*	MMP3-serum	8	14	57.14	1	11	9.09	13.33	1.32–134.62	0.0330
NRP2 92646:KDR 17693	*GG-CT*	VEGF-serum	4	5	80.00	0	6	0.00	17.50	1.22–250.37	0.0152
VEGF+936:NRP2 13581	*CC-GT*	VEGF-ex	5	6	83.33	3	11	27.27	13.33	1.07–166.38	0.0498
MMP2-1306:MMP9-1562	*TC-CC*	VEGF-ex	3	6	50.00	0	11	0.00	12.00	1.02–141.34	0.0294
MMP9-1562:NRP2 13581	*CC-GT*	VEGF-ex	5	6	83.33	1	11	9.09	50.00	2.56–977.02	0.0054
MMP9-1562:KDR 14011	*CC-TC*	VEGF-ex	4	6	66.67	1	11	9.09	20.00	1.39–287.61	0.0276
NRP2 13581:NRP2 92646	*GT-GA*	VEGF-ex	3	6	50.00	0	11	0.00	12.00	1.02–141.34	0.0294
MMP9-1562:NRP2 92646	*CC-GG*	TIMP3-serum	7	13	53.85	0	9	0.00	11.43	1.15–113.12	0.0167
VEGF+936:NRP2 68279	*CC-AA*	TIMP3-ex	10	11	90.91	3	10	30.00	23.33	1.99–273.31	0.0075
MMP2-1306:NRP2 68279	*CC-AA*	TIMP3-ex	8	11	72.73	0	10	0.00	24.75	2.33–262.60	0.0010
MMP9-1562:NRP2 68279	*CC-AA*	TIMP3-ex	7	11	63.64	1	10	10.00	15.75	1.42–174.25	0.0237
NRP2 68279:NRP2 92646	*AA-GG*	TIMP3-ex	8	11	72.73	1	10	10.00	24.00	2.06–279.64	0.0075
MMP2-1306	*CC*	MMP3-serum	6	14	42.86	10	11	90.91	0.08	0.01–0.76	0.0330
NRP2 68279	AC	TIMP3-ex	0	11	0.00	6	10	60.00	0.06	0.01–0.62	0.0039
MMP2-1306:NRP2 92646	*CC-GG*	MMP3-serum	3	14	21.43	9	11	81.82	0.06	0.01–0.45	0.0048

Note. OR – odds ratio; OR's 95 % CI – 95 % confidence interval for OR; P(tmF$_2$) – level of statistical significance (p) of differences according to the exact Fisher test (two–sided); ex – extracellular fluid.

in the regulation and development of vascular networks of the lymphatic and circulatory systems play a significant role in the development of primary lymphedema.

The VEGFR3 receptor performs the main function in the development and formation of the lymphatic system. Autosomal dominant mutations of *VEGFR3,* which interfere with the functioning of the receptor as a homodimer, not only cause one of the main forms of hereditary primary lymphostasis, namely primary lymphedema (Milroy's disease), but also participate in predisposition to the development of common cyanotic congenital heart defects, demonstrating a new function of VEGFR3 in the early development of heart tissues (Monaghan et al., 2021). This explains the interest in studying the role of a number of *VEGFR* family genes (especially *NRP-2*) associated

with the development of angiogenesis and vasculogenesis, endothelial growth factor genes *VEGF*, metalloproteinases *MMP* and their tissue inhibitors *TIMP*.

As a result of research, there is increasing evidence that in the development of secondary lymphedema, which develops as a result of surgical vascular disorders such as mastectomy, genetic factors of predisposition to the development of lymphatic edema of the extremities also play a role, which allows us to hope for the creation of prognostic criteria for identifying groups at increased risk of their development and preventive measures. The practical significance and prospects of such studies are the positive results of the developed therapy of lymphedema, including with the help of inducers of lymphangiogenesis with VEGF drugs (Forte et al., 2019).

Conclusion

Among patients with both primary and secondary lymphedema, there are significant deviations from the normative indicators established for the control group of healthy individuals in the frequency of distribution of a number of complex genotypes of the *MMP 2*, *3*, *9* and *VEGF* genes, which indicates a significant influence of the studied fragment of the patient's genotype on predisposition to these types of lymphatic edematous syndrome.

The groups of patients with primary and secondary lymphedema differ significantly in the nature of the distribution of a number of complex genotypes of the *MMP 2*, *3*, *9* and *VEGF* genes, which indicates numerous ways of realizing a genetic predisposition to the development of these pathological conditions.

Comparative analysis revealed no significant differences in the level of matrix metalloproteinases, their tissue inhibitors and vascular endothelial growth factor in serum and extracellular fluid of patients with primary and secondary lymphedema.

In both primary and secondary lymphedema, various associative relationships have been established between the studied combined genotypes of gene polymorphism of angiogenesis regulation factors and the level of protein products of these genes in serum and extracellular fluid, which in turn indicates the presence of certain genomic and metabolomic mechanisms for the realization of a genetic predisposition to the development of lymphatic edema.

Data on an increase in the frequency of homozygous *MMP2-1306 CC* in primary lymphedema and an increase in the frequency of homozygous *MMP3-1171 5A5A* in secondary lymphedema were obtained. Both of these polymorphisms are associated with quantitative indicators of the content of protein products MMP, TIMP and VEGF in various variants of limb lymphedema development.

References

Bassiouni W., Ali M., Schulz R. Multifunctional intracellular matrix metalloproteinases: implications in disease. *FEBS J.* 2021;288(24):7162-7182. DOI 10.1111/febs.15701

Cabral-Pacheco G.A., Garza-Veloz I., Castruita-De la Rosa C., Ramirez-Acuña J.M., Perez-Romero B.A., Guerrero-Rodriguez J.F., Martinez-Avila N., Martinez-Fierro M.L. The roles of matrix metalloproteinases and their inhibitors in human diseases. *Int. J. Mol. Sci.* 2020;21(24):9739. DOI 10.3390/ijms21249739

Chen L.-H., Chiu K.-L., Hsia T.-C., Lee Y.-H., Shen T.-C., Li C.-H., Shen Y.-C., Chang W.-S., Tsai C.-W., Bau D.-T. Significant association of MMP2 promoter genotypes to asthma susceptibility in Taiwan. *In Vivo.* 2020;34(6):3181-3186. DOI 10.21873/invivo.12153

Detry B., Erpicum C., Paupert J., Blacher S., Maillard C., Bruyère F., Pendeville H., Remacle T., Lambert V., Balsat C., Ormenese S., Lamaye F., Janssens E., Moons L., Cataldo D., Kridelka F., Carmeliet P., Thiry M., Foidart J.M., Struman I., Noël A. Matrix metalloproteinase-2 governs lymphatic vessel formation as an interstitial collagenase. *Blood.* 2012;119(21):5048-5056. DOI 10.1182/blood-2011-12-400267

Executive Committee of the International Society of Lymphology. The diagnosis and treatment of peripheral lymphedema: 2020 Consensus Document of the International Society of Lymphology. *Lymphology*. 2020;53(1):3-19

Forte A.J., Boczar D., Huayllani M.T., Lu X., McLaughlin S.A. Pharmacotherapy agents in lymphedema treatment: A systematic review. *Cureus*. 2019;11(12):e6300. DOI 10.7759/cureus.6300

Gao X., Ma F., Hao H., Dai Y., Liu W., Xiao X., Gao P., Li S. Association of *VEGFA* polymorphisms with necrotizing enterocolitis in Chinese Han population. *Pediatr. Neonatol.* 2019;60(2):129-134. DOI 10.1016/j.pedneo.2018.07.002

Gao Y., Lee Y., Almazyad A., Birsner A., Li D., Wong S., Wen A., D'Amato R., Adam R.M., Dixon J.B., Srinivasan R.S., Chen H., Bielenberg D.R. Loss of neuropilin 2 in adult lymphatic endothelium promotes lymphedema. *FASEB J.* 2020;34(S1):1. DOI 10.1096/fasebj.2020.34.s1.06345

Gordon K., Mortimer P.S., van Zanten M., Jeffery S., Ostergaard P., Mansour S. The St George's classification algorithm of primary lymphatic anomalies. *Lymphat. Res. Biol.* 2021;19(1):25-30. DOI 10.1089/lrb.2020.0104

Luo Y., Luo J., Peng H. Associations between genetic polymorphisms in the *VEGFA*, *ACE*, and *SOD2* genes and susceptibility to diabetic nephropathy in the Han Chinese. *Genet. Test. Mol. Biomarkers.* 2019;23(9):644-651. DOI 10.1089/gtmb.2018.0320

Miller A. Lymphedema-clinical picture and therapy. Review. *Hautarzt.* 2020;71(1):32-38. DOI 10.1007/s00105-019-04523-z

Monaghan R.M., Page D.J., Ostergaard P., Keavney B.D. The physiological and pathological functions of VEGFR3 in cardiac and lymphatic development and related diseases. *Cardiovasc. Res.* 2021;117(8):1877-1890. DOI 10.1093/cvr/cvaa291

Mucka P., Levonyak N., Geretti E., Zwaans B., Li X., Adini I., Klagsbrun M., Adam R., Bielenberg D. Inflammation and lymphedema are exacerbated and prolonged by neuropilin 2 deficiency. *Am. J. Pathol*. 2016;186(11):2803-2812. DOI 10.1016/j.ajpath.2016.07.022

Poveshchenko A.F., Nimaev V.V., Lyubarsky M.S., Konenkov V.I. Medical and genetical aspects of lymphedema. *Meditsinskaya Genetika = Medical Genetics.* 2010;9(99):3-9 (in Russian)

Quirion E. Recognizing and treating upper extremity lymphedema in postmastectomy/lumpectomy patients: a guide for primary care providers. *J. Am. Acad. Nurse. Pract.* 2010;22(9):450-459. DOI 10.1111/j.1745-7599.2010.00542.x

Rauniyar K., Jha S.K., Jeltsch M. Biology of vascular endothelial growth factor C in the morphogenesis of lymphatic vessels. *Front. Bioeng. Biotechnol*. 2018;6:7. DOI 10.3389/fbioe.2018.00007

Saik O.V., Nimaev V.V., Usmonov D.B., Demenkov P.S., Ivanisenko T.V., Lavrik I.N., Ivanisenko V.A. Prioritization of genes involved in endothelial cell apoptosis by their implication in lymphedema using an analysis of associative gene networks with ANDSystem. *BMC Med. Genomics.* 2019;12(Suppl. 2):47. DOI 10.1186/s12920-019-0492-9

Shevchenko A.V., Prokofyev V.F., Konenkov V.I., Khapaev R.S., Nimaev V.V. Polymorphism of vascular endothelial growth factor gene (*VEGF*) and matrix metalloproteinase (*MMP*) genes in primary limb lymphedema. *Meditsinskaya Immunologiya = Medical Immunology*. 2020;22(3):497-506. DOI 10.15789/1563-0625-POV-1913 (in Russian)

Stevens M., Oltean S. Modulation of receptor tyrosine kinase activity through alternative splicing of ligands and receptors in the VEGF-A/VEGFR axis. *Cells* 2019;8(4):288. DOI 10.3390/cells8040288

Vaahtomeri K., Karaman S., Mäkinen T., Alitalo K. Lymphangiogenesis guidance by paracrine and pericellular factors. *Genes Dev.* 2017;31(16):1615-1634. DOI 10.1101/gad.303776.117

Vignes S. Lymphedema: from diagnosis to treatment. *Rev. Med. Interne.* 2017;38(2):97-105. DOI 10.1016/j.revmed.2016.07.005

Watson C.J., Webb N.J.A., Bottomley M.J., Brenchley P.E.C. Identification of polymorphisms within the vascular endothelial growth factor (VEGF) gene: Correlation with variation in VEGF protein production. *Cytokine.* 2000;12(8):1232-1235. DOI 10.1006/cyto.2000.0692

Yap R.W.K., Lin M.H., Shidoji Y., Yap W.S. Association of stress, mental health, and *VEGFR-2* gene polymorphisms with cardiometabolic risk in Chinese Malaysian adults. *Nutrients.* 2019;11(5):1140. DOI 10.3390/nu11051140

Zhang K., Chen X., Zhou J., Yang C., Zhang M., Chao M., Zhang L., Liang C. Association between MMP2-1306 C/T polymorphism and prostate cancer susceptibility: a meta-analysis based on 3906 subjects. *Oncotarget*. 2017;8(27):45020-45029. DOI 10.18632/oncotarget.16972

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

Received December 5, 2023. Revised February 26, 2024. Accepted February 26, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):443-455

DOI 10.18699/vjgb-24-50

A pipeline for processing hyperspectral images, with a case of melanin-containing barley grains as an example

I.D. Busov[1, 2], M.A. Genaev[1, 2], E.G. Komyshev[1], V.S. Koval[1], T.E. Zykova[1, 2], A.Y. Glagoleva[1], D.A. Afonnikov[1, 2]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia

i.busov@g.nsu.ru

Abstract. Analysis of hyperspectral images is of great interest in plant studies. Nowadays, this analysis is used more and more widely, so the development of hyperspectral image processing methods is an urgent task. This paper presents a hyperspectral image processing pipeline that includes: preprocessing, basic statistical analysis, visualization of a multichannel hyperspectral image, and solving classification and clustering problems using machine learning methods. The current version of the package implements the following methods: construction of a confidence interval of an arbitrary level for the difference of sample averages; verification of the similarity of intensity distributions of spectral lines for two sets of hyperspectral images on the basis of the Mann–Whitney U-criterion and Pearson's criterion of agreement; visualization in two-dimensional space using dimensionality reduction methods PCA, ISOMAP and UMAP; classification using linear or ridge regression, random forest and catboost; clustering of samples using the EM-algorithm. The software pipeline is implemented in Python using the Pandas, NumPy, OpenCV, SciPy, Sklearn, Umap, CatBoost and Plotly libraries. The source code is available at: https://github.com/igor2704/Hyperspectral_images. The pipeline was applied to identify melanin pigment in the shell of barley grains based on hyperspectral data. Visualization based on PCA, UMAP and ISOMAP methods, as well as the use of clustering algorithms, showed that a linear separation of grain samples with and without pigmentation could be performed with high accuracy based on hyperspectral data. The analysis revealed statistically significant differences in the distribution of median intensities for samples of images of grains with and without pigmentation. Thus, it was demonstrated that hyperspectral images can be used to determine the presence or absence of melanin in barley grains with great accuracy. The flexible and convenient tool created in this work will significantly increase the efficiency of hyperspectral image analysis.

Key words: hyperspectral images; machine learning; statistical analysis; barley grains; pigment composition.

For citation: Busov I.D., Genaev M.A., Komyshev E.G., Koval V.S., Zykova T.E., Glagoleva A.Y., Afonnikov D.A. A pipeline for processing hyperspectral images, with a case of melanin-containing barley grains as an example. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(4):443-455. DOI 10.18699/vjgb-24-50

Funding. Development of the pipeline structure, algorithms and programs was supported by RSF, project No. 22-74-00122.

Acknowledgements. Testing of the pipeline and data processing were performed using computational resources of the CDC "Bioinformatics" (supported by budget project No. FWNR-2022-0020).

Конвейер обработки гиперспектральных изображений на примере исследования зерен ячменя, содержащих меланин

И.Д. Бусов[1, 2], М.А. Генаев[1, 2], Е.Г. Комышев[1], В.С. Коваль[1], Т.Е. Зыкова[1, 2], А.Ю. Глаголева[1], Д.А. Афонников[1, 2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский государственный университет, Новосибирск, Россия

i.busov@g.nsu.ru

Аннотация. Анализ гиперспектральных изображений представляет большой интерес при изучении растений. В настоящее время такой анализ используется все более широко, поэтому создание методов обработки гиперспектральных изображений является актуальной задачей. В статье представлен конвейер для работы с гиперспектральными изображениями, который включает: предварительную обработку, базовый статистический анализ, визуализацию многоканального гиперспектрального изображения, а также решение задач классификации и кластеризации с применением методов машинного обучения. В текущей версии пакета программ реализованы следующие методы: построение доверительного интервала произвольного уровня для разницы выборочных средних; проверка сходства распределений интенсивности линий спектра для двух наборов гиперспектральных изображений на основе U-критерия Манна–Уитни и критерия согласия Пирсона; визуализация в двухмерном пространстве с применением методов понижения размерности PCA, ISOMAP и UMAP; классификация с использованием линейной или гребневой регрессии, случайного леса и градиентного бустинга; кластеризация образцов с помощью EM-алгоритма. Программ-

ный конвейер реализован на языке Python с использованием библиотек Pandas, NumPy, OpenCV, SciPy, Sklearn, Umap, CatBoost и Plotly. Исходный код доступен по адресу: https://github.com/igor2704/Hyperspectral_images. Данный конвейер был применен для идентификации пигмента меланина в оболочке зерен ячменя на базе гиперспектральных данных. Визуализация на основе методов PCA, UMAP и ISOMAP, а также использование алгоритмов кластеризации показали, что на базе гиперспектральных данных с высокой точностью можно провести линейное разделение образцов зерен с пигментацией и без нее. Анализ выявил статистически значимые различия в распределении медиан интенсивности для выборок изображений зерен с пигментом и без него. Таким образом, продемонстрировано, что с помощью гиперспектральных изображений с большой точностью можно определить наличие или отсутствие меланина в зернах ячменя. Созданный в данной работе гибкий и удобный инструмент позволит существенно повысить эффективность анализа гиперспектральных изображений.

Ключевые слова: гиперспектральные изображения; машинное обучение; статистический анализ; зерна ячменя; пигментный состав.

Introduction

The presence of pigments in the grain shell affects its various technological properties. For example, flavonoids, anthocyanins and carotenoids have a number of valuable properties, are antioxidants and affect the nutritional value of the grain. The addition of wheat bran with purple pericarp or blue aleurone layer to flour can improve the quality of bakery products through taste, texture and color characteristics (Machálková et al., 2017). Phlobaphenes, which impart red coloration to the grain pericarp, have a positive effect on the duration of grain dormancy and prevent preharvest germination (Flintham et al., 2002). Therefore, wheat genotypes with red grain coloration are used in breeding as donors of genes for resistance to preharvest grain germination (Krupnov et al., 2013; Fakthongphan et al., 2016).

Genetic control of color formation of both grains and other plant organs is carried out by genes encoding enzymes involved in pigment biosynthesis, as well as regulatory genes (Khlestkina, 2014; Lachman et al., 2017; Shoeva et al., 2018). For a number of pigments, these genes have been investigated quite well, to the point of fully deciphering their nucleotide sequences and location in the genome. However, for some pigments, such as melanin, which determines the black coloration of barley grains, the molecular mechanisms of biosynthesis are not yet fully known (Glagoleva et al., 2017; Shoeva et al., 2018).

High-performance, non-destructive and accurate measurement techniques play an important role in assessing seed quality and improving agricultural production (Afonnikov et al., 2016, 2022). Hyperspectral and multispectral imaging techniques covering visible, near-infrared wavelength ranges provide spectral and spatial information for each image pixel. Hyperspectral images represent reflected intensity values for hundreds of wavelength intervals, which is significantly larger than for multispectral images with multiple wavelength ranges (Gowen et al., 2007).

By reducing the total amount of data, multispectral imaging systems aim to rapidly acquire images with relatively low spatial resolution and can be used in real time. Hyperspectral images, on the other hand, are typically used as datasets from which optimal wavelength ranges can be determined, which will be further used in multispectral imaging for a specific application problem (Qin et al., 2013). Such technologies allow obtaining more accurate information about the characteristics of reflected radiation of objects, compared to digital RGB images.

Hyperspectral data analysis has been successfully applied to crop yield estimation and prediction. L. Serrano et al. predicted biomass and yield of winter wheat using spectral indices (Serrano et al., 2000). W.S. Weber et al. (Weber et al., 2012) predicted grain yield using spectra (495–1,853 nm) of canopy and leaf reflectance of maize plants grown under different water regimes and obtained the most appropriate wavelengths for yield prediction. X. Zhang and Y. He (Zhang, He, 2013) developed a method for early and rapid seed yield estimation using hyperspectral images of oilseed rape leaves in the visible and near-infrared regions (380–1,030 nm). Soybean (*Glycine max*) seed yield was predicted based on hyperspectral data (395–1,005 nm) and machine learning algorithms: multilayer perseptron, support vector method and random forest, which also identified the most significant reflectance spectrum (395 nm) (Yoosefzadeh-Najafabadi et al., 2021).

Hyperspectral reflectance analysis can provide reliable information on seed viability of both weedy (Matzrafi et al., 2017) and cultivated plants: rice (He et al., 2019; Jin et al., 2022), wheat (Zhang et al., 2018), maize (Ambrose et al., 2016; Wakholi et al., 2018), peanut (Zou et al., 2023), melon (Kandpal et al., 2016), Japanese spinach mustard (Ma et al., 2020).

Based on hyperspectral technologies, innovative methods for diagnosing plant diseases are being developed (Cheshkova, 2022). Hyperspectral imaging technology covering the visible and near-infrared wavelength range (400–1,000 nm) was used to analyze rice to detect discolored, diseased seeds infected with bacterial panicle blight (*Burkholderia glumae*). It has been shown that determining the intensity of reflected radiation in a small number of wavelength bands is sufficient for accurate (>90 %) classification of pathogen-affected and healthy plants (Baek et al., 2019).

Hyperspectral images are used to determine the chemical composition of seeds of cultivated plants. Near-infrared (895–2,504 nm) reflectance analysis has been shown to have potential in predicting anthocyanin content in black rice grains (Amanah et al., 2021). C. Liu et al. (Liu et al., 2020) demonstrated the feasibility of using near-infrared (930–2,500 nm) hyperspectral data analysis to determine the starch content of maize grains. G. Yang et al. (Yang et al., 2018) applied Raman hyperspectral technology with line scanning to determine the chemical composition of maize seeds. It was found that the characteristic Raman peaks identified at 477, 1,443, 1,522, 1,596 and 1,654 nm in the spectrum from 380 to 1,800 nm were associated with corn starch, oil and starch mixture, zeaxanthin, lignin and oil in corn seeds, respectively.

A method for non-destructive estimation of the concentrations and spatial distribution of moisture, protein and sugars at different developmental stages of vigna seeds has been proposed based on multispectral data from 20 discrete wavelengths in the ultraviolet, visible and near-infrared regions (ElMasry et al., 2022). Handheld near-infrared spectroscopy and hyperspectral imaging techniques have been used to quantify oil and fatty acid content and to classify seed species of the genus Brassica (da Silva Medeiros et al., 2022). Hyperspectral images have been used to solve the classification problem for grains of rice (Díaz-Martínez et al., 2023), ryegrass (Reddy et al., 2023) and many other crops important for the agricultural industry.

Platforms are being developed to provide hyperspectral information on seeds, such as HyperSeed, which includes a high-throughput line-scan spectrograph (600–1,700 nm) and open-source software based on a graphical user interface. The system was used to classify rice seeds (with 97.5 % accuracy) grown under heat stress and in control environments using both traditional machine learning and neural network (3D CNN) models (Gao et al., 2021).

Thus, the analysis of hyperspectral images is of great interest in various tasks related to plant research. However, developing algorithms to analyze such data is a time-consuming task.

This paper presents a hyperspectral image analysis pipeline, the use of which can significantly reduce the time cost in hyperspectral imaging-related research. We applied the developed pipeline to determine the melanin content of barley grains. Although the presence of melanin accounts for the dark coloration of the grain, in practice, visual determination of its presence is difficult. The dark color of the grain may be associated with the accumulation of anthocyanin pigments, which accumulate in the aleurone of the grain, giving ripe grains a gray color. Barley grains can also darken during storage. Therefore, accurate determination of the presence of melanin requires additional analysis, for example, immersion of grains in alkali solution for its extraction.

In this paper, we present a tool for hyperspectral image research, a pipeline, the use of which can significantly reduce time costs in such research. The capabilities of the developed pipeline are demonstrated on the example of the task of melanin content determination in barley grains. The task of studying the spectrum of melanin-containing and non-melanin-containing grains was chosen for testing, since it is known that there are significant differences in their spectrum. Our analysis also showed significant differences in the spectrum of grains containing melanin and samples without this pigment. Unlike other works in this area, in addition to classifying the samples, we had the task of implementing a pipeline to facilitate and automate the acquisition of hyperspectral images. The developed pipeline allows us to visualize and cluster the input data, as well as to perform their statistical analysis.

Materials and methods

Plant material. Seeds of 313 barley (*Hordeum vulgare*) accessions were selected for the study, of which 117 accessions contained melanin and the remaining 196 accessions lacked this pigment (Supplementary Material)[1]. The material was obtained from the barley collection of the All-Russian Institute of Plant Genetic Resources named after N.I. Vavilov (VIR, https://www.vir.nw.ru), barley collection of the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences (ICG, http://www.bionet.nsc.ru). Material from the Oregon Wolfe Barleys population (OWB, https://barleyworld.org/owb) was also used. Biochemical analysis of samples with stained grain, as well as a detailed description of the melanin detection method were performed by A.Y. Glagoleva et al. (Glagoleva, et al. 2022).

Chemical method for determination of pigment composition of grains. To determine the qualitative presence of melanin in the grain, extraction with 2 % NaOH followed by blackening of the solution was performed. Based on this method, each of the samples was assigned a pigmentation type based on the presence of pigment: “contain melanins” or “do not contain melanins”.

Image acquisition. Hyperspectral images of grains were obtained using a Cubert S185 camera with a Cinegon 1.8/16 lens. For this purpose, a plastic petri dish with a diameter of 55 mm filled with grains without gaps was placed on a white matte sheet of A3 paper. A diffusing light was placed on the sides, and the camera was fixed on a tripod from above, with the lens vertically downward. At the output, the camera produced a 138-channel hyperspectral image, each channel of which corresponded to the reflection intensity in a certain wavelength range (Fig. 1). The size of the hyperspectral image was: 50 by 50 pixels, spectral range: 450–998 n.m., spectral channel width: 4 n.m. The images were saved in tiff format.

Thus, the hyperspectral image obtained by a Cubert S185 camera is a hypercube, in which indices i, j ($i, j = 1, ... 50$) correspond to spatial coordinates (image pixels), index $k = 1, ... 138$, corresponds to hyperspectral lines with a certain wavelength. Each element of this hypercube corresponds to the intensity of reflected radiation from the subject for a pixel in the image with spatial coordinates i, j and spectral line with serial number k.

Images for the study of the pigment composition of barley grains were obtained from several series of surveys over several days.

Pipeline description. The input data for the pipeline are hyperspectral images in tiff format described in the previous section and calibration hyperspectral images (black and white background images in tiff format).

Multichannel hyperspectral image analysis is performed in several steps including preprocessing, feature extraction, normalization and direct data analysis (Fig. 2).

Hyperspectral image preprocessing and feature extraction. The nature of ambient light can affect the reflected spectrum intensities (Zahavi et al., 2019). In order that the reflected emission intensities on different spectrum lines could be compared for different imaging conditions, we used image calibration according to the following formula:

$$R_{ijk} = \frac{S_{ijk} - D_{ijk}}{W_{ijk} - D_{ijk}},$$

[1] Supplementary Material is available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Busov_Engl_28_4.pdf

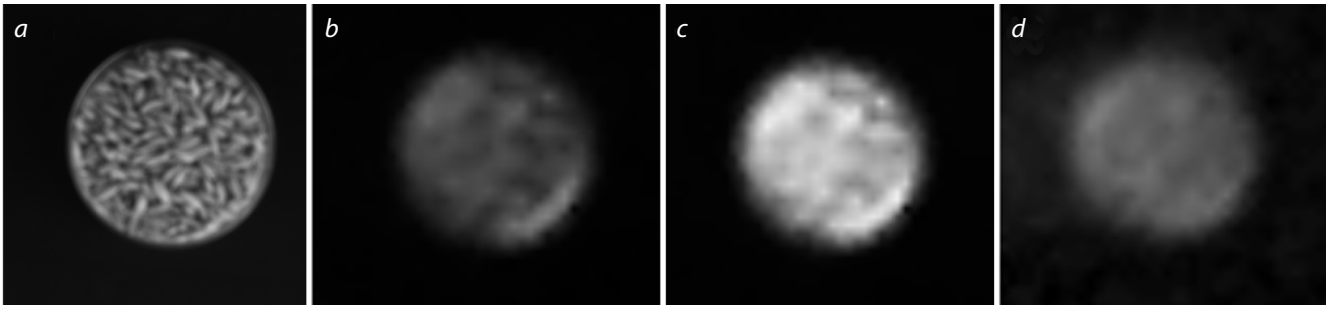


Fig. 1. Image of barley grains in a Petri dish in shades of gray (*a*) and visualization of reflected radiation intensity in the wavelength intervals of 450 nm (*b*), 554 nm (*c*), and 986 nm (*d*).

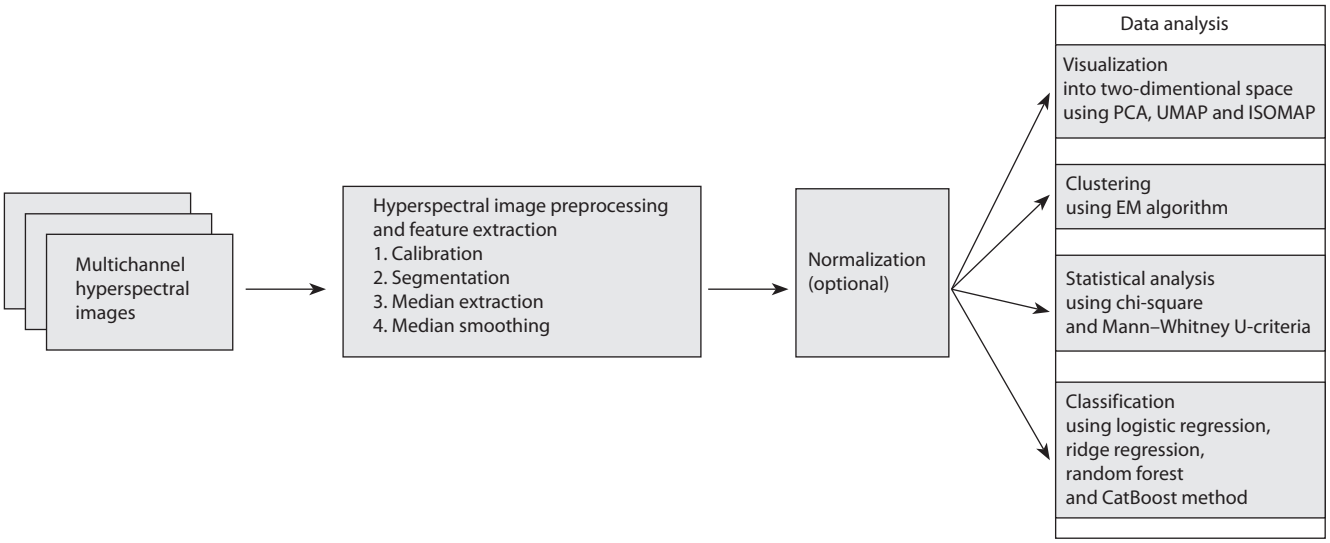


Fig. 2. Pipeline schematic for hyperspectral image analysis.

where S_{ijk} is the barley hyperspectral image hypercube element, D_{ijk} is the black background calibration image element, W_{ijk} is the white background calibration image element, R_{ijk} is the calibrated image element.

The calibrated images are converted to a three-channel image approximating RGB based on intensities for wavelengths 450 nm (blue), 510 nm (green), 630 nm (red) using a threshold transformation (OpenCV library threshold() function (Howse J., 2013)). This image is converted to a grayscale image (OpenCV library function cvtColor()) and binarized to highlight the Petri dish region with grains. If necessary, the pipeline allows you to use your own implementation of segmentation, but for the task at hand, segmentation by threshold value is sufficient.

Then, for each image, the medians for each hyperspectral channel are calculated from the pixel values in the segmented area occupied by grains. The Savitzky–Golay filter (Savitzky, Golay, 1964) is used to smooth the median values. The obtained vector of medians characterizes hyperspectral data for each studied sample.

Normalization. In order to eliminate differences arising between imaging series, 2 methods of image normalization were implemented in the pipeline. The first way of normalization is standardization (subtraction of the sample mean and division by standard deviation) by identical samples of each image (vectors of medians). The second method is standardization by identical groups (samples containing/not containing melanin), within each series.

Data analysis. Dimensionality reduction methods. The pipeline uses 3 dimensionality reduction methods: PCA (principal component analysis) (Jolliffe, 2002), ISOMAP (isometric mapping) (Balasubramanian, Schwartz, 2002), and UMAP (uniform manifold approximation and projection) (McInnes, et al., 2018) to visualize samples clearly in hyperspectral data space. PCA is a linear dimensionality reduction method that preserves the largest percentage of variance.

ISOMAP, UMAP are nonlinear dimensionality reduction methods. The UMAP method builds a weighted graph where only the nearest neighbors are connected by edges (the number of neighbors is given as a pipeline parameter). The ISOMAP method first constructs a sparse graph where, just as in the graph for UMAP, only the nearest neighbors are connected by edges (the number of neighbors is given as a pipeline parameter). Then, either the Dijkstra algorithm (Cormen et al., 2002) or the Floyd–Worshall algorithm (Cormen et al., 2002) is used to compute the distances between objects in the sparse graph for the ISOMAP method. After constructing the graphs and the distance matrix for them, the UMAP and ISOMAP

methods are used to determine the position of the samples in a space of lower dimensionality (usually 2 or 3) that preserves the distances between objects. The dimensionality reduction methods were implemented using the Sklearn (Hao et al., 2019) and Umap (Becht et al., 2019) libraries.

Visualization. After the preprocessing and feature extraction stages, each sample (hyperspectral image) is presented as a vector lying in a dimensionality space equal to the number of hyperspectral image channels. The elements of the vector correspond to the reflected radiation intensity for the corresponding channel. After obtaining the coordinates of the samples in lower dimensionality spaces, visualization in the form of a scatter diagram was performed using the plotly.express. scatter function of the Plotly library (Stančin I. et al., 2019).

Clustering. The pipeline implemented clustering using the EM algorithm (Dempster et al., 1977). It was assumed that each sample could belong to each cluster with a probability obeying the Gaussian distribution mixture model. The parameters of the distributions were found using the maximum likelihood method, using the EM algorithm. The main hyperparameters of clustering are: dimensionality of the space in which clustering takes place, method of dimensionality reduction, method of initialization of weights (random initialization, initialization by the *k*-means method). The pipeline returns a table with information about the most frequent group in each cluster and the percentage of samples in it. The Sklearn library was used to implement clustering.

Statistical analysis. In the created pipeline for the difference of sample averages of two groups of images, it is possible to determine the confidence interval at a given level of significance, which is based on the central limit theorem (CLT). According to the CLT, if the sample size is sufficient, we can assume that the difference of sample averages is normally distributed. For this random variable, the sample mean and sample variance are calculated, and thus confidence intervals of arbitrary level are constructed.

Tests based on the Mann–Whitney U-criterion (Wilcoxon, 1945) and chi-square criterion (Greenwood, Nikulin, 1996) were added to the pipeline to test the hypothesis that the distributions of the two groups coincide. Statistical analysis was implemented using the SciPy library (Nunez-Iglesias et al., 2017).

Classification. The developed pipeline classifies hyperspectral images using methods such as logistic regression (Norman, Harry, 2007), ridge regression (Norman, Harry, 2007), random forest (Ho,1995) and gradient boosting (Prokhorenkova et al., 2017). The pipeline returns tables with classification results on metrics such as accuracy, F1, precision and recall, as well as error matrices for each classifier. The first table contains classification results for macro metrics and the second, for micro metrics. If a function that converts a group into a vector is passed to the pipeline, the pipeline returns a third table with the averaged binary classification results for each individual component of the vector. Classification is implemented using the Sklearn and CatBoost libraries (Hancock, Khoshgoftaar, 2020).

Results

Sample images for pigment composition analysis were obtained from three series of surveys. In two series, grains containing melanin were absent. In one series, both grains with melanin and grains without this pigment were present. There were no identical samples in different series of imaging. For each sample, two images were obtained: a hyperspectral image and a high-resolution image. Since samples without pigment were present in all imaging series, normalization by samples of grains with no pigment was performed.

Median graph

The obtained medians were used to plot the dependence of intensity on wavelength for each image (Fig. 3). As can be noted, the hyperspectrum of grains containing melanin differs markedly from the hyperspectrum of grains without this pigment.

The plot without normalization for the median curves shows local maxima in the 600–700 nm range, and local minima in

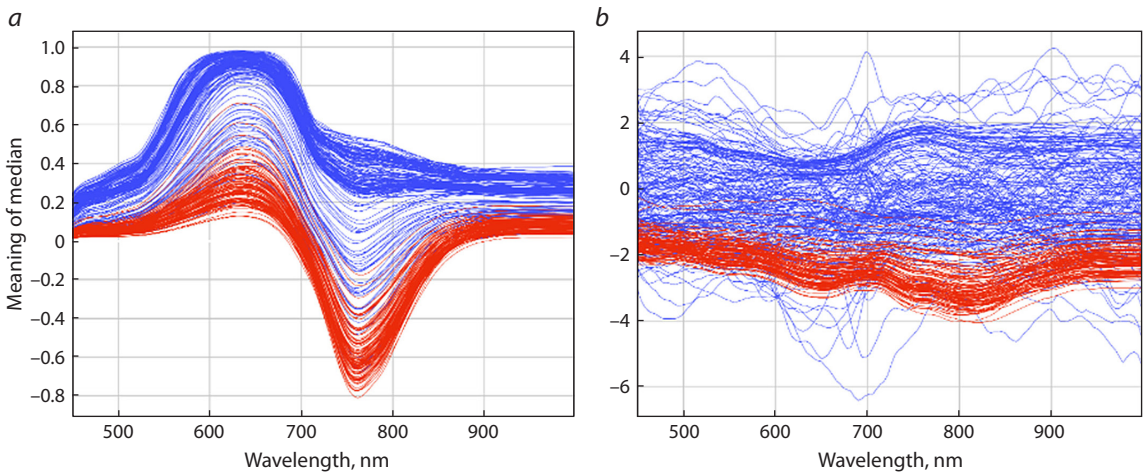


Fig. 3. Graph of the dependence of the median intensity of reflected radiation for barley grain samples as a function of wavelength.

On the left is the graph (*a*) without normalization, on the right is the graph (*b*) of medians after normalization by identical groups. Blue lines correspond to medians of images of barley grains without melanin, and red lines, to medians of images of grains with melanin.

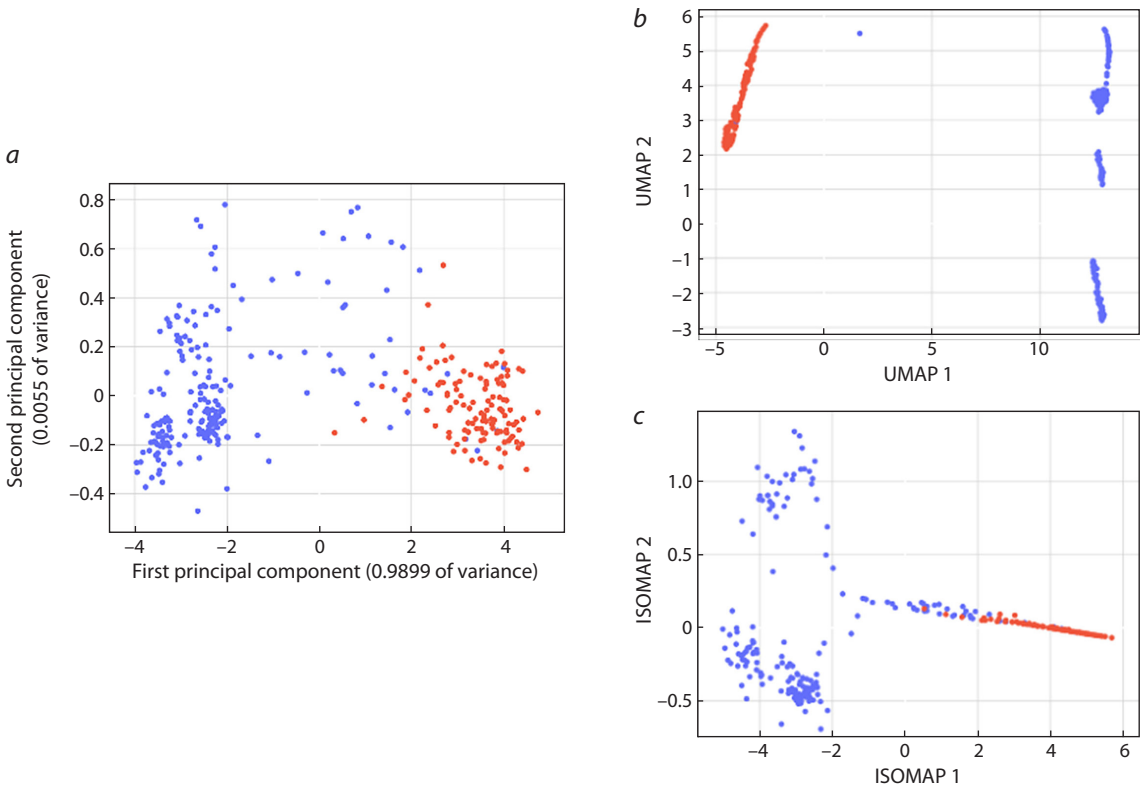


Fig. 4. Plots of distributions in two-dimensional space obtained with PCA (*a*), UMAP (*b*) and ISOMAP (*c*).
Blue points correspond to samples without melanin and red points, to samples with melanin.

the 700–800 nm range. Most of the median curves of grains with melanin are more tightly clustered (wavelength-averaged dispersion is smaller) and have smaller mean values than the curves of samples without pigment over the entire wavelength range. Despite the partial overlap, most of the median curves of the samples with pigment are distinguishable from the median curves of the samples without pigment.

Visualization in two-dimensional space

In the PCA (Fig. 4*a*) and ISOMAP (Fig. 4*c*) plots, it can be observed that the dispersion in grains without melanin is larger than in grains with this pigment.

Clustering results

Clustering was performed into 2 clusters representing samples with melanin and samples without pigment. Clustering confirms that the medians of hyperspectral images are separable with high accuracy (Fig. 5, Table 1).

The samples with and without pigment were least clearly separated in the PCA plot (Fig. 4*a*): samples without pigment (blue dots) are present near the cluster of samples with melanin (red dots on the right). These samples were assigned to the second cluster (green dots) during clustering (Fig. 5*a*). In contrast, in the UMAP plot, all samples with pigment were arranged in isolation (Fig. 4*b*), on the top left, while samples without pigment formed clusters of dots on the right. However, in the clustering plot (Fig. 5*b*), single samples on the left were assigned to the second cluster. The ISOMAP plot (Fig. 4*c*) shows a good clustering of samples with melanin, while samples without pigment were distributed on the left, and to a lesser extent, on the right side of the plot, partially overlapping with samples with melanin. In clustering (Fig. 5*c*), some of these samples were assigned to the second cluster (green dots in the right part of the graph).

Table 1 numerically confirms that the median vectors of hyperspectral images of grains of different classes (containing and not containing melanin) in clustering mainly fall into different clusters, which indicates the existence of significant differences in the spectrum of grains with and without pigment. It is also worth noting that the first cluster includes samples exclusively without melanin.

Statistical analysis

Figure 6*a* shows the differences of sample mean values of reflected radiation intensity for barley samples for all wavelength intervals. As can be noted, the mean values of different groups of grains are statistically significantly different in the whole wavelength interval under consideration. Figure 6*b* shows a plot of the dependence of the logarithm of the reliability of differences (p-value) on wavelength for the Mann–Whitney U-criterion. This criterion (taking into account the Bonferroni correction) allowed us to detect statistically significant differences for the entire hyperspectrum under study.

Classification results

The task of classifying hyperspectral grain images based on melanin content is a binary classification task. Table 2 shows the classification accuracy estimates for accuracy, F1, preci-

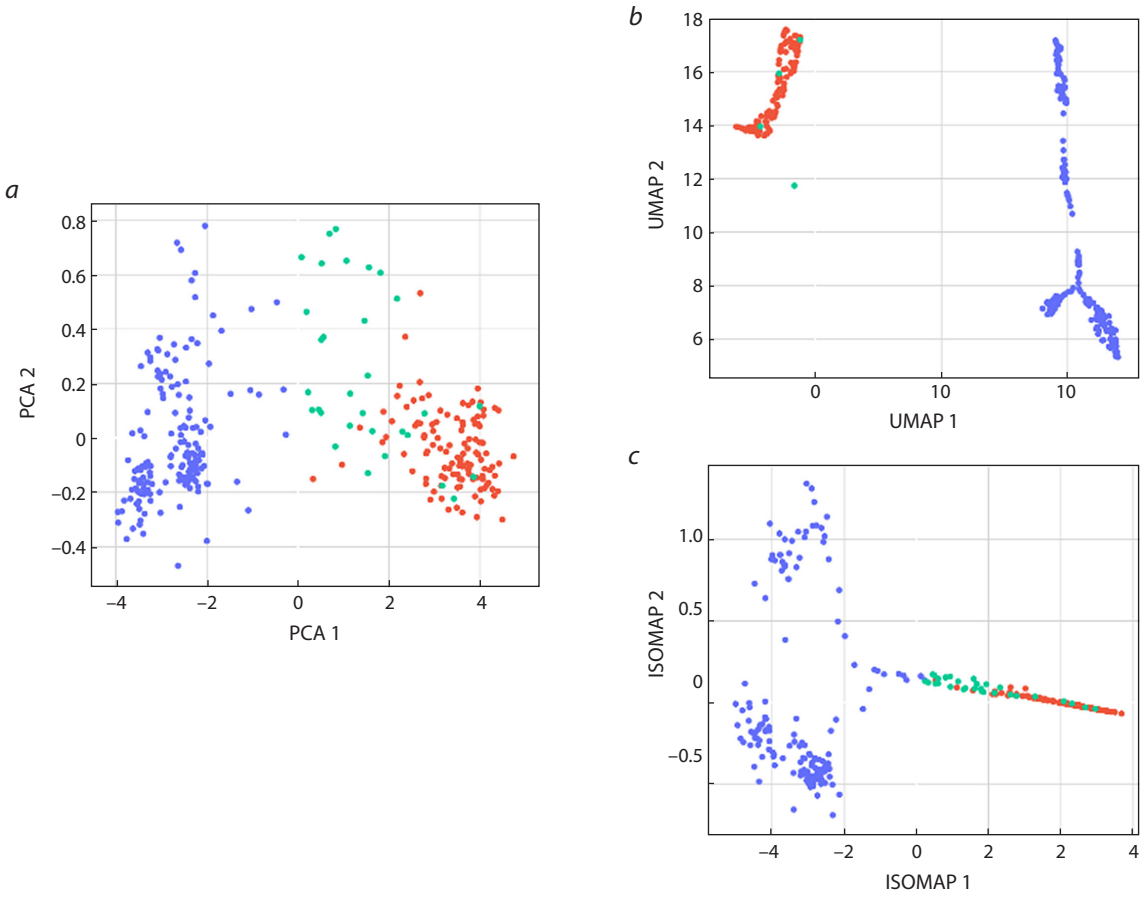


Fig. 5. Visualization of the clustering results using the EM algorithm. Initialization was performed using the *k*-means method in a space of dimensionality 15, using the dimensionality reduction methods PCA (*a*), UMAP (*b*) and ISOMAP (*c*).

Blue dots correspond to grains that do not contain melanin and are in the first cluster. Red dots correspond to grains that contain melanin and are in the second cluster. Green dots stand for grains that do not contain melanin and belong to the second cluster. Samples containing melanin but assigned to the first cluster were absent.

Table 1. Clustering accuracy by the EM algorithm with random initialization in dimension space 15, using the UMAP dimensionality reduction method

Cluster	Prevailing class	Dimensionality reduction method	Frequency of the most frequently occurring class in the cluster
2	Melanin	PCA	0.79
1	Melanin-free	PCA	1.00
2	Melanin	UMAP	0.96
1	Melanin-free	UMAP	1.00
2	Melanin	ISOMAP	0.80
1	Melanin-free	ISOMAP	1.00

Note. The prevalent class is the most frequent class of samples in the cluster.

sion and recall metrics for each dimensionality reduction method. The test sample size was 47 samples and the training sample size was 266 samples. The *k*-fold cross validation ($k = 4$) was used in training. 18 samples in the test sample contained melanin; 29 samples were without melanin; 99 samples in the training sample were with melanin; 167 samples were without this pigment.

The studied grain samples contained anthocyanins in addition to melanin, which allowed us to study the possibility of differentiation between melanins and anthocyanins. Samples were classified in the 15-dimensional space previously obtained by PCA using logistic regression (266 samples for the training sample and 47 samples for the test sample). As a result, classification errors occurred mainly between the classes

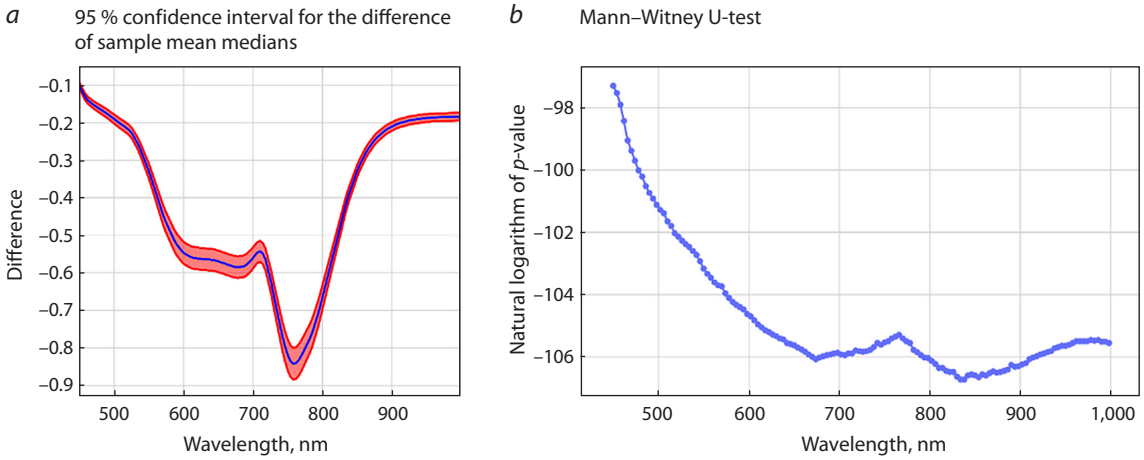


Fig. 6. (*a*) 95 percent confidence interval for the difference of sample mean medians. The blue line is the values of the difference of sample mean differences. The red area is the 95 percent confidence interval. (*b*) Logarithm *p*-value plot for the Mann–Whitney U-criterion for the difference in mean values of reflected spectrum intensity for grain samples with and without melanin for different wavelength intervals.

Table 2. Classification results of the test sample in dimension space 15, using PCA, UMAP and ISOMAP for dimensionality reduction

Model	Accuracy	F1	Precision	Recall
Logistic regression	0.979	0.971	0.944	1.000
Ridge regression	0.979	0.971	0.944	1.000
Random forest	0.979	0.971	0.944	1.000
Gradient boosting (catboost)	0.979	0.971	0.944	1.000

Note. For the obtained training and test samples, the results when using different dimensionality reduction methods on the test sample were the same.

"without pigments" and "with anthocyanins only", as well as in identifying samples containing both pigments and grains containing only melanin (Fig. 8).

Based on the results of the statistical analysis, no statistically significant differences (p-value < 0.05/138, taking into account the Bonferroni correction) were found across the spectrum for grains containing only melanin and grains with both pigments. The lowest p-value for the Mann–Whitney criterion for these groups was reached at 774 nm and was 0.0438 (Fig. 9*a*). For grains containing only anthocyanins and grains without pigments, statistically significant differences (p-value < 0.05/138, taking into account the Bonferroni correction) were found at wavelengths falling in the red and infrared bands (> 714 nm) (Fig. 9*b*).

Discussion

Pipelines in the field of hyperspectral image processing

There are many state-of-the-art approaches to automate the process of hyperspectral data analysis. They utilize a wide range of machine learning, computer vision and advanced data processing techniques. Hyperspectral images are characterized by high dimensionality, large data volume, are affected by noise, require calibration and normalization, and are more difficult to visualize compared to RGB images. In addition, there is a problem of training sample size. To solve this problem, various methods of increasing the size of training sets (augmentation) are used. On the other hand, the high dimensionality of hyperspectral data can easily lead to a high level of data redundancy. To solve this problem, algorithms for ranking and filtering significant features, as well as for selecting groups of significant spectra are used.

The acquired hyperspectral raw data are preprocessed: outlier detection using principal component analysis (PCA), group averaging, scaling and centering (Yoosefzadeh-Najafabadi et al., 2021); calibration of the acquired images using reference images (dark and white); normalization; Savitzky–Golay filtering; and parameter ranking and filtering for classification to improve model accuracy and generality (Amanah et al., 2021).

The use of dimensionality reduction techniques may lead to a decrease in classification accuracy, however, it may be justified in order to increase the generality of the models – to avoid overfitting them. Thus, the development of approaches for solving individual problems using hyperspectral data requires multi-stage processing, the realization of which is possible in a software pipeline architecture, where each individual stage is replaceable and can be carefully tuned and adapted.

To solve such problems, pipeline approaches are currently being actively developed. For example, in the work of F. Zhu et al. (Zhu et al., 2024), the authors investigated ways to preprocess spectral data to effectively reduce the effect of different illumination on chlorophyll estimation in basil crops grown under different light intensities. The authors determined the

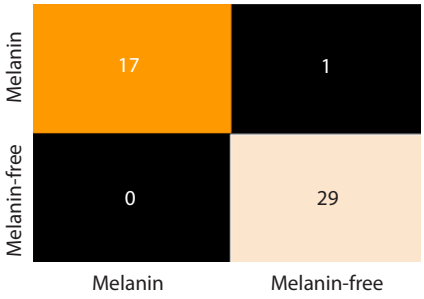


Fig. 7. Error matrix for the test sample in dimension space 15.

For the obtained training and test samples, the results using different dimensionality reduction methods and different classification models on the test sample were the same.

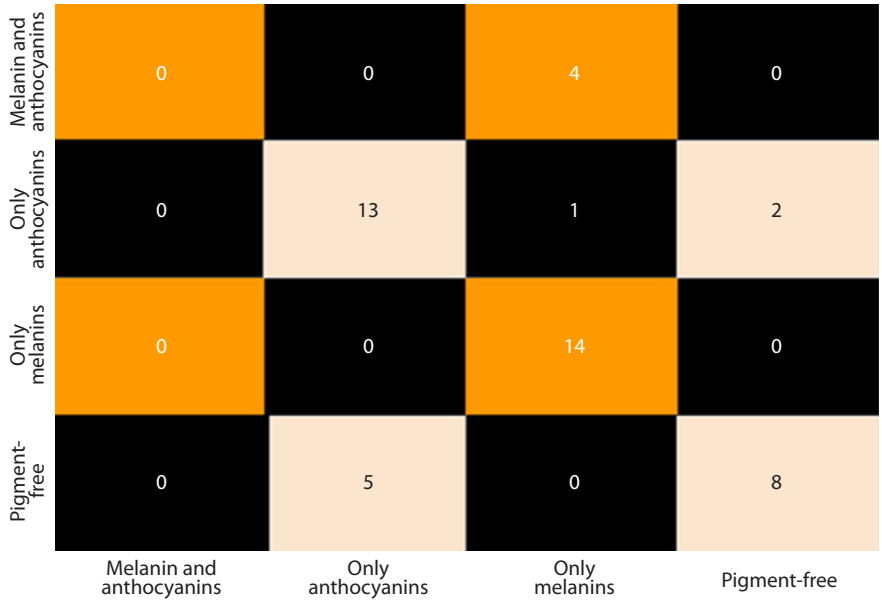


Fig. 8. Classification error matrix based on logistic regression of grain samples into 4 classes: containing melanin and anthocyanins, only anthocyanins, only melanin and without pigments.

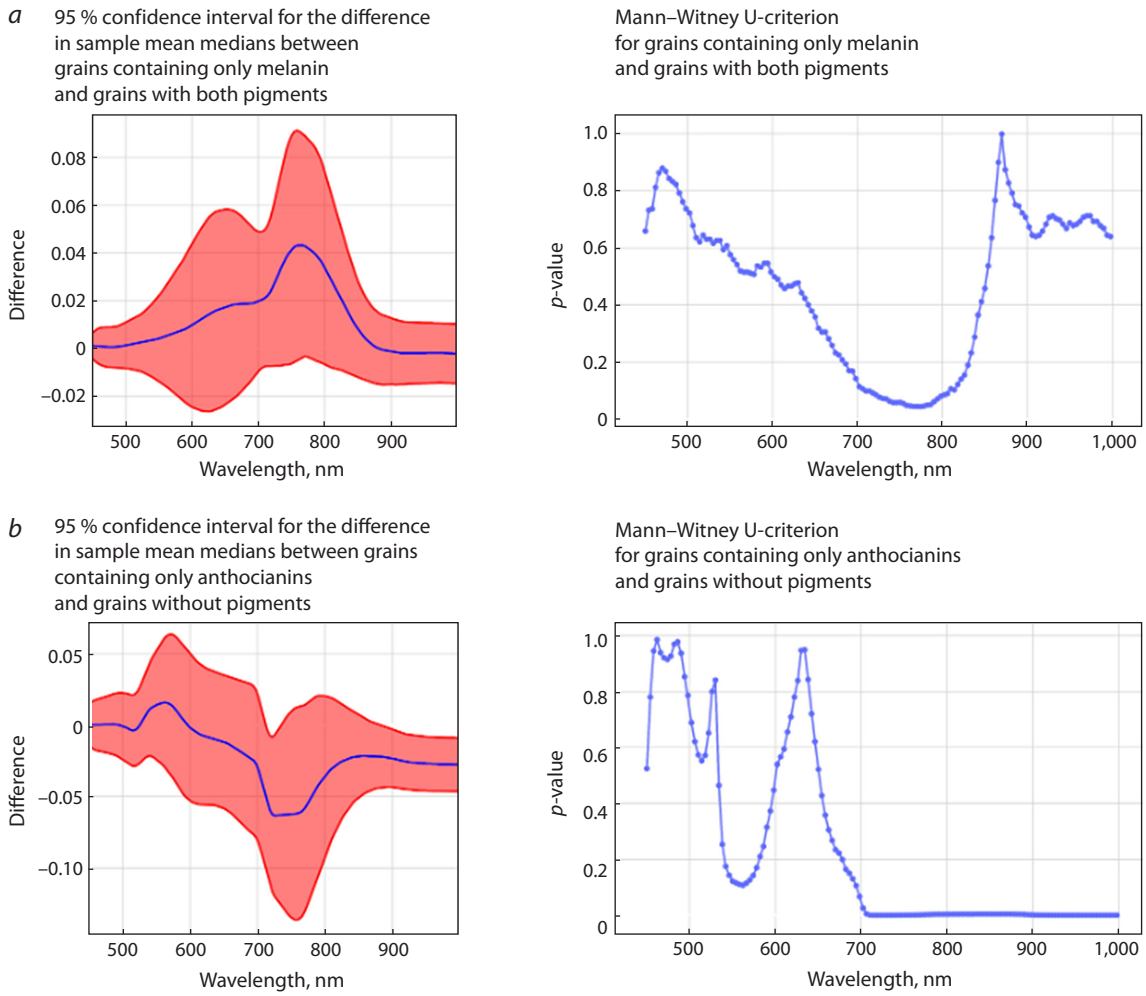


Fig. 9. Plots of the logarithm of the *p*-value for the Mann–Whitney U-criterion for the difference in mean values of reflected spectrum intensity and 95 percent confidence intervals for the difference in sample mean medians.

optimal analysis pipeline for near-field hyperspectral imaging data by evaluating the performance of regression modeling and obtaining satisfactory chlorophyll distribution maps consistent with observed differences in chlorophyll levels.

In their work, H. Feng et al. (Feng et al., 2017) developed an integrated image analysis pipeline for automatic processing of large volumes of hyperspectral data. Models were built to accurately quantify 4 pigments (chlorophyll a, chlorophyll b, total chlorophyll, and carotenoids) from rice leaves and identified important wavelength groups (700–760 nm) associated with these pigments. At the tillering stage, the R^2 values and mean absolute percentage errors of the models were 0.827–0.928 and 6.94–12.84 %, respectively.

By establishing a four-stage image processing and data analysis management pipeline, the applicability of hyperspectral remote sensing for early detection of drought stress and root-knot nematodes (RKN) infestation in tomato plants was evaluated (Žibrat et al., 2019). The pipeline included: image acquisition, data extraction, preprocessing and analysis. By combining discriminant analysis based on partial least squares and support vector machine with time series analysis, the authors achieved 100 % classification success in determining irrigation regime and infestation rate. Thus, the development of pipelined solutions for hyperspectral data analysis is an actively developing area at the moment.

The hyperspectral data analysis example presented in this paper also uses a pipeline approach, which includes preprocessing and dimensionality reduction data analysis (principal component analysis, group averaging, calibration using reference images, normalization, Savitzky–Golay filtering). The pipeline structure allows the use of different dimensionality reduction methods: PCA (Jolliffe, 2002), ISOMAP (Balasubramanian, Schwartz, 2002) and UMAP (McInnes, et al., 2018) in combination with different classification methods: logistic regression (Norman, Harry, 2007), ridge regression (Norman, Harry, 2007), random forest (Ho, 1995), gradient boosting (Prokhorenkova et al., 2017).

Methods of plant image classification based on hyperspectral data

Hyperspectral images are used to classify the physiological state of plants. T. Zhang et al. (2018) investigated the feasibility of using hyperspectral imaging techniques in the visible and near-infrared ranges (VIS/NIR, 400–1,000 nm) to recognize viable and non-viable wheat seeds. For this purpose, classification models, partial least squares discriminant analysis (PLS-DA) and support vector machines (SVM) combined with some preprocessing techniques and sequential projection algorithm (SPA) were used. The results showed that the standard normal variation (SNV)-SPA-PLS-DA model had high classification accuracy for whole seeds (>85.2 %) and viable seeds (>89.5 %).

Y. Lu et al. (Lu et al., 2022) were able to achieve up to 99.6 % accuracy in differentiating five cannabis varieties, and 100 % accuracy in distinguishing between five growth stages and two plant organs (leaves and flowers) using a desktop hyperspectral imaging system in the spectral range of 400–1,000 nm and machine learning based on regularized linear discriminant analysis.

The work published by B.C. da Silva et al. (2024) evaluated the performance of five ML algorithms and the sensitivity of 90 spectra in the task of predicting the content of nitrogen and pigments (chlorophyll and carotenoids) in maize leaves at different phenological stages to optimize nitrogen fertilization. In predicting the contents of chlorophyll a and b, the value of Pearson correlation coefficient between predicted and observed data was about 0.6, and the mean absolute error (MAE) was below 0.5. When flavonoid content was predicted, the value of the correlation coefficient between predicted and observed data was about 0.6 and the MAE was 0.07. When nitrogen content was predicted, the correlation coefficient values were above 0.35 and the MAE was below 2.75.

In the paper published by Changyeun Mo et al. (2014), the authors developed a method to assess the viability of pepper (*Capsicum annuum* L.) seeds based on hyperspectral imaging in the 400–700 nm range obtained using blue, green, red and RGB LED illumination. For this purpose, a partial least squares discriminant analysis (PLS-DA) model was developed based on the standard normal variant of RGB LED illumination (400–700 nm), which provided recognition accuracies ranging from 96.7 to 100 %.

R. Falcioni et al. (2023) developed a method to estimate pigments such as chlorophylls, carotenoids, anthocyanins and flavonoids in six agronomic crops: maize, sugarcane, coffee, rapeseed, wheat and tobacco based on hyperspectral data. Clustering based on principal component analysis (PCA) and Kappa coefficient analysis yielded accuracies ranging from 92 to 100 % in the ultraviolet (UV-VIS), near-infrared (NIR) and shortwave infrared (SWIR) bands.

In our study, we obtained quite high precision values: accuracy = 0.979, F1 = 0.971 with precision = 0.944 and recall = 1.000 for all prediction models, which is comparable to similar values in other works, in particular, those described above. The resulting estimates were similar for all models, probably due to the fact that the sample size was small and homogeneous. As a result, with the resulting partitioning, all models in the test sample made one error, misclassifying one sample. On the other hand, this demonstrates the high stability of the predictions based on hyperspectral data and the proposed models.

In our previous work (Komyshev et al., 2023), we developed a method for estimating the presence of anthocyanins and melanin in barley grain shells based on the analysis of digital RGB images using computer vision and machine learning algorithms. We used a similar imaging protocol using Petri dishes for grains, but imaging was performed with a conventional RGB camera. The samples were taken from a similar collection. In that case, the best accuracy (accuracy = 0.821) was shown by the U-Net model based on the EfficientNetB0 topology. Thus, even when using deep machine learning methods, the classification accuracy was lower than in the present work. It can be concluded that more hyperspectral images allow more accurate classification of plant grains by pigment content using less resource-intensive "shallow" machine learning methods.

We studied the effect of the presence of anthocyanins on the accuracy of melanin determination in barley samples. The accuracy of melanin determination in samples contain-

ing anthocyanins was lower (accuracy = 0.95) compared to samples without this pigment (accuracy = 1) (Fig. 8). Thus, the presence of anthocyanins insignificantly reduces the accuracy of melanin determination in samples.

The ability to differentiate samples with only melanin from those with both melanin and anthocyanins was poor (Fig. 9*a*). Determination of anthocyanins, based on the hyperspectral data obtained, seems to be possible with high accuracy due to the spectrum in wavelengths falling in the red and infrared ranges (>714 nm) (Fig. 9*b*). Thus, this approach allows differentiating grains without pigments from grains with anthocyanins, but does not allow determining the presence of anthocyanins in samples with melanin.

Our goal was to explore the possibility of distinguishing between melanin-containing and non-melanin-containing seed samples using hyperspectral data alone. We also tested several approaches consisting of interchangeable methods that form a typical hyperspectral data processing pipeline and formed it into a software tool. This software tool can be used to quickly build a hyperspectral data analysis algorithm that includes the main data processing steps such as image loading, preprocessing, analysis and visualization.

Conclusion

Visualization based on the PCA, UMAP and ISOMAP methods, as well as clustering in dimension space 15, showed that barley samples with and without melanin could be divided into two respective classes with high accuracy on the basis of hyperspectral images. The analysis revealed statistically significant differences in the distribution of reflected intensity for these samples for all hyperspectral lines.

Advantages of using the developed pipeline over classical and more accurate biochemical methods of solving the classification problem are low time and labor costs, as well as objectivity of the obtained results. Neural networks/deep machine learning methods were not used in this version of the package for classification. The disadvantages of neural network approaches compared to the methods implemented in the pipeline may be the difficult interpretability of the prediction results, as well as the need for a training sample of a very large volume.

In this paper, an open-source Python-based computational pipeline has been developed for hyperspectral image analysis, which includes visualization in two-dimensional space, clustering, basic statistical analysis and classification. The proposed software package can significantly reduce the time cost in studies involving hyperspectral image analysis. The developed pipeline was tested in the task of investigating the effect of melanin on the hyperspectrum of barley grains.

References

Afonnikov D.A., Genaev M.A., Doroshkov A.V., Komyshev E.G., Pshenichnikova T.A. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments. *Russ. J. Genet.* 2016;52(7):688-701. DOI 10.1134/S1022795416070024]

Afonnikov D.A., Komyshev E.G., Efimov V.M., Genaev M.A., Koval V.S., Gierke P.U., Börner A. Relationship between the characteristics of bread wheat grains, storage time and germination. *Plants.* 2021;11(1):35. DOI 10.3390/plants11010035

Amanah H.Z., Wakholi C., Perez M., Faqeerzada M.A., Tunny S.S., Masithoh R.E., Choung M.G., Kim K.H., Lee W.H., Cho B.K. Near-infrared hyperspectral imaging (NIR-HSI) for nondestructive prediction of anthocyanins content in black rice seeds. *Appl. Sci.* 2021;11(11):4841. DOI 10.3390/app11114841

Ambrose A., Kandpal L.M., Kim M.S., Lee W.H., Cho B.K. High speed measurement of corn seed viability using hyperspectral imaging. *Infrared Phys. Technol.* 2016;75:173-179. DOI 10.1016/j.infrared.2015.12.008

Baek I., Kim M.S., Cho B.K., Mo C., Barnaby J.Y., McClung A.M., Oh M. Selection of optimal hyperspectral wavebands for detection of discolored, diseased rice seeds. *Appl. Sci.* 2019;9(5):1027. DOI 10.3390/app9051027

Balasubramanian M., Schwartz E.L. The isomap algorithm and topological stability. *Science.* 2002;295(5552):7. DOI 10.1126/science.295.5552.7a

Becht E., McInnes L., Healy J., Dutertre C.A., Kwok I.W., Ng L.G., Ginhoux F., Newell E.W. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol*. 2019;37(1):38-44. DOI 10.1038/nbt.4314

Cheshkova A.F. A review of hyperspectral image analysis techniques for plant disease detection and identification. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(2):202-213. DOI 10.18699/VJGB-22-25 (in Russian)]

Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. Introduction to Algorithms. Cambridge, Massachusetts: The MIT Press, 2022

da Silva B.C., de Mello Prado R., Baio F.H.R., Campos C.N.S., Teodoro L.P.R., Teodoro P.E., Santana D.C., Fernandes T.F.S., da Silva J.C.A., de Souza Loureiro E. New approach for predicting nitrogen and pigments in maize from hyperspectral data and machine learning models. *Remote Sens. Appl. Soc. Environ*. 2024;33:101110. DOI 10.1016/j.rsase.2023.101110

da Silva Medeiros M.L., Cruz-Tirado J.P., Lima A.F., de Souza Netto J.M., Ribeiro A.P.B., Bassegio D., Godoy H.T., Barbin D.F. Assessment oil composition and species discrimination of *Brassicas* seeds based on hyperspectral imaging and portable near infrared (NIR) spectroscopy tools and chemometrics. *J. Food Compos. Anal.* 2022;107:104403. DOI 10.1016/j.jfca.2022.104403

Dempster A.P., Laird N.M., Rubin D.B. Maximum likelihood from incomplete data via the *EM* algorithm. *J. R. Statist. Soc. B.* 1977;39(1):1-22. DOI 10.1111/j.2517-6161.1977.tb01600.x

Díaz-Martínez V., Orozco-Sandoval J., Manian V., Dhatt B.K., Walia H. A deep learning framework for processing and classification of hyperspectral rice seed images grown under high day and night temperatures. *Sensors.* 2023;23(9):4370. DOI 10.3390/s23094370

ElMasry G., Mandour N., Ejeez Y., Demilly D., Al-Rejaie S., Verdier J., Belin E., Rousseau D. Multichannel imaging for monitoring chemical composition and germination capacity of cowpea (*Vigna unguiculata*) seeds during development and maturation. *Crop J.* 2022;10(5):1399-1411. DOI 10.1016/j.cj.2021.04.010

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

Falcioni R., Antunes W.C., Demattê J.A.M., Nanni M.R. Reflectance spectroscopy for the classification and prediction of pigments in agronomic crops. *Plants.* 2023;12(12):2347. DOI 10.3390/plants12122347

Feng H., Chen G., Xiong L., Liu Q., Yang W. Accurate digitization of the chlorophyll distribution of individual rice leaves using hyperspectral imaging and an integrated image analysis pipeline. *Front. Plant Sci*. 2017;8:1238. DOI 10.3389/fpls.2017.01238

Flintham J., Adlam R., Bassoi M., Holdsworth M., Gale M. Mapping genes for resistance to sprouting damage in wheat. *Euphytica.* 2002;126:39-45. DOI 10.1023/A:1019632008244

Gao T., Chandran A.K.N., Paul P., Walia H., Yu H. HyperSeed: an end-to-end method to process hyperspectral images of seeds. *Sensors.* 2021;21(24):8184. DOI 10.3390/s21248184

Glagoleva A.Y., Shmakov N.A., Shoeva O.Y., Vasiliev G.V., Shatskaya N.V., Börner A., Afonnikov D.A., Khlestkina E.K. Metabolic pathways and genes identified by RNA-seq analysis of barley near-isogenic lines differing by allelic state of the *Black lemma* and pericarp (*Blp*) gene. *BMC Plant Biol.* 2017;17(Suppl. 1):182. DOI 10.1186/s12870-017-1124-1

Glagoleva A.Y., Novokreschyonov L.A., Shoeva O.Y., Kovaleva O.N., Khlestkina E.K. Studying grain color diversity in the barley collection of VIR. *Trudy po Prikladnoy Botanike, Genetike i Selektsii = Proceedings on Applied Botany, Genetics, and Breeding.* 2022; 183(3):76-84. DOI 10.30901/2227-8834-2022-3-76-84 (in Russian)]

Gowen A.A., O'Donnell C.P., Cullen P.J., Downey G., Frias J.M. Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends Food Sci. Technol.* 2007;18(12): 590-598. DOI 10.1016/j.tifs.2007.06.001

Greenwood P.E., Nikulin M.S. A Guide to Chi-Squared Testing. New York: Wiley, 1996;196-202

Hancock J.T., Khoshgoftaar T.M. CatBoost for big data: an interdisciplinary review. *J. Big Data.* 2020;7(1):94. DOI 10.1186/s40537-020-00369-8

Hao J., Ho T.K. Machine learning made easy: a review of *Scikit-learn* package in python programming language. *J. Educ. Behav. Stat.* 2019;44(3):348-361. DOI 10.3102/1076998619832248

He X., Feng X., Sun D., Liu F., Bao Y., He Y. Rapid and nondestructive measurement of rice seed vitality of different years using near-infrared hyperspectral imaging. *Molecules.* 2019;24(12):2227. DOI 10.3390/molecules24122227

Ho T.K. Random decision forests. In: Proceedings of the Third International Conference on Document Analysis and Recognition. 1995;1: 278-282. DOI 10.1109/ICDAR.1995.598994

Howse J. OpenCV Computer Vision with Python. Birmingham: Packt Publishing, 2013

Jin B., Qi H., Jia L., Tang Q., Gao L., Li Z., Zhao G. Determination of viability and vigor of naturally-aged rice seeds using hyperspectral imaging with machine learning. *Infrared Phys. Technol.* 2022; 122:104097. DOI 10.1016/j.infrared.2022.104097

Jolliffe I.T. Principal component analysis for special types of data. In: Principal Component Analysis. Springer Series in Statistics. New York, NY: Springer, 2002;338-372. DOI 10.1007/0-387-22440-8_13

Kandpal L.M., Lohumi S., Kim M.S., Kang J.S., Cho B.K. Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds. *Sens. Actuators B.* 2016;229:534-544. DOI 10.1016/j.snb.2016.02.015

Khlestkina E.K. Current applications of wheat and wheat-alien precise genetic stocks. *Mol. Breed.* 2014;34(2):273-281. DOI 10.1007/s11032-014-0049-8

Komyshev E.G., Genaev M.A., Busov I.D., Kozhekin M.V., Artemenko N.V., Glagoleva A.Y., Koval V.S., Afonnikov D.A. Determination of the melanin and anthocyanin content in barley grains by digital image analysis using machine learning methods. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(7):859-868. DOI 10.18699/VJGB-23-99 (in Russian)]

Krupnov V.A. Genetic complexity and context specificity of traits improving wheat yield under drought conditions. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2013;17(3):524-534 (in Russian)]

Lachman J., Martinek P., Kotíková Z., Orsák M., Šulc M. Genetics and chemistry of pigments in wheat grain. A review. *J. Cereal Sci*. 2017;74:145-154. DOI 10.1016/j.jcs.2017.02.007

Liu C., Huang W., Yang G., Wang Q., Li J., Chen L. Determination of starch content in single kernel using near-infrared hyperspectral images from two sides of corn seeds. *Infrared Phys. Technol.* 2020; 110:103462. DOI 10.1016/j.infrared.2020.103462

Lu Y., Young S., Linder E., Whipker B., Suchoff D. Hyperspectral imaging with machine learning to differentiate cultivars, growth stages, flowers, and leaves of industrial hemp (*Cannabis sativa* L.). *Front. Plant Sci*. 2022;12:810113. DOI 10.3389/fpls.2021.810113

Ma T., Tsuchikawa S., Inagaki T. Rapid and non-destructive seed viability prediction using near-infrared hyperspectral imaging coupled with a deep learning approach. *Comput. Electron. Agric.* 2020;177: 105683. DOI 10.1016/j.compag.2020.105683

Machálková L., Janečková M., Hřivna L., Dostálová Y., Hernandez K., Joany L., Mrkvicová E., Vyhnánek T., Trojan V. Impact of added colored wheat bran on bread quality. *Acta Univ. Agric. Silvic. Mendelianae Brun.* 2017;65(1):99-104. DOI 10.11118/actaun201765010099

Matzrafi M., Herrmann I., Nansen C., Kliper T., Zait Y., Ignat T., Siso D., Rubin B., Karnieli A., Eizenberg H. Hyperspectral technologies for assessing seed germination and trifloxysulfuron-methyl response in *Amaranthus palmeri* (Palmer amaranth). *Front. Plant Sci.* 2017;8:474. DOI 10.3389/fpls.2017.00474

McInnes L., Healy J., Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. *ArXiv.* 2018;1802. 03426. DOI 10.48550/arXiv.1802.03426

Mo C., Kim G., Lee K., Kim M.S., Cho B.K., Lim J., Kang S. Non-destructive quality evaluation of pepper (*Capsicum annuum* L.) seeds using LED-induced hyperspectral reflectance imaging. *Sensors*. 2014;14(4):7489-7504. DOI 10.3390/s140407489

Norman R.D., Harry S. Applied Regression Analysis. Williams, 2007

Nunez-Iglesias J., Van der Walt S., Dashnow H. Elegant SciPy: The Art of Scientific Python. Sebastopol, CA: O'Reilly Media, 2017

Prokhorenkova L., Gusev G., Vorobev A., Dorogush A.V., Gulin A. CatBoost: unbiased boosting with categorical features. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc., 2018; 6639-6649

Qin J., Chao K., Kim M.S., Lu R., Burks T.F. Hyperspectral and multispectral imaging for evaluating food safety and quality. *J. Food Eng.* 2013;118(2):157-171. DOI 10.1016/j.jfoodeng.2013.04.001

Reddy P., Panozzo J., Guthridge K.M., Spangenberg G.C., Rochfort S.J. Single seed near-infrared hyperspectral imaging for classification of perennial ryegrass seed. *Sensors.* 2023;23(4):1820. DOI 10.3390/s23041820

Savitzky A., Golay M.J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 1964;36(8):1627-1639. DOI 10.1021/ac60214a047

Serrano L., Filella I., Penuelas J. Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Sci*. 2000; 40(3):723-731. DOI 10.2135/cropsci2000.403723x

Shoeva O.Yu., Strygina K.V., Khlestkina E.K. Genes determining the synthesis of flavonoid and melanin pigments in barley. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2018;22(3):333-342. DOI 18699/VJ18.369 (in Russian)]

Stančin I., Jović A. An overview and comparison of free Python libraries for data mining and big data analysis. In: 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2019;977-982. DOI 10.23919/MIPRO.2019.8757088

Wakholi C., Kandpal L.M., Lee H., Bae H., Park E., Kim M.S., Mo C., Lee W.H., Cho B.K. Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics. *Sens. Actuators B.* 2018;255:498-507. DOI 10.1016/j.snb.2017.08.036

Weber V.S., Araus J.L., Cairns J.E., Sanchez C., Melchinger A.E., Orsini E. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. *Field Crops Res.* 2012;128:82-90. DOI 1016/j.fcr.2011.12.016

Wilcoxon F. Individual comparisons by ranking methods. In: Kotz S., Johnson N.L. (Eds.). Breakthroughs in Statistics. Springer Series in Statistics. New York, NY: Springer, 1992;196-202. DOI 10.1007/978-1-4612-4380-9_16

Yang G., Wang Q., Liu C., Wang X., Fan S., Huang W. Rapid and visual detection of the main chemical compositions in maize seeds based on Raman hyperspectral imaging. *Spectrochim. Acta A. Mol. Biomol. Spectrosc*. 2018;200:186-194. DOI 10.1016/j.saa.2018.04.026

Yoosefzadeh-Najafabadi M., Earl H.J., Tulpan D., Sulik J., Eskandari M. Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean. *Front. Plant Sci.* 2021;11:624273. DOI 10.3389/fpls.2020.624273

Zahavi A., Palshin A., Liyanage D.C., Tamre M. Influence of illumination sources on hyperspectral imaging. In: 20th International Conference on Research and Education in Mechatronics (REM). Wels, Austria, 2019;1-5. DOI 10.1109/REM.2019.8744086

Zhang X., He Y. Rapid estimation of seed yield using hyperspectral images of oilseed rape leaves. *Ind. Crops Prod.* 2013;42:416-420. DOI 10.1016/j.indcrop.2012.06.021

Zhang T., Wei W., Zhao B., Wang R., Li M., Yang L., Wang J., Sun Q. A reliable methodology for determining seed viability by using hyperspectral data from two sides of wheat seeds. *Sensors*. 2018; 18(3):813. DOI 10.3390/s18030813

Zhu F., Qiao X., Zhang Y., Jiang J. Analysis and mitigation of illumination influences on canopy close-range hyperspectral imaging for the *in situ* detection of chlorophyll distribution of basil crops. *Comput. Electron. Agric*. 2024;217:108553. DOI 10.1016/j.compag.2023.108553

Žibrat U., Susič N., Knapič M., Širca S., Strajnar P., Razinger J., Vončina A., Urek G., Stare B.G. Pipeline for imaging, extraction, preprocessing, and processing of time-series hyperspectral data for discriminating drought stress origin in tomatoes. *MethodsX*. 2019; 6:399-408. DOI 10.1016/j.mex.2019.02.022

Zou Z., Chen J., Wu W., Luo J., Long T., Wu Q., Wang Q., Zhen J., Zhao Y., Wang Y., Chen Y., Zhou M., Xu L. Detection of peanut seed vigor based on hyperspectral imaging and chemometrics. *Front. Plant Sci.* 2023;14:1127108. DOI 10.3389/fpls.2023.1127108

Conflict of interest. The authors declare no conflict of interest.
Received July 23, 2023. Revised March 22, 2024. Accepted March 26, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(4):456-462

DOI 10.18699/vjgb-24-51

The BLUP method in evaluation of breeding values of Russian spring wheat lines using micro- and macroelements in seeds

N.A. Potapova [1, 2, 3#], A.S. Zlobin [1#], I.N. Leonova [4], E.A. Salina [1], Y.A. Tsepilov [1]

[1] Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
[2] Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Moscow, Russia
[3] Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical-Biological Agency, Moscow, Russia
[4] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
nadezhdalpotapova@gmail.com; tsepilov@bionet.nsc.ru

Abstract. Genomic selection is a technology that allows for the determination of the genetic value of varieties of agricultural plants and animal breeds, based on information about genotypes and phenotypes. The measured breeding value (BV) for varieties and breeds in relation to the target trait allows breeding stages to be thoroughly planned and the parent forms suitable for crossing to be chosen. In this work, the BLUP method was used to assess the breeding value of 149 Russian varieties and introgression lines (4 measurements for each variety or line, 596 phenotypic points) of spring wheat according to the content of seven chemical elements in the grain – K, Ca, Mg, Mn, Fe, Zn, Cu. The quality of the evaluation of breeding values was assessed using cross-validation, when the sample was randomly divided into five parts, one of which was chosen as a test population. The following average values of the Pearson correlation were obtained for predicting the concentration of trace elements: K – 0.67, Ca – 0.61, Mg – 0.4, Mn – 0.5, Fe – 0.38, Zn – 0.46, Cu – 0.48. Out of the 35 models studied, the *p*-value was below the nominal significant threshold (p-value < 0.05) for 28 models. For 11 models, the p-value was significant after correction for multiple testing (p-value < 0.001). For Ca and K, four out of five models and for Mn two out of five models had a *p*-value below the threshold adjusted for multiple testing. For 30 varieties that showed the best varietal values for Ca, K and Mn, the average breeding value was 296.43, 785.11 and 4.87 mg/kg higher, respectively, than the average breeding value of the population. The results obtained show the relevance of the application of genomic selection models even in such limited-size samples. The models for K, Ca and Mn are suitable for assessing the breeding value of Russian wheat varieties based on these characteristics.

Key words: genomic selection; BLUP; wheat; microelements; macroelements.

For citation: Potapova N.A., Zlobin A.S., Leonova I.N., Salina E.A., Tsepilov Y.A. The BLUP method in evaluation of breeding values of Russian spring wheat lines using micro- and macroelements in seeds. *Vavilovskii Zhurnal Genetiki i Selektsii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(4):456-462. DOI 10.18699/vjgb-24-51

Funding. The research was carried out at the expense of the grant from the Russian Science Foundation No. 23-16-00041 (https://rscf.ru/project/23-16-00041/).

Acknowledgements. The plant material was grown at the Plant Reproduction Center of the ICG SB RAS with the financial support of the budget project FWNR-2022-0017.

Использование метода BLUP для оценки селекционной ценности образцов мягкой яровой пшеницы по содержанию микро- и макроэлементов в зерне

Н.А. Потапова [1, 2, 3#], А.С. Злобин [1#], И.Н. Леонова [4], Е.А. Салина [1], Я.А. Цепилов [1]

[1] Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия
[2] Институт проблем передачи информации им. А.А. Харкевича Российской академии наук, Москва, Россия
[3] Федеральный научно-клинический центр физико-химической медицины им. академика Ю.М. Лопухина Федерального медико-биологического агентства, Москва, Россия
[4] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
nadezhdalpotapova@gmail.com; tsepilov@bionet.nsc.ru

Аннотация. Геномная селекция – это технология, позволяющая определять генетическую ценность сортов сельскохозяйственных растений и пород животных, опираясь на информацию о генотипах и фенотипах. Измеренная селекционная ценность по отношению к целевому признаку дает возможность грамотно планировать этапы се-

# Equal contribution to the study

лекции и выбирать подходящие для скрещивания родительские формы. В настоящей работе использован метод BLUP для оценки селекционной ценности 149 российских сортов и интрогрессивных линий (4 измерения для каждого сорта или линии, 596 фенотипических точек) яровой пшеницы по содержанию семи химических элементов в зерне – K, Ca, Mg, Mn, Fe, Zn, Cu. Качество оценки селекционной ценности было определено с помощью кросс-валидации методом случайного разделения выборки на пять частей, одна из которых выступала в качестве тестовой популяции. Средние значения коэффициента корреляции Пирсона для предсказания концентрации микроэлементов составили: K – 0.67, Ca – 0.61, Mg – 0.4, Mn – 0.5, Fe – 0.38, Zn – 0.46, Cu – 0.48. Для 28 из 35 исследуемых моделей значение *p*-value было ниже номинального значимого порога (*p*-value < 0.05). Для 11 моделей *p*-value было значимо после коррекции на множественное тестирование (*p*-value < 0.001). Четыре из пяти моделей для Ca и K, и две из пяти для Mn имели *p*-value ниже порога, поправленного на множественное тестирование. Для 30 сортов, показавших лучшие значения сортовой ценности, средняя селекционная ценность для Ca, K и Mn была выше на 296.43, 785.11 и 4.87 мг/кг соответственно, чем средняя селекционная ценность популяции. Полученные результаты демонстрируют возможность применения моделей геномной селекции на ограниченных по размеру выборках образцов. Модели для K, Ca и Mn, показавшие наилучший результат, пригодны для оценки селекционной ценности российских сортов пшеницы для данных признаков.
Ключевые слова: геномная селекция; BLUP; пшеница; микроэлементы; макроэлементы.

Introduction

Since time immemorial, classical selection methods have been used to develop new varieties of agricultural plants and animal breeds based on the hybridization of samples with economically valuable traits, followed by selection based on phenotype. With the evolution of genome sequencing technologies and methods for developing molecular markers, it became possible to use differences in the structure of genomes, find marker-trait associations, and use the information obtained to identify the relationship between genotypic polymorphisms and phenotypic variations. New approaches actively being developed in plants include marker-assisted selection (MAS) and genomic selection (GS) (Charmet, Storlie, 2012; Bhat et al., 2016; Bartholomé et al., 2022; Miller et al., 2023); they are also used for animals (Kuznetsov, 1999; Melucci et al., 2009; Suslina et al., 2019; Stolpovsky et al., 2020; Zhumanov et al., 2021; Johnsson, 2023).

Despite the fact that MAS is quite effective for searching for and introducing genes with a high contribution to the phenotypic manifestation of a trait, the main disadvantage of the method is the insufficient accuracy of predicting traits with polygenic inheritance. As an alternative to MAS, and for the purpose of overcoming the limitations of this method, genomic selection has been proposed. One of its main advantages is the use of predictive models to assess the breeding value (BV).

Among the main approaches for genomic selection the following methods can be distinguished: BLUP methods (Best Linear Unbiased Prediction) (Charmet, Storlie, 2012; Hoffstetter et al., 2016; Lozada, Carter, 2020; Plavšin et al., 2022), a group of Bayesian methods (Juliana et al., 2022) and a group of methods that use machine learning (Wang et al., 2018). Index methods are also used (Lopez-Cruz et al., 2020). Genomic selection methods in agriculture can change and increase the accuracy of the approach to breeding new plant varieties and animal breeds.

Significant benefits of the use of GS have been observed in livestock production due to the high cost of reproduction. The adoption of GS for crop production began much later, but the general potential of current approaches and the potential of GS itself has now been explored in major crops such as wheat, corn, barley and soybean.

Wheat plays an important role in food security around the world. In addition to nutritional components, wheat grains contain elements such as calcium, zinc, magnesium, etc. Elemental deficiency, also known as "hidden hunger," occurs as a result of consuming food with low concentrations of elements and vitamins (Liu et al., 2019) and can lead to various diseases and even death. For this reason, the ability to intelligently select wheat varieties and increase the concentration of essential elements in the grain is an important way to combat nutritional deficiencies around the world. According to the elements' content in the human body, elements are divided into macroelements (the content in the human body is hundredths of a percent or more), microelements (the content is from hundred thousandths to thousandths of a percent) and ultramicroelements (millionths of a percent or less). Of the seven elements analyzed in this paper, macroelements include calcium, potassium and magnesium, and microelements include iron, manganese, copper and zinc.

Genomic selection methods are applied to different populations and varieties of wheat for a variety of traits: from grain element content and yield to disease resistance (Hoffstetter et al., 2016). The most actively used method is BLUP and its various modifications – rrBLUP, gBLUP, egBLUP, wBLUP and others (Zhao et al., 2014; Martini et al., 2017; Berkner et al., 2022; Rabieyan et al., 2022). This method has proven itself over several decades of use in plant and animal breeding. The Bayesian method and its modifications are also used – BRR (Bayesian Ridge Regression), BL (Bayesian Lasso), BA (Bayes A), BB (Bayes B), BC (Bayes C). Recently, machine learning and deep learning methods have been applied to genomic breeding of wheat (Sandhu et al., 2021a, b; Sirsat et al., 2022). Comparisons of results between methods showed that they mainly overlap and the generally accepted methods of the BLUP group are in no way inferior to Bayesian methods and machine learning methods (Tsai et al., 2020; Berkner et al., 2022; Juliana et al., 2022).

Previously, using a panel of varieties and introgressive lines of bread wheat, we performed associative mapping of genetic factors that determine the content of seven chemical elements in wheat grain (Potapova et al., 2023).

The purpose of this work was to study the BV of samples from this collection based on the content of chemical ele-

ments and to obtain unbiased estimates of the effects of genetic polymorphisms to determine the BV of other Russian varieties.

Materials and methods

A panel of 157 Russian varieties and introgressive lines of spring soft wheat was used in the work. A list of plant material, information about the origin of samples and phenotyping conditions are available in (Leonova et al., 2020; Potapova et al., 2023).

Genotyping of samples with SNP markers was performed using the Illumina Infinium 15 K platform (TraitGenetics Section, Germany, www.sgs-institut-fresenius.de). After alignment of markers to the reference wheat genome to determine their location (chromosome and position), quality control checks and subsequent data filtering (quality of SNP genotyping < 5 %, minor allele frequency < 1 %, quality of sample genotyping < 5 %), 149 wheat lines and 11,405 SNPs (single nucleotide polymorphisms) remained. Detailed information about the analysis of genotypes is available in (Potapova et al., 2023).

The content of micro- and macroelements (Zn, Mg, Mn, Ca, Cu, Fe and K) was determined by flame spray atomic absorption spectrometry on a Contra 800 D device (Analytik Jena, Germany), as described in (Potapova et al., 2023). Statistical processing of the results was carried out using the Statistica v.10.0 software package.

The assessment of the chemical elements content was carried out using seed material from the collection grown under the conditions of 2018–2019. At the same time, 4 measurements were performed, 2 for each year. Heritability was calculated using the formula V(G)/V(P) in the plink program (v.1.90b6.26, Purcell et al., 2007).

For each of these phenotypes (element content in grain), average values were obtained among 4 measurement points. The obtained average values were used in further analysis.

BLUP and cross-validation. Each element was analysed separately, analysis flowchart is presented in Figure 1.

For 149 cultivars, the genetic relationship matrix for SNPs was estimated with genotyping quality above 98 % using GCTA software (version 1.94) (Yang et al., 2011). Breeding values were then obtained for each of the accessions using the --reml-pred-rand option from GCTA. This function estimates the variance of a trait explained by all analyzed SNPs. To estimate the weight of each SNP (SNP coefficients) separately, the --blup-snp function from GCTA was used.

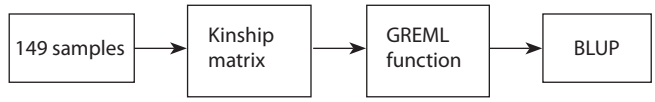


Fig. 1. Analysis flowchart.

To check the validity of the obtained results, we applied a model using the k-fold cross-validation method (Fig. 2).

The sample of 149 units was randomly divided into five subsamples. Each of the subsamples acted once as a test population, while the remaining four subsamples acted as a training population. Thus, for each of the chemical elements, five models were used to obtain SNP coefficients to assess the breeding value of the element content in wheat. In each case, there were 119 varieties in the training set, and 30 varieties in the test set.

Breeding value assessment was carried out using GCTA software as described above. Next, the coefficients of single nucleotide polymorphisms were obtained to assess the breeding value of the content of elements in wheat using the BLUP method implemented in the GCTA software.

For the obtained coefficients, the BV was estimated for all accessions from the test population using plink software (version 1.90b6.26) (Purcell et al., 2007). The obtained values were used to assess the quality of the prediction by calculating the correlation coefficients between the estimated BV and the actual phenotypic data. Confidence intervals for the values of correlation coefficients were also estimated using the Fisher *z*-transformation of the distribution.

For visualization, we used R programming language (version 2022.07.0, build 548). The regression line was constructed using the formula BV ~ phenotype, where BV is the estimated breeding value, and phenotype is the real values of the phenotypes.

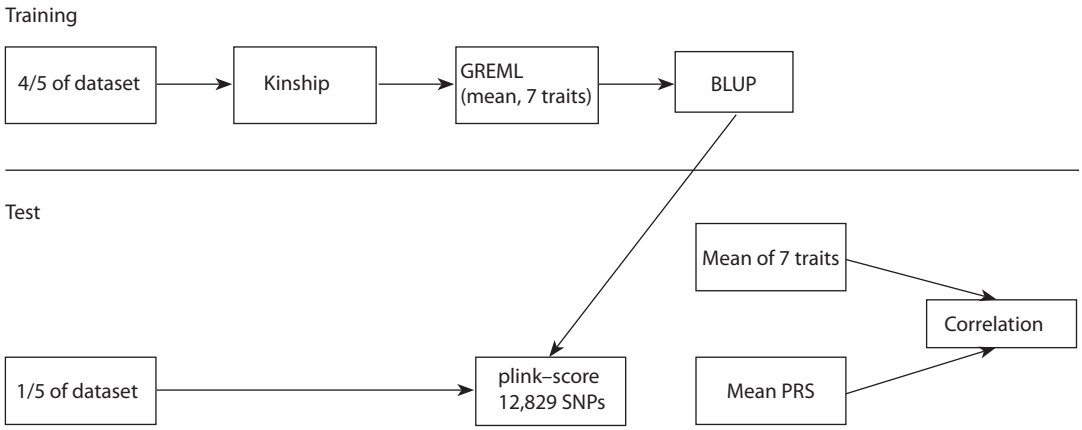


Fig. 2. Analysis flowchart using the k-block cross-validation method.

Results

Table 1 and Table 2 show the phenotypic mean and the heritability of each element in 2018 and 2019, respectively. A graph of correlations between phenotypic means is presented in Figure 3.

Data on estimates of BV and mean content of seven elements for each variety are presented in Supplementary Material 1[1].

Data on the estimated correlation coefficients and confidence intervals, as well as the *p*-value between the estimated BV and the actual phenotypic data for all seven elements studied, are presented in Supplementary Material 2.

Average values of Pearson correlation coefficients were obtained to predict the concentration of microelements with real phenotypes: K – 0.67, Ca – 0.61, Mg – 0.4, Mn – 0.5, Fe – 0.38, Zn – 0.46, Cu – 0.48. The maximum correlation coefficient was 0.75 (*p*-value = 1.85e-07) and was obtained for model 4 for potassium. The minimum is 0.22 for model 5 for iron (*p*-value = 0.24).

It was assumed that the prediction of BV for an element is significant if for at least one out of the five models the *p*-value is below the threshold adjusted for multiple testing (*p*-value < 0.001), and for the remaining four models the *p*-value is below the nominal level of significance (*p*-value < 0.05). Thus, we obtained significant estimates of the BV for calcium, potassium and manganese.

The absolute values of the correlation coefficient for the other four micro- and macroelements (Fe, Mg, Zn, Cu) and models were included in the estimated confidence intervals of each model for each of the studied elements, and were also significantly different from zero for 28 out of the 35 estimated models. For iron, in three out of the five models (models numbered 1, 4, 5), the *p*-values were above the nominal significance level of 0.05. Also, correlation coefficient values insignificant at the *p*-value level were obtained for model 3 for copper, model 2 for magnesium, model 4 for manganese and model 1 for zinc. The resulting scatterplots are presented in Supplementary Materials 3–9.

For the 30 varieties with the highest estimated BV, the response to selection was assessed (compared with the average values of BV for the population) (Table 3). A comparison was carried out for 30 varieties with the highest values of micro- and macroelements. Response to selection for phenotypes was adjusted for heritability. Only for calcium, the response to selection obtained while accounting for the BV was higher than the response to selection obtained for phenotypes while taking into account heritability (Table 3). The response to selection was estimated as (Ptop – Pmean)*h^2, where Ptop is the average value of the phenotype for 30 varieties with the highest value of the estimated BV, Pmean is the average value of the phenotype in the study population, h^2 is the heritability indicator of this phenotype.

The resulting estimates of the breeding value of Russian wheat varieties in the form of coefficients for SNP were registered in the Unified Register of Russian Programs for Electronic Computers and Databases and are available upon request to the copyright holder (Institute of Cytology and Genetics SB RAS) (Supplementary Material 3).

[1] Supplementary Materials 1–10 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Potapova_Engl_28_4.pdf

Table 1. Phenotypic mean values for each year separately (2018 and 2019) and averaged between years, the standard error is indicated in parentheses

Element	Mean element content, mg/kg		
	2018	2019	2018 and 2019
Ca	772.67 (18.03)	666.68 (17.53)	719.67 (12.74)
Cu	3.58 (0.06)	4.41 (0.05)	3.99 (0.04)
Fe	43.17 (0.47)	46.61 (0.45)	44.89 (0.33)
K	4025.70 (37.10)	4186.73 (48.39)	4106.21 (30.63)
Mg	1588.54 (11.44)	1329.75 (11.59)	1459.14 (9.64)
Mn	37.83 (0.31)	39.05 (0.46)	38.44 (0.28)
Zn	35.89 (0.40)	46.27 (0.64)	41.08 (0.43)

Table 2. Heritability of traits by year (in fractions of one, where zero is the absence of a genetic contribution to the trait, one is a completely genetically determined trait) and for the average between years, the standard error is indicated in parentheses

Element	Heritability of element content		
	2018	2019	2018 and 2019
Ca	0.66 (0.05)	0.53 (0.06)	0.70 (0.04)
Cu	0.79 (0.04)	0.82 (0.03)	0.81 (0.03)
Fe	0.89 (0.02)	0.75 (0.04)	0.84 (0.03)
K	0.72 (0.05)	0.83 (0.03)	0.89 (0.02)
Mg	0.73 (0.05)	0.55(0.06)	0.72 (0.04)
Mn	0.85 (0.03)	0.76 (0.04)	0.83 (0.03)
Zn	0.82 (0.03)	0.67(0.05)	0.73 (0.04)

Discussion

In this work, we conducted a study of unbiased estimates of the effects of genetic polymorphisms and their use to assess the genomic potential of Russian spring bread wheat samples for the content of seven micro- and macroelements – K, Ca, Mg, Mn, Fe, Zn, Cu. The best linear unbiased prediction (BLUP) was chosen as a method, and an approach of dividing the sample into several parts (*k*-fold cross-validation) was selected to check the quality of the model. The choice of model and method was due to the wide dissemination and application of them in genomic selection of plants and animals (Piepho et al., 2008; Molenaar et al., 2018; Tajalifar, Rasooli, 2022).

The sample was randomly divided into five subsamples. Correlation was used as a quality metric for the obtained SNP coefficients to assess the BV. The minimum correlation coefficient value was 0.22 for model 5 for iron content (*p*-value = 0.24). At the same time, the *p*-values of model 5

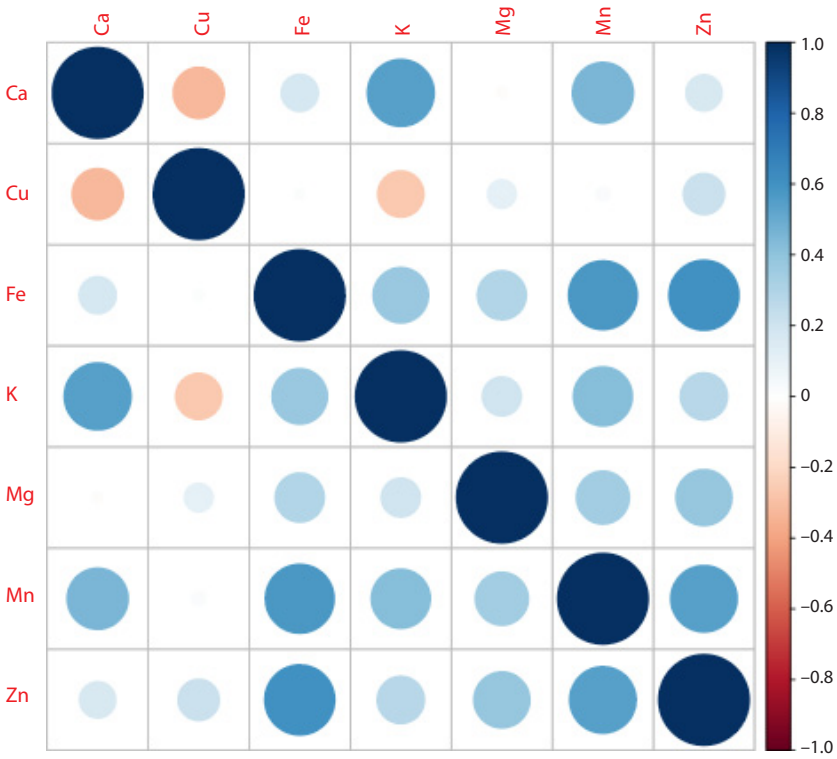


Fig. 3. Plot of correlations between phenotypes for 149 varieties.
For each variety, the phenotype value was calculated as the average between the four points.

Table 3. Expected response to selection using 30 cultivars with the highest estimated BV and 30 cultivars with the highest content of micro- and macronutrients as parent population

Element	Response to selection by BV	Response to selection by phenotypes (accounting for heritability)
Ca	296.43	262.24
Cu	0.60	0.90
Fe	4.99	7.25
K	785.11	873.41
Mg	88.57	131.66
Mn	4.87	5.95
Zn	6.79	7.54

for all elements were higher than the nominal value of 0.05 for iron in only one case out of seven. Moreover, out of 35 *p*-values obtained for the correlation coefficients of the estimated BV and real phenotypes, only 7 were equal to or above the nominal significance level of 0.05. This indicates a stable estimate of BV between different parts of the sample.

It is worth noting that for calcium, potassium, and magnesium, at least one out of the five models had a correlation coefficient that was significant using the threshold adjusted for multiple testing (p-value < 0.001), and the remaining models were significant using the nominal significance level (p-value < 0.05). Based on this, we established that the estimates of breeding value for these three elements are significant. The lack of significance according to a given criterion for the remaining four elements can be observed due to many factors, such as small sample size, heterogeneity of the selected population according to the estimated BV, etc. We also measured confidence intervals for each obtained value of the correlation coefficient. For each of the seven elements studied, all correlation coefficient values for all models were within the estimated confidence intervals.

One of the advantages of using genomic selection, and using BLUP in particular, is the ability to evaluate the expected increase in a trait in the next generation (response to selection). We assessed breeding differentials and response to selection for 30 varieties with the highest BV values and 30 varieties with the highest content of micro- and macroelements in wheat. In the case of selection based on BV the expected response to selection is comparable to the expected response to selection based on phenotypes. This statement is appliable in the case when the response to selection based on phenotypes is weighted by the heritability of the trait according to the breeder's equation. In our study we showed higher response for selection based on BV for calcium. The obtained high values of the selection differential for selection by phenotypes may be associated with high heritability and heterogeneity in the distribution of phenotypes in the studied population.

Previously, we conducted a genome-wide association study for seven micro- and macroelements in varieties and introgression lines of wheat (Potapova et al., 2023), and identified four significant loci. One of them was associated with the content of potassium and calcium, two with the content of iron and manganese, and one with all the studied elements. The results of this work demonstrate that, indeed, by using data from wheat accessions, it is possible to obtain estimated BV numbers for predicting calcium and potassium content (for calcium and potassium, all p-values obtained were less than the nominally significant threshold of 0.05). However, for three out of five models for iron and one out of five for manganese, the p-values exceeded the nominally significant threshold. This may be due to a limited sample size or many other factors, such as the complex genetic structure of a trait (for example, polygenicity or pleiotropy), insufficient data for prediction (number of varieties or SNPs), etc. At this time, there is a lack of scientific publications analyzing the breeding value of varieties for the content of the elements we studied. In this regard, it is difficult to compare our results with previous ones.

The main limitation of this work is a relatively small sample size. There are currently no reliable estimates of what minimum sample size is needed to create genomic selection models. In this article, we empirically showed that it makes sense to carry out such studies even on small samples (149 varieties with four measurements for each, a total of 596 phenotypic points). It is expected that as the sample size increases, the quality of the models will also increase. The second limitation of our work is the use of microarray genotyping data to construct models. If the genotyping test data are obtained using another array or technology, the model we used will most likely be inapplicable due to low overlap in polymorphisms. The use of genetic imputation methods can potentially solve this problem and increase genotyping coverage (Nyine et al., 2019; Song et al., 2019; Munyengwa et al., 2021; Bonnett et al., 2022; Kriaridou et al., 2023), and testing these methods on wheat is the scope of future work in this direction.

Conclusion

Thus, in this work, estimates of the BV were obtained for Russian wheat varieties, regarding the content of seven chemical elements in the grain (K, Ca, Mg, Mn, Fe, Zn, Cu). Our results can be useful primarily for breeders when carrying out work on the selection and breeding of varieties with a high content of micro- and macroelements in the grain. Using the values of the estimated BV, it becomes possible to rank and select the best samples from the populations under study. Additionally, this work can be methodologically useful in creating models for genomic selection of other agricultural plants. Also, these assessments can be used in practice developing breeding schemes and directly breeding new varieties.

References

Bartholomé J., Prakash P.T., Cobb J.N. Genomic prediction: progress and perspectives for rice improvement. In: Ahmadi N., Bartholomé J. (Eds.). Genomic Prediction of Complex Traits. Methods in Molecular Biology. V. 2467. New York: Humana, 2022;569-617. DOI 10.1007/978-1-0716-2205-6_21

Berkner M.O., Schulthess A.W., Zhao Y., Jiang Y., Oppermann M., Reif J.C. Choosing the right tool: Leveraging of plant genetic resources in wheat (*Triticum aestivum* L.) benefits from selection of a suitable genomic prediction model. *Theor. Appl. Genet.* 2022; 135(12):4391-4407. DOI 10.1007/s00122-022-04227-4

Bhat J.A., Ali S., Salgotra R.K., Mir Z.A., Dutta S., Jadon V., Tyagi A., Mushtaq M., Jain N., Singh P.K., Singh G.P., Prabhu K.V. Genomic selection in the era of next generation sequencing for complex traits in plant breeding. *Front. Genet.* 2016;7:221. DOI 10.3389/fgene.2016.00221

Bonnett D., Li Y., Crossa J., Dreisigacker S., Basnet B., Pérez-Rodríguez P., Alvarado G., Jannink J.L., Poland J., Sorrells M. Response to early generation genomic selection for yield in wheat. *Front. Plant Sci.* 2022;12:718611. DOI 10.3389/fpls.2021.718611

Charmet G., Storlie E. Implementation of genome-wide selection in wheat. *Russ. J. Genet. Appl. Res*. 2012;2(4):298-303. DOI 10.1134/S207905971204003X

Hoffstetter A., Cabrera A., Huang M., Sneller C. Optimizing training population data and validation of genomic selection for economic traits in soft winter wheat. *G3 (Bethesda)*. 2016;6(9):2919-2928. DOI 10.1534/g3.116.032532

Johnsson M. Genomics in animal breeding from the perspectives of matrices and molecules. *Hereditas.* 2023;160(1):20. DOI 10.1186/s41065-023-00285-w

Juliana P., He X., Marza F., Islam R., Anwar B., Poland J., Shrestha S., Singh G.P., Chawade A., Joshi A.K., Singh R.P., Singh P.K. Genomic selection for wheat blast in a diversity panel, breeding panel and full-sibs panel. *Front. Plant Sci.* 2022;12:745379. DOI 10.3389/fpls.2021.745379

Kriaridou C., Tsairidou S., Fraslin C., Gorjanc G., Looseley M.E., Johnston I.A., Houston R.D., Robledo D. Evaluation of low-density SNP panels and imputation for cost-effective genomic selection in four aquaculture species. *Front. Genet.* 2023;14:1194266. DOI 10.3389/fgene.2023.1194266

Kuznetsov V.M. The best linear unbiased forecast of the breeding value of roosters by the quality of offspring. *Vestnik Rossiiskoy Akademii Sel'skokhozyaystvennykh Nauk = Vestnik of the Russian Academy of Agricultural Sciences*. 1999;2:61-63 (in Russian)

Leonova I.N., Skolotneva E.S., Orlova E.A., Orlovskaya O.A., Salina E.A. Detection of genomic regions associated with resistance to stem rust in Russian spring wheat varieties and breeding germplasm. *Int. J. Mol. Sci.* 2020;21(13):4706. DOI 10.3390/ijms21134706

Liu J., Wu B., Singh R.P., Velu G. QTL mapping for micronutrients concentration and yield component traits in a hexaploid wheat mapping population. *J. Cereal Sci.* 2019;88:57-64. DOI 10.1016/j.jcs.2019.05.008

Lopez-Cruz M., Olson E., Rovere G., Crossa J., Dreisigacker S., Mondal S., Singh R., Campos G.L. Regularized selection indices for breeding value prediction using hyper-spectral image data. *Sci. Rep.* 2020;10(1):8195. DOI 10.1038/s41598-020-65011-2

Lozada D.N., Carter A.H. Genomic selection in winter wheat breeding using a recommender approach. *Genes.* 2020;11(7):779. DOI 10.3390/genes11070779

Martini J.W.R., Gao N., Cardoso D.F., Wimmer V., Erbe M., Cantet R.J., Simianer H. Genomic prediction with epistasis models: on the marker-coding-dependent performance of the extended GBLUP and properties of the categorical epistasis model (CE). *BMC Bioinformatics*. 2017;18(1):3. DOI 10.1186/s12859-016-1439-1

Melucci L.M., Birchmeier A.N., Cappa E.P., Cantet R.J. Bayesian analysis of selection for greater weaning weight while maintaining birth weight in beef cattle. *J. Anim. Sci.* 2009;87(10):3089-3096. DOI 10.2527/jas.2009-1801

Miller M.J., Song Q., Fallen B., Li Z. Genomic prediction of optimal cross combinations to accelerate genetic improvement of soybean (*Glycine max*). *Front. Plant Sci.* 2023;14:1171135. DOI 10.3389/fpls.2023.1171135

Molenaar H., Boehm R., Piepho H.-P. Phenotypic selection in ornamental breeding: It's better to have the BLUPs than to have the BLUEs. *Front. Plant Sci.* 2018;9:1511. DOI 10.3389/fpls.2018.01511

Munyengwa N., Le Guen V., Bille H.N., Souza L.M., Clément-Demange A., Mournet P., Masson A., Soumahoro M., Kouassi D., Cros D. Optimizing imputation of marker data from genotyping-by-sequencing (GBS) for genomic selection in non-model species: Rubber tree (*Hevea brasiliensis*) as a case study. *Genomics.* 2021; 113(2):655-668. DOI 10.1016/j.ygeno.2021.01.012

Nyine M., Wang S., Kiani K., Jordan K., Liu S., Byrne P., Haley S., Baenziger S., Chao S., Bowden R., Akhunov E. Genotype imputation in winter wheat using first-generation haplotype map SNPs improves genome-wide association mapping and genomic prediction of traits. *G3 (Bethesda).* 2019;9(1):125-133. DOI 10.1534/g3.118.200664

Piepho H.P., Möhring J., Melchinger A.E., Büchse A. BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica.* 2008;161:209-228. DOI 10.1007/s10681-007-9449-8

Plavšin I., Gunjača J., Galić V., Novoselović D. Evaluation of genomic selection methods for wheat quality traits in biparental populations indicates inclination towards parsimonious solutions. *Agronomy.* 2022;12(5):1126. DOI 10.3390/agronomy12051126

Potapova N.A., Timoshchuk A.N., Tiys E.S., Vinichenko N.A., Leonova I.N., Salina E.A., Tsepilov Y.A. Multivariate genome-wide association study of concentrations of seven elements in seeds reveals

four new loci in Russian wheat lines. *Plants.* 2023;12(17):3019. DOI 10.3390/plants12173019

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., de Bakker P.I., Daly M.J., Sham P.C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 2007;81(3):559-575. DOI 10.1086/519795

Rabieyan E., Bihamta M.R., Moghaddam M.E., Mohammadi V., Alipour H. Genome-wide association mapping and genomic prediction of agronomical traits and breeding values in Iranian wheat under rain-fed and well-watered conditions. *BMC Genomics.* 2022;23(1): 831. DOI 10.1186/s12864-022-08968-w

Sandhu K.S., Lozada D.N., Zhang Z., Pumphrey M.O., Carter A.H. Deep learning for predicting complex traits in spring wheat breeding program. *Front. Plant Sci.* 2021a;11:613325. DOI 10.3389/fpls.2020.613325

Sandhu K., Patil S.S., Pumphrey M., Carter A. Multitrait machine- and deep-learning models for genomic selection using spectral information in a wheat breeding program. *Plant Genome.* 2021b;14(3): e20119. DOI 10.1002/tpg2.20119

Sirsat M.S., Oblessuc P.R., Ramiro R.S. Genomic prediction of wheat grain yield using machine learning. *Agriculture.* 2022;12(9):1406. DOI 10.3390/agriculture12091406

Song H., Ye S., Jiang Y., Zhang Z., Zhang Q., Ding X. Using imputation-based whole-genome sequencing data to improve the accuracy of genomic prediction for combined populations in pigs. *Genet. Sel. Evol.* 2019;51(1):58. DOI 10.1186/s12711-019-0500-8

Stolpovsky Y.A., Piskunov A.K., Svishcheva G.R. Genomic selection. I. Latest trends and possible ways of development. *Russ. J. Genet.* 2020;56(9):1044-1054. DOI 10.1134/S1022795420090148

Suslina Ye.N., Novikov A.A., Pavlova S.V., Bashmakova N.V., Fedin G.I., Alekseyeva S.I. Evaluation of breeding qualities of hog producers using the BLUP method. *Izvestiya Timiryazevskoy Sel'skokhozyaystvennoy Akademii = Izvestiya of Timiryazev Agricultural Academy*. 2019;6:150-161. DOI 10.34677/0021-342x-2019-6-150-161 (in Russian)

Tajalifar M., Rasooli M. Importance of BLUP method in plant breeding. *J. Plant Sci. Phytopathol.* 2022;6(2):40-42. DOI 10.29328/journal.jpsp.1001072

Tsai H.Y., Janss L.L., Andersen J.R., Orabi J., Jensen J.D., Jahoor A., Jensen J. Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat. *Sci. Rep.* 2020;10(1):3347. DOI 10.1038/s41598-020-60203-2

Wang X., Xu Y., Hu Z., Xu C. Genomic selection methods for crop improvement: Current status and prospects. *Crop J.* 2018;6(4):330-340. DOI 10.1016/j.cj.2018.03.001

Yang J., Lee S.H., Goddard M.E., Visscher P.M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 2011;88(1): 76-82. DOI 10.1016/j.ajhg.2010.11.011

Zhao Y., Mette M.F., Gowda M., Longin C.F., Reif J.C. Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity.* 2014;112(6):638-645. DOI 10.1038/hdy.2014.1

Zhumanov K.Z., Karymsakov T.N., Kineev M.A., Baimukanov A.D. Development and optimization of the equations of the mixed BLUP model for the evaluation of the breed value of bulls-producers of the golstin black-motioned breed of the Republic of Kazakhstan. *Agrarnaya Nauka = Agrarian Science.* 2021;2:33-36. DOI 10.32634/0869-8155-2021-345-2-33-36 (in Russian)

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

Received November 29, 2023. Revised March 6, 2024. Accepted March 12, 2024.