2020 • 24 • 2

Print ISSN 2500-0462
Online ISSN 2500-3259

Научный рецензируемый журнал

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.
Периодичность 8 выпусков в год
DOI 10.18699/VJ20.605

Учредители

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

Сибирское отделение Российской академии наук

2020 • 24 • 2

Print ISSN 2500-0462
Online ISSN 2500-3259

Scientific Peer Reviewed Journal

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

Founded in 1997
Published 8 times annually
DOI 10.18699/VJ20.605

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

СОДЕРЖАНИЕ • 2020 • 24 • 2

Генетика растений

Селекция растений на иммунитет

Генетика микроорганизмов

Генетика животных

Физиологическая генетика

Генетика человека

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2020 • 24 • 2

Вавиловский журнал генетики и селекции. 2020;24(2):115-122
DOI 10.18699/VJ20.606

Identification of genome compositions in allopolyploid species of the genus *Elymus* (Poaceae: Triticeae) in the Asian part of Russia by CAPS analysis

A.V. Agafonov[1] ✉, E.V. Shabanova (Kobozeva)[1], S.V. Asbaganov[1], A.V. Mglinets[2], V.S. Bogdanova[2]

[1] Central Siberian Botanical Garden of Siberian Branch of the Russian Academy of Sciences, Russia, Novosibirsk, Russia
[2] Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ e-mail: agalex@mail.ru

Abstract. The genus *Elymus* L., together with wheat, rye, and barley, belongs to the tribe Triticeae. Apart from its economic value, this tribe is characterized by abundance of polyploid taxa formed in the course of remote hybridization. Single-copy nuclear genes are convenient markers for identification of source genomes incorporated into polyploids. In the present work, a CAPS-marker is developed to distinguish basic St, H, and Y genomes comprising polyploid genomes of Asiatic species of the genus *Elymus*. The test is based on electrophoretic analysis of restriction patterns of a PCR-amplified fragment of the gene coding for beta-amylase. There are about 50 *Elymus* species in Russia, and most of them are supposed to possess one of three haplome combinations, StH, StY and StHY. Boreal StH-genomic species endemic for Russia are the least studied. On the basis of nucleotide sequences from public databases, *Taq*I restrictase was selected, as it produced patterns of restriction fragments specific for St, H, and Y haplomes easily recognizable in agarose gel. A sample of 68 accessions belonging to 32 species was analyzed. In 15 species, the earlier known genomic constitutions were confirmed, but in *E. kamoji* this assay failed to reveal the presence of H genome. This unusual H genome was suggested to originate from a different *Hordeum* species. In 16 species, genomic constitutions were identified for the first time. Fifteen accessions from Asian Russia possessed the genomic constitution StStHH, and *E. amurensis*, phylogenetically close to the StY-genomic species *E. ciliaris*, had the genomic constitution StStYY. It is inferred that the center of species diversity of the StH-genomic group is shifted to the north as compared to the center of origin of StY-genomic species, confined to China.
Key words: *Elymus*; taxonomy; allopolyploids; genome constitution; CAPS markers.

For citation: Agafonov A.V., Shabanova (Kobozeva) E.V., Asbaganov S.V., Mglinets A.V., Bogdanova V.S. Identification of genome compositions in allopolyploid species of the genus *Elymus* (Poaceae: Triticeae) in the Asian part of Russia by CAPS analysis. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020; 24(2):115-122. DOI 10.18699/VJ20.606

Выявление геномного состава аллополиплоидных видов рода *Elymus* (Poaceae: Triticeae) Азиатской России с помощью CAPS-анализа

А.В. Агафонов[1] ✉, Е.В. Шабанова (Кобозева)[1], С.В. Асбаганов[1], А.В. Мглинец[2], В.С. Богданова[2]

[1] Центральный сибирский ботанический сад Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: agalex@mail.ru

Аннотация. Род *Elymus* L. наряду с пшеницей, рожью и ячменем принадлежит к трибе Triticeae. Помимо своего хозяйственного значения, эта триба характеризуется широким распространением аллополиплоидных таксонов, которые формируются в ходе межвидовой и межродовой гибридизации и последующих преобразований вовлеченных в гибридизацию диплоидных геномов. Для идентификации исходных геномов в составе полиплоидов удобны малокопийные ядерные гены, менее подверженные процессам реорганизации, чем повторенные некодирующие элементы. В настоящей работе разработан удобный CAPS-маркер для различения базисных геномов St, H, Y, входящих в состав азиатских видов рода *Elymus*, с помощью электрофоретического анализа фрагментов рестрикции ПЦР-амплифицированного участка гена, кодирующего β-амилазу. В России распространено около 50 видов *Elymus* предположительно трех гапломных комбинаций: StH, StY и StHY. Наименее изученными остаются бореальные StH-геномные виды – эндемики Российской Федерации. По результатам анализа ранее изученных разными авторами нуклеотидных последовательностей гена β-амилазы была отобрана эндонуклеаза рестрикции *Taq*I, которая имела различающиеся по положению сайты узнавания в составе вышеуказанного фрагмента из геномов St, H и Y. В результате расщепления ПЦР-продукта эндонуклеазой *Taq*I у каждого из исходных гапломов формировался специфичный паттерн фраг-

ментов рестрикции, легко визуализируемый в агарозном геле. Проанализирована выборка из 68 образцов, принадлежащих 32 видам. У 15 видов была подтверждена ранее известная геномная конституция, у *E. kamoji* этот метод не позволил выявить присутствие генома H. Предполагается возможное происхождение данного варианта генома H от другого вида *Hordeum*. У 16 видов геномная конституция определена впервые. Показано, что большинство изученных видов бореальной группы видов из Сибири и Российского Дальнего Востока имеют геномную конституцию StStHH. Исключение составил *E. amurensis*, филогенетически близкий к StY-геномному виду *E. ciliaris* и также имеющий геномный состав StStYY. Сделан вывод, что основное видовое разнообразие StH-геномной группы находится севернее центра происхождения большинства StY-геномных видов рода.

Ключевые слова: *Elymus*; таксономия; аллополиплоиды; геномная конституция; CAPS-маркеры.

Introduction

The genus *Elymus* L. is the largest in the tribe Triticeae Dum. and, according to different estimates, counts from 150 to 200 species (Dewey, 1984; Barkworth, 2000). It is represented only by alloploid taxa with genome compositions including several basic genomes (haplomes) in different combinations. The genetic basis of the genus *Elymus* is formed by five haplomes descending from different genera of the tribe Triticeae: (St) *Pseudoroegneria*, (H) *Hordeum*, (P) *Agropyron*, (W) *Astralopyrum*, (Y) donor unknown. Genome constitution was proposed as a stable genetic criterion for taxonomic classification of *Elymus* species (Löve, 1984). Within a relatively short span of time, substantial changes occurred in the taxonomy of the tribe Triticeae on the basis of the genomic system of classification suggested by D.R. Dewey (1984). During the next 20 years, six genera were identified according to variants of genome constitution: *Douglasdeweya* C. Yen, J.L. Yang & B.R. Baum (PPStSt), *Roegneria* C. Koch (StStYY), *Anthosachne* Steudel (StStWWYY), *Kengylia* C. Yen & J.L. Yang (PPStStYY), *Campeiostachys* Drobow (HHStStYY), and *Elymus* L. (StStHH, StStStHH, StStHHHH).

However, departing from A. Löve's principles, many botanists still attribute several genome combinations to the single genus *Elymus* s. l. With all this, genome constitutions are not yet determined in about 40 % of species (Okito et al., 2009). According to current evidence, 53 species of the genus *Elymus* subdivided into four sections occur in Russia (Tsvelyov, 2008; Tsvelyov, Probatova, 2010). Two of the sections, *Elymus* and *Goulardia* (Husn.) Tzvelev, contain species with different genomic constitutions, which obviously contradicts the phylogenetic principle of their formulation. We suppose that Russia is home to species with only three haplome combinations: StH, StY, and StHY (Agafonov et al., 2015). Boreal StH-genomic endemics of Russia are less studied. According to the taxonomic system based on the genome constitution, the *Elymus* species should be attributed to three genera: *Elymus*, *Roegneria*, and *Campeiostachys.* However, in our view, the division of the species inhabiting Russia into three genera is impractical due to the difficulties of morphologic identification of these genera. With all this, taxonomic classification within the genus based on genome constitutions is indispensable for the construction of a phylogenetically oriented taxonomy of the genus.

Earlier, Cleaved Amplified Polymorphic Sequences (CAPS) markers were used to distinguish individual genomes in representatives of the tribe Triticeae (Gostimsky et al., 2005; Li et al., 2007; Hu et al., 2014; Shavrukov, 2016). Some advantages of CAPS markers are their codominance, moderate sensitivity to the amount of genomic DNA, and relatively low cost.

We were first to use CAPS-markers to identify the genomic constitutions of species of the genus *Elymus* (Kobozeva et al., 2017). For this purpose, primers were designed based on the known sequences of the gene coding for β amylase (Mason-Gamer, 2013), which included 38 sequences of haplome St, 23 of haplome H, and 15 of haplome Y, belonging to 24 *Elymus* species. Of them, 14 species had the genomic composition StStHH; 9, StStYY; and 1, StStHHUkUk (*Elytrigia repens*). Variable positions were sought that would discriminate representatives of an individual genome from the other two. Special attention was paid to those genome-specific sequence variants that resulted in appearance/disappearance of recognition sites for restriction endonucleases. It was found that digestion of the PCR products with *Taq*I endonuclease resulted in the formation of genome-specific restriction patterns. In the present work, we apply CAPS analysis to a large sample of *Elymus* species from Asian Russia to reveal their genome constitutions unknown hitherto.

Materials and methods

Plant material included 68 accessions of the species with known (Table 1) and unknown (Table 2) genome constitutions found in Russia. The species nomenclature is given according to N.N. Tsvelyov and N.S. Probatova (2010). The accessions analyzed were received from the scientific collection of biological resources of the Central Siberian Botanic Garden SB RAS "Collections of living plants indoors and outdoors"; their identification numbers are given in Tables 1 and 2. Prefixes correspond to the geographic origin of the accessions.

Total DNA was extracted from 20 mg of dried green matter with the use of NucleoSpin Plant II Kit (Macherey-Nagel, Germany) according to manufacturer's recommendations. Amplification of the β amylase gene fragment was made in a C-1000 thermocycler (Bio-Rad, USA) with the following primers: El_balg_F4 (5′-GGTACCATCGTGGACATTGAA-3′) and El_balg_R4 (5′-CTGTACCACCAGTGAATGCC-3′) (Kobozeva et al., 2017). The PCR reaction mixture of 15 μL in volume contained 1× buffer for Taq polymerase, 0.2 mM each dNTP, 1.5 mM $MgCl_2$, 1 μM each of primers, 20 ng of genomic DNA, and 1 U of HS Taq DNA polymerase (Evrogene, RF). The following settings were used: predenaturation at 94 °C for 4 min; 40 cycles: denaturation at 94 °C for 20 s, primer annealing at 60 °C for 25 s, elongation at 72 °C for 90 s; postextension at 72 °C for 5 minutes. CAPS-analysis (Konieczny, Ausubel, 1993) was made as follows: 8 μL of the PCR reaction mixture was mixed with MQ-H_2O and *Taq*I buffer up to 1× concentration in a volume of 15 μL, and 1 unit of *Taq*I restrictase (Thermo Scientific, USA) was added. The mixture was incubated at 65 °C for 1 hour and resolved in

Table 1. Accessions of *Elymus* species with known genomic constitutions determined by the classical cytogenetic method

No	*Elymus* species; accession	Genomic constitutions	Location and collectors
1	*E. kamoji*; CCH-1395	StHY (Zhou et al., 1999)	People's Republic of China (PRC), prov. Sichuan, Chengdu outskirts, a forest patch in Panda park, alt. 527 m, N 30°44.253′ E 104°8.453′ (A. Agafonov, E. Kobozeva)
2	*E. kamoji*; SLA-1276	»	Russia, Primorskiy Krai, Khasan raion, Slavyanka Town, coastal meadow, alt. 1 m, N 42°52.101′ E 131°22.987′ (A. Agafonov)
3	*E. dahuricus*; BUD-0781	StHY (Dewey, 1984)	Russia, Republic of Buryatia, Dzhirga raion, Nizhniy Torey village outskirts, the Toreyka River valley, meadow slope, alt. 863 m, N 50°34.567′ E 104°52.571′ (N. Badmaeva)
4	*E. dahuricus*; CHJ-1516	»	People's Republic of China (PRC), Jilin Province, nearby Lake Tsagan-Nur, forest plantation, alt. 138 m, N 45°12.106′ E 124°25.074′ (S. Asbaganov)
5	*E. schrenkianus*; AKA-0702	StHY (Lu, Bothmer, 1992)	Russia, Altai Republic, Kosh-Agach raion, stony meadow, Kalanegir River valley, alt. 2283 m, N 49°37.896′ E 88°29.441′ (S. Lukjanchikov)
48	*E. schrenkianus*; AUK-0652	»	Russia, Altai Republic, Kosh-Agach raion, stony meadow in a brook valley, alt. 2445 m, N 49°30.418′ E 88°5.012′ (A. Agafonov, B. Salomon)
6	*E. pendulinus*; VOK-0738	StY (Jensen, 1990)	Russia, Vladivostok City outskirts, Okeanskaya Station, shrubbery; alt. 7 m, N 43°14.10′ E 132°0.19′ (A. Agafonov)
7	*E. pendulinus*; CHE-1044	»	Russia, Altai Republic, Chemal raion, confluence of the Chemal and Katun Rivers, alt. 434 m, N 51°23.533′ E 86°00.197′ (E. Kobozeva)
8	*E. gmelinii*; AUS-1013	StY (Jensen, Hatch, 1989)	Russia, Altai Republic, Shebalino raion, Ust-Sema settlement outskirts, shingle bank in the Katun River left floodplain, alt. 341 m, N 51°37.620′ E 85°45.923′ (A. Agafonov, E. Kobozeva)
9	*E. sibiricus*; ACH-1601	StH (Dewey, 1974)	Russia, Altai Republic, Chike-Taman Pass, right side of the road, 500 m N of the viewpoint, alt. 1250 m, N 50°38.911′ E 86°18.789′ (E. Kobozeva)
10	*E. caninus*; OSE-1423	StH (Dewey, 1968)	Russia, North Osetia-Alania Republic, Alagir raion, ecological path in the Tseydon River valley, alt. 1951 m, N 42°47.139′ E 43°53.605′ (A. Agafonov, M. Agafonova)
11	*E. caninus*; OSE-1427	»	Russia, North Osetia-Alania Republic, Iraf raion, Stur-Digora village outskirts, Great Caucasus Mt. Range northern slope, the valley of a right Urukh River tributary, alt. 1996 m, N 42°52.898′ E 43°35.959′ (S. Asbaganov)
12	*E. caninus*; UKU-1617	»	Russia, Republic of Bashkortostan, Beloretskiy district, Novoabzakovo village outskirts, dry meadow in a mixed forest, alt. 618 m, N 53°48.718′ E 58°40.377′ (A. Agafonov, S. Asbaganov)
13	*E. gmelinii*; BKA-0962	StY (Jensen, Hatch, 1989)	Russia, Primorskiy Krai, Bolshoy Kamen Town outskirts, alt. 41 m, N 43°7.513′ E 132°25.133′ (A. Agafonov)
14	*E. fedtschenkoi*; KSA-0935	StY (Liu, Dewey, 1983)	Kazakhstan, South Altai Mt. Range, the W principal slope, alt. 1791 m, N 49°05.077′ E 86°04.483′ (D. Gerus)
49	*E. fedtschenkoi*; KME-1729	»	Kazakhstan, Trans-Ili Alatau Mts., Medeu village outskirts, stony highland meadow, alt. 3223 m, N 43°07.101′ E 77°06.828′ (S. Asbaganov)
50	*E. fedtschenkoi*; AUR-1714	»	Russia, Altai Republic, Ulagan raion, the Yarly-Amry River, stony highland meadow, alt. 2180 m, N 50°19.044′ E 087°43.049′ (E. Kobozeva)
51	*E. nevskii*; KME-1728	StY (Dewey, 1980)	Kazakhstan, Trans-Ili Alatau Mts., Medeu village outskirts, stony highland meadow, alt. 3223 m, N 43°07.101′ E 77°06.828′ (S. Asbaganov)
15	*E. fibrosus*; ABZ-1602	StH (Dewey, 1984)	Russia, Republic of Bashkortostan, Beloretsk raion, Novoabzakovo village outskirts, a ground road side, alt. 546 m, N 53°47.845′ E 58°37.291′ (A. Agafonov, S. Asbaganov)
16	*E. mutabilis*; KHA-1210	»	Russia, Krasnoyarskiy Krai, Shushenskoe raion, Shushenskiy Bor pine forest, alt. 495 m, N 52°49.622′ E 91°26.609′ (S. Asbaganov)
17	*E. mutabilis*; ABZ-1607	»	Russia, Republic of Bashkortostan, Beloretsk raion, Novoabzakovo village outskirts, a ground road side, alt. 546 m, N 53°47.845′ E 58°37.291′ (A. Agafonov, S. Asbaganov)
26	*E. abolinii*; BUD-0780	StY (Jensen, 1989)	Russia, Republic of Buryatia, Dzhida raion, Nizhniy Torey village outskirts, the Toreyka River valley, meadowy slope, alt. 863 m, N 50°34.634′ E 104°52.781′ (N. Badmaeva)
27	*E. ciliaris*; VOK-0711	StY (Dewey, 1984)	Russia, Vladivostok Sity env., between Sanatornaya and Okeanskaya Stations, shrubbery, alt. 4 m, N 43°13.94′ E 131°59.95′ (D. Gerus, A. Agafonov)
38	*E. panormitanus*; H4152*	StY (Lu, Salomon, 1992)	USSR: Ukraine, Crimea
39	*E. caucasicus*; H3207*	StY (Jensen, Wang, 1991)	USSR: Armenia, Dilidjan city outskirts
66	*E. confusus*; BUM-0505	StH (Lu et al., 1995)	Russia, Republic of Buryatia, Tunka raion, Mondy village outskirts, the Tunka Range southern slope, a stony brook bank, alt. 1738 m, N 51°42.610′ E 100°59.967′ (D. Gerus, A. Agafonov)
67	*E. confusus*; TAR-0730	»	Russia, Republic of Tyva, Tes-Khem raion, Khorumnug-Taiga Mt. Range, Shuurmak-Samagaltay Pass, spruce-larch forest, alt. 1545 m, N 50°36.870′ E 95°10.729′ (I. Artemov)

Note. The numbering of accessions corresponds to the lane numbering in Fig. 2.
* Accessions kindly provided by Dr. B. Salomon (Swedish University of Agricultural Sciences, Department of Plant Breeding, Alnarp, Sweden).

Table 2. Accessions of *Elymus* species with unknown genomic constitutions collected in Russia

No.	*Elymus* species; accession	Location and collectors
18	*E. uralensis*; UKU-1617	Republic of Bashkortostan, Beloretsk raion, Novoabzakovo village outskirts, dry meadow in a mixed forest, alt. 618 m, N 53°48.718′ E 58°40.377′ (A. Agafonov, S. Asbaganov)
19	*E. viridiglumis*; UKU-1618	Republic of Bashkortostan, Beloretsk raion, Novoabzakovo village outskirts, tall herbage meadow in a birch open stand, alt. 619 m, N 53°48.718′ E 58°40.377′ (A. Agafonov, S. Asbaganov)
20	*E. transbaicalensis*; AKU-0422	Altai Republic, Kosh-Agach raion, 10 km N of Chagan-Uzun village along the Chuya Highway, Kuyaktanar valley, alt. 1815 m, N 50°9.783′ E 88°19.054′ (A. Agafonov, D. Gerus)
21	*E. transbaicalensis*; GAR-0530	Republic of Buryatia, Oka raion, the road to Orlik town, shingle bank of the Gargan River, alt. 1610 m, N 52°05.947′ E 100°23.005′ (A. Agafonov, D. Gerus)
22	*E. margaritae*; GUK-1009	Altai Republic, Ust-Koksa raion, Krasnaya Mt., a complex of screes and highland meadows, alt. 2028 m, N 50°4.495′ E 85°13.073′ (D. Nikonova, E. Kobozeva)
23	*E. margaritae*; AUK-0650	Altai Republic, Kosh-Agach raion, Ukok Plateau, stony meadow in a brook valley, alt. 2438 m, N 49°30.418′ E 88°05.012′ (A. Agafonov, B. Salomon)
24	*E. komarovii*; AKU-0458	Altai Republic, Kosh-Agach raion, 10 km N of Chagan-Uzun village along the Chuya Highway, Kuyaktanar valley, alt. 1815 m, N 50°9.783′ E 88°19.054′ (A. Agafonov, D. Gerus)
25	*E. transbaicalensis*; TUV-9697	Republic of Tyva, Todzha raion, Azas State Nature Reserve, Ilgi-Chul ranger post (D. Shaulo)
28	*E. komarovii*; AKT-0417	Altai Republic, Kosh-Agach raion, North-Chuya Range, Aktry Gorge, ground road edge at forest margin, alt. 2061 m, N 50°6.518′ E 87°48.193′ (A. Agafonov, D. Gerus)
29	*E. komarovii*; GAR-0501	Republic of Buryatia, Oka raion, the Oka River valley, forest glade 50 m from the Gargan River mouth, alt. 1607 m, N 52°05.947′ E 100°23.005′ (A. Agafonov, D. Gerus)
30	*E. komarovii*; JPO-1505	Republic of Sakha-Yakutia, Khangalas raion, Pokrovsk Town outskirts, a meadow at the gas station, alt. 131 m, N 61°29.367′ E 129°08.225′ (E. Kobozeva, S. Asbaganov)
31	*E. subfibrosus*; ANA-1118	Chukotskiy Autonomous district, Anadyr′ Town outskirts (D. Lysenko)
32	*E. subfibrosus*; LEN-1524	Republic of Sakha-Yakutia, Khangalas raion, the nature reserve "Lenskie Stolby", alt. 156 m, N 61°6.370′ E 127°21.593′ (E. Kobozeva, S. Asbaganov)
33	*E. macrourus*; 12-0135	Taymyr Peninsula, the shingle floodplain of the Bolshaya Lesnaya Rassokha River at its mouth, alt. 2 m, N 72°37.363′ E 101°17.793′ (E. Pospelova)
34	*E. jacutensis*; 13-0443	Taymyr Peninsula, the Anabar Plateau margin, Eriechka and Nyamakit-Daldyn Rivers, a small meadow below rocks, alt. 218 m, N 71°15.250′ E 105°37.452′ (I. Pospelov)
35	*E. sajanensis*; ZUN-0502	Republic of Buryatia, Oka raion, Zun-Kholbo village outskirts, alt. 1682 m, N 52°10.092′ E 100°57.581′ (A. Agafonov, D. Gerus)
36	*E. sajanensis*; ART-0202	Altai Republic, Kosh-Agach raion, Chikhacheva Range, shingle bank of a Buguzun River left tributary, alt. 2254 m, N 50°1.914′ E 89°23.620′ (I. Artemov)
37	*E. amurensis*; MES-1111	Primorskiy Krai, Khasan raion, Andreevka village outskirts, meadow patch at a ground road edge, alt. 93 m, N 42°37.045′ E 131°8.650′ (E. Kobozeva, A. Agafonov)
40	*E. transbaicalensis*; AKT-0628	Altai Republic, Kosh-Agach raion, North-Chuya Range, Aktry Gorge, willow thickets at the mountaineering camp, alt. 2118 m, N 50°5.038′ E 87°46.820′ (A. Agafonov, D. Gerus)
41	*E. kronokensis*; BER-0804	Republic of Buryatia, Eravnoe raion, SE of the temporary settlement Ozernyy, larch forest, alt. 1154 m, N 52°58.625′ E 111°38.166′ (O. Anenkhonov)
42	*E. kronokensis*; MMA-1103	Magadan oblast, Madaun village outskirts, a burnt area in the Arman′ Rover floodplain, alt. 627 m, N 60°35.861′ E 150°40.862′ (D. Lysenko)
43	*E. kronokensis*; KES-9603	Kamchatka Krai, Bystraya raion, southern slope of a mountain N of Esso village, alt. 627 m, N 55°55.945′ E 158°41.275′ (A. Agafonov, B. Salomon)
44	*E. lenensis*; 12-0125	Taymyr Peninsula, Bolshaya Rassokha and Novaya Rivers, alt. 39 m, N 72°39.613′ E 101°17.079′ (I. Pospelov)
45	*E. kamczadalorum*; KSO-9623	Kamchatka Krai, Elizovo raion, Sosnovka village outskirts, alt. 247 m, N 53°5.046′ E 158°17.918′ (A. Agafonov, B. Salomon)
46	*E. charkeviczii*; KES-9670	Kamchatka Krai, Bystraya raion, Esso village outskirts, ground road margin, alt. 484 m, N 55°55.014′ E 158°42.116′ (A. Agafonov, B. Salomon)
47	*E. charkeviczii*; MSN-1202	Magadan City, Snezhnyy settlement, path at a forest margin, alt. 145 m, N 59°43.466′ E 150°52.677′ (N. Badmaeva)
52	*E. lenensis*; LEN-1520	Republic of Sakha (Yakutia), Khangalas raion, meadow slope at the Lena River right bank, alt. 114 m, N 61°6.369′ E 127°21.593′ (E. Kobozeva, S. Asbaganov)

Table 2 (end)

No.	*Elymus* species; accession	Location and collectors
53	*E. lenensis*; ALD-1539-3	Republic of Sakha (Yakutia), Aldan raion, the Aldan River bank, shrubbery at a sandy bank, alt. 228 m, N 58°40.878′ E 128°33.081′ (E. Kobozeva, S. Asbaganov)
54	*E. kronokensis*; KRT-1611	Krasnoyarsk Kray, Evenk raion, Tura village outskirts, the Nizhnyaya Tunguska River, path side in a larch forest, alt. 169 m, N 64°16.478′ E 100°16.445′ (L. Krivobokov)
55	*E. kronokensis*; 12-0137	Taymyr Peninsula, the Bolshaya Rassokha River bank bluff, alt. 2 m, N 72°35.808′ E 101°15.900′ (E. Pospelova)
56	*E. kronokensis*; TAL-0602	Altai Republic, Kosh-Agach raion, the Taldura River valley, larch forest on a mountain slope, alt. 2095 m, N 49°57.472′ E 87°57.552′ (D. Gerus, A. Agafonov)
57	*E. subfibrosus*; KRT-1612	Krasnoyarsk Kray, Evenk raion, Tura village, a ground road edge, alt. 309 m, N 64°16.920′ E 100°14.880′ (L. Krivobokov)
58	*E. subfibrosus*; JRO-1733	Republic of Sakha (Yakutia), Tompo raion, the Kolyma Riad, Verkhoyanskiy Mt. Range southern spurs, the Rosomakha River valley, alt. 460 m, N 63°2.879′ E 137°52.610′ (N. Badmaeva)
59	*E. jacutensis*; ALU-1711	Altai Republic, Ulagan raion, herbaceous meadow at the Chulyshman River left bank under Katu-Yaryk Pass, alt. 733 m, N 50°55.497′ E 088°12.226′ (E. Kobozeva)
60	*E. jacutensis*; GAN-1516	Altai Republic, Chemal raion, Anos village outskirts, slope above the Anos River left bank at the bridge, alt. 380 m, N 51°30.014′ E 85°57.160′ (E. Kobozeva)
61	*E. jacutensis*; ALD-1541	Republic of Sakha (Yakutia), Aldan raion, the Aldan River bank, shrubbery at a sandy bank, alt. 228 m, N 58°40.878′ E 128°33.081′ (E. Kobozeva, S. Asbaganov)
62	*E. macrourus*; MTE-1210	Magadan oblast, Tenka raion, roadside at a mixed forest margin, alt. 970 m, N 60°26.034′ E 150°58.558′ (N. Badmaeva)
63	*E. macrourus*; LEN-1524_1	Republic of Sakha (Yakutia), Khangalas raion, meadow slope at the Lena River right bank, alt. 114 m, N 61°6.369′ E 127°21.593′ (E. Kobozeva, S. Asbaganov)
64	*E. turuchanensis*; KRE-1440	Krasnoyarsk Kray, Turukhansk raion, Bor village, Yenisey River sandy bank, alt. 30 m, N 61°36.265′ E 90°0.143′ (M. Lomonosova)
65	*E. peschkovae*; MJA-1106	Magadan oblast, Khasyn raion, Yablonovyy Pass, floodplain meadow at a road, alt. 755 m, N 60°19.467′ E 151°10.540′ (D. Lysenko)
68	*E. peschkovae*; AMU-8804	Amur oblast, 50 km downstream the Gilyuy River from Tynda City, sandy bank, alt. 445 m, N 54°56.216′ E 125°21.854′ (O. Potemkin)

Note. The numbering of accessions corresponds to the lane numbering in Fig. 2.

1.7 % agarose gel in TAE buffer. Molecular weight marker: 100+ bp DNA Ladder (Evrogen, RF).

Results and discussion

The comparative analysis of sequences of the β amylase gene published in R. Mason-Gamer (2013) showed that the studied fragment of Y genome of about 1100 bp in length did not contain recognition sites for *Taq*I endonuclease, while St genome contained one recognition site in the fragment of interest at a distance of about 170 bp from the primer El_balg_R4. The same site was present in some H genomes; besides, all H genomes contained a recognition site at a distance of about 280 bp from the primer El_balg_F4. Visualized on gels, restriction patterns of the studied genomes were differentiated according to the lengths of the longest fragments: H genome was distinguished by the presence of a band at about 650 bp; St genome, 930 bp; and Y genome, 1100 bp (Fig. 1).

Restriction patterns of the CAPS marker employed were studied in 68 accessions (see Tables 1, 2). Electrophoretic patterns formed after *Taq*I digestion are shown in Fig. 2. Based on the results of CAPS analysis, genomic constitutions of the accessions studied were determined. Previously known genomic constitutions were confirmed in 15 species of 16, *E. kamoji* being the only exception. In 16 species, genomic compositions were determined for the first time: 15 of them had the genomic constitution StStHH, and one species, *E. amurensis*, had StStYY (Table 3). However, some limitations of the approach were met. For example, in two accessions of *E. kamoji* CAPS-analysis revealed only two haplomes, St and Y (Fig. 2, lanes *1* and *2*), whereas it is known to be hexaploid according to the number of chromosomes, thus, it should contain three basic genomes (haplomes). It is improbable that the absence of restriction fragments corresponding to haplome H was due to incomplete digestion. Since all representatives of the genus contain St haplome, possessing a recognition site for *Taq*I endonuclease, the presence of St-specific fragments serves as an

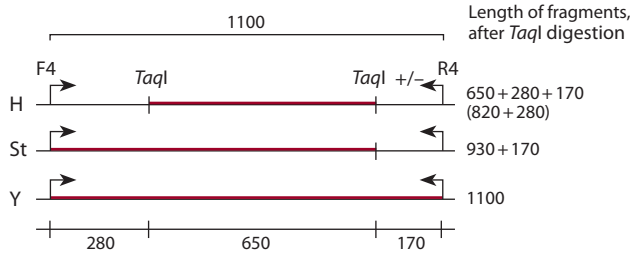


Fig. 1. Map of recognition sites for *Taq*I endonuclease in the β amylase gene fragment amplified from the basic haplomes constituting the polyploid *Elymus* genome.

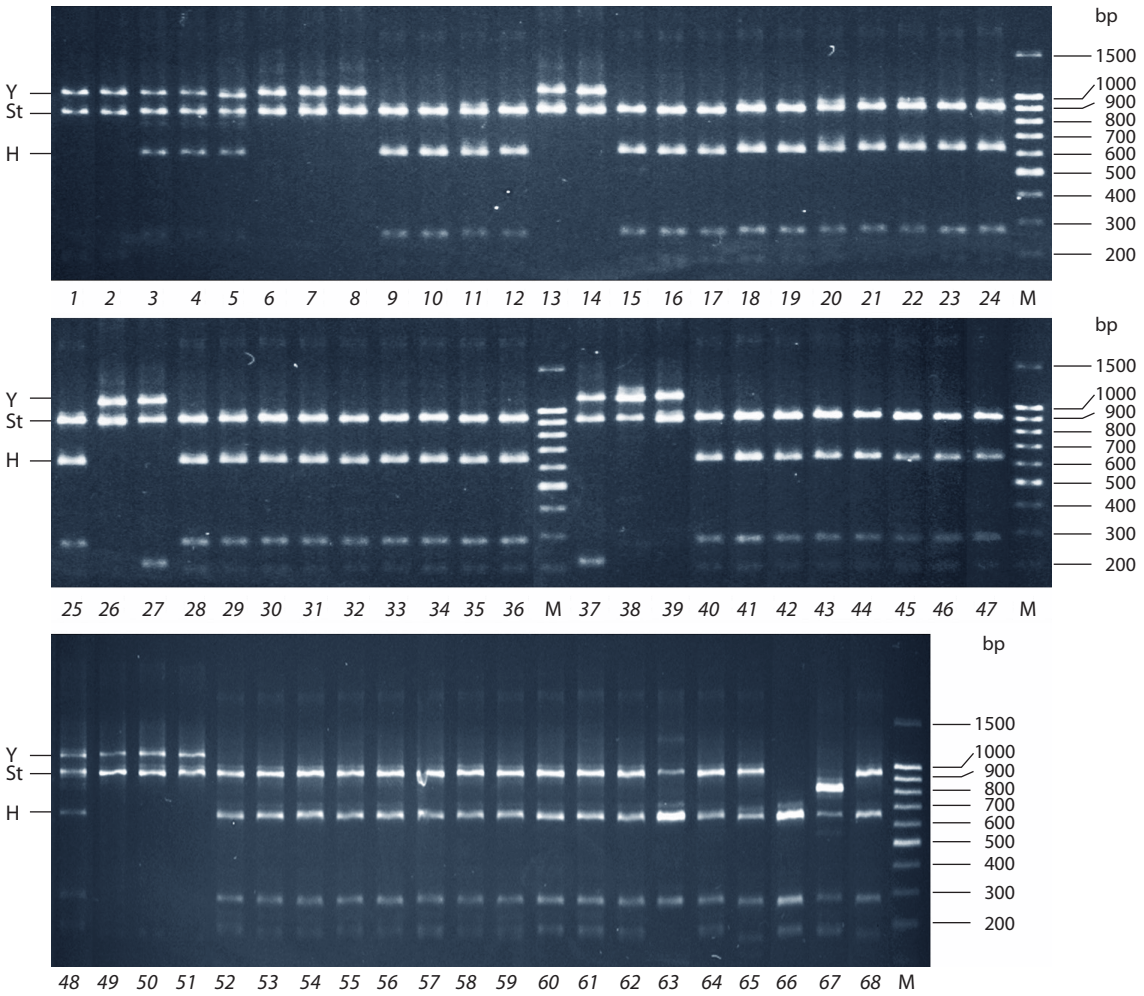


Fig. 2. Polymorphism of restriction fragment lengths (CAPS) after *Taq*I digestion of the PCR-amplified fragment of the β amylase gene in species of the genus *Elymus*.

Lane numbers correspond to the accession numbering in Tables 1 and 2. M – molecular weight ladder: 100+bp DNA Ladder (Evrogen).

Table 3. The list of boreal *Elymus* species in Asian Russia in which genome constitutions (GC) were determined by the CAPS method

No.	*Elymus* species	Number of accessions studied	GC	No.	*Elymus* species	Number of accessions studied	GC
1	*E. amurensis*	1	StY	9	*E. margaritae*	2	StH
2	*E. charkeviczii*	2	StH	10	*E. sajanensis*	2	StH
3	*E. jacutensis*	4	StH	11	*E. subfibrosus*	4	StH
4	*E. kamczadalorum*	1	StH	12	*E. transbaicalensis*	4	StH
5	*E. komarovii*	4	StH	13	*E. uralensis*	1	StH
6	*E. kronokensis*	6	StH	14	*E. viridiglumis*	1	StH
7	*E. lenensis*	3	StH	15	*E. turuchanensis*	1	StH
8	*E. macrourus*	3	StH	16	*E. peschkovae*	2	StH

internal control for the completeness of hydrolysis. According to the classification system based on genomic compositions, *E. kamoji* belongs to the genus *Campeiostachys* (Baum et al., 2011) which embraces species with the genomic composition StHY. In fact, we performed a cytological analysis, which showed that both accessions of *E. kamoji* possessed the chromosome number $2n = 42$, corresponding to hexaploid. The presence of the H genome lacking two recognition sites for *Taq*I endonuclease in *E. kamoji* brings its origin into a question. It is not inconceivable that different representatives of

the genus received their H genomes from different ancestor species, which agrees with the assumption of polyphyly of the donors of basic haplomes (Mason-Gamer, 2013).

An interesting pattern of restriction fragments was observed in two accessions of *E. confusus* (see Fig. 2, lanes *66* and *67*), with the genome constitution formerly determined as StStHH (Lu et al., 1995). In accession TAR-0730 (see Fig. 2, lane *67*), the longer fragment corresponding to the allele from St genome is truncated, possibly, as the result of a deletion or acquisition of an additional restriction site. The spectrum of restriction fragments in accession BUM-0505 (see Fig. 2, lane *66*) lacks the fragment of about 930 bp characteristic of St genome, while the smaller fragment of about 170 bp corresponding to this haplome is clearly seen. This phenomenon might be attributed to a mutation in the St genome of the accession, for example, appearance of a recognition site for *Taq*I. Another possibility is a recombination and/or introgression between genomes St and H in the course of intense microevolutionary processes indirectly confirmed by the high morphologic variability within this species.

According to the CAPS analysis undertaken in the present work, almost all newly studied accessions of the boreal group of species from Siberia and Russian Far East have the StH genomic composition. One exception was *E. amurensis*, phylogenetically close to the StY-genomic species *E. ciliaris* and possessing the genome composition StY. This implies that the center of species diversity of the Asiatic StH-genome group is shifted to the north as compared to that of the StY-genome group, which is considered to be situated in China (Lu, Salomon, 1992). In this context, it is worth noting that in North America, the genus *Elymus* is also represented mainly by StH-genome species (except for *Elymus californicus* with unclear origin) (Mason-Gamer, 2001). Besides, in that territory a number of adventive Asiatic StHY- and StY-genome species were found (Barkworth et al., 2007).

In general, the applied method showed a high accuracy: in the present work earlier known genome constitutions were confirmed by CAPS analysis in 15 *Elymus* species of 16. For 10 species, the genomic composition newly determined by CAPS analysis as StH, was independently corroborated by sequencing of a cloned fragment of the *GBSS1* (*waxy*) gene (Kobozeva et al., 2018; Agafonov et al., 2019). It should be noted that the sequencing of DNA from polyploid species has a disadvantage, as it is rather laborious, requiring additional gene cloning manipulations.

Conclusion

The main advantage of CAPS markers is the ease of their methodic implementation, which permits one to analyze many specimens with extensive morphologic and genetic variability from broad ranges. The present work involves CAPS analysis with the use of a fragment of the gene for β amylase and demonstrates rather good predictive power of the method. However, it should be kept in mind that no molecular marker taken by itself can unambiguously identify a genome or species; it serves as a marker, not diagnostic. Therefore, the development of additional simple and accessible approaches for genome identification in new and poorly studied biotypes from local habitats remains vital.

References

Agafonov A.V., Asbaganov S.V., Shabanova (Kobozeva) E.V., Morozov I.V., Bondar A.A. Genome constitution and differentiation of subgenomes in Siberian and Far Eastern endemic species of the genus *Elymus* (Poaceae) according to the sequencing of the nuclear gene *waxy*. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(7):817-826. DOI 10.18699/VJ19.555. (in Russian)

Agafonov A.V., Kobozeva E.V., Asbaganov S.V., Shmakov N.A. Current achievements and prospects of construction of a phylogenetically oriented taxonomy of the genus *Elymus* (Poaceae: Triticeae). Proceedings of the 14th International Scientific and Practical Conference "Problems of Botany of Southern Siberia and Mongolia". Barnaul. 2015;314-322. (in Russian)

Barkworth M.E. Changing perceptions of the Triticeae. In: Jacobs S.W.L., Everett J. (Eds.). Grasses: Systematics and Evolution. Melbourne: CSIRO, 2000;110-120.

Barkworth M.E., Cambell J.J.N., Salomon B. *Elymus* L. In: Barkworth M.E., Capels K.M., Long S., Anderton L.K., Piep M.B. (Eds.). Flora of North America. New York; Oxford: Oxford University Press, 2007;24:288-343.

Baum B.R., Yang J.-L., Yen C., Agafonov A.V. A taxonomic revision of the genus *Campeiostachys* Drobov. J. Syst. Evol. 2011;49(2):146-159.

Dewey D.R. Synthetic *Agropyron-Elymus* hybrids. III. *Elymus canadensis* × *Agropyron caninum*, *A. trachycaulum* and *A. striatum.* Amer. J. Bot. 1968;55:1133-1139.

Dewey D.R. Cytogenetics of *Elymus sibiricus* and its hybrids with *Agropyron tauri*, *Elymus canadensis* and *Agropyron caninum.* Bot. Gaz. 1974;135:80-87.

Dewey D.R. Cytogenetics of *Agropyron ugamicum* and six of its interspecific hybrids. Bot. Gaz. 1980;141:305-312.

Dewey D.R. The genomic system of classification as a guide to intergeneric hybridization with the perennial Triticeae. In: Gustafson J.P. (Ed.). Gene Manipulation in Plant Improvement. New York: Plenum Publ. Corp., 1984;209-279.

Gostimsky S.A., Kokaeva Z.G., Konovalov F.A. Studying plant genome variation using molecular markers. Russ. J. Genet. 2005;41(4):378-388. DOI 10.1007/s11177-005-0101-1.

Hu C.Y., Tsai Y.Z., Lin S.F. Development of STS and CAPS markers for variety identification and genetic diversity analysis of tea germplasm in Taiwan. Bot. Stud. 2014;55:12. DOI 10.1186/1999-3110-55-12.

Jensen K.B. Cytology, fertility, and origin of *Elymus abolinii* (Drob.) Tzvelev and its F_1 hybrids with *Pseudoroegneria spicata*, *E. lanceolatus*, *E. dentatus* ssp. *ugamicus*, and *E. drobovii* (Poaceae: Triticeae). Genome. 1989;32:468-474. DOI 10.1139/g89-470.

Jensen K.B. Cytology and morphology of *Elymus pendulinus*, *E. pendulinus* ssp. *multiculmis*, and *E. parviglume* (Poaceae: Triticeae). Bot. Gaz. 1990;151(2):245-251.

Jensen K.B., Hatch S. Genome analysis, morphology, and taxonomy of *Elymus gmelinii* and *E. strictus* (Poaceae: Triticeae). Bot. Gaz. 1989;150(1):84-92.

Jensen K.B., Wang R.R.C. Cytogenetics of *Elymus caucasicus* and *Elymus longearistatus* (Poaceae: Triticeae). Genome. 1991;34:860-867.

Kobozeva E.V., Asbaganov S.V., Agafonov A.V. Genome composition and assessment of the divergence between Russian boreal species in the genus *Elymus* (Poaceae) detected on the basis of sequencing of the nuclear gene GBSSI. In: Prospects of Development and Challenges of Modern Botany. BIO Web Conf. 2018;11.00023. DOI 10.1051/bioconf/20181100023.

Kobozeva E.V., Mglinets A.V., Agafonov A.V. Identification of the genomic composition in allopolyploid species of the genus *Elymus* (Poaceae: Triticeae) with CAPS-analysis. In: Proceedings of the 6th International Conference "Issues in the Study of the Vegetation Cover in Siberia". Tomsk. 2017;155-157. DOI 10.17223/9785946216371/51. (in Russian)

Konieczny A., Ausubel F.M. A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J. 1993;4(2):403-410.

Li X.-M., Lee B.S., Mammadov A.C., Koo B.-C., Mott I.W., Wang R.R.C. CAPS markers specific to Eb, Ee, and R genomes in the tribe Triticeae. Genome. 2007;50:400-411.

Liu C.W., Dewey D.R. The genome constitution of *Elymus fedtschenkoi.* Acta Genet. Sinica. 1983;10:20-27.

Löve A. Conspectus of the Triticeae. Feddes Repert. 1984;95:425-521.

Lu B.-R., Bothmer R. von. Interspecific hybridization with *Elymus himalayanus* and *E. schrenkianus*, and other *Elymus* species (Triticeae: Poaceae). Genome. 1992;35:230-237.

Lu B.-R., Salomon B. Differentiation of the SY genomes in Asiatic *Elymus.* Hereditas. 1992;116:121-126.

Lu B.-R., Salomon B., Bothmer R. von. Interspecific hybridization with *Elymus confusus* and *E. dolichaterus*, and their genomic relationships (Poaceae: Triticeae). Plant Syst. Evol. 1995;197:1-17.

Mason-Gamer R.J. Origin of North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. Syst. Bot. 2001;26:757-768.

Mason-Gamer R.J. Phylogeny of a genomically diverse group of *Elymus* (Poaceae) allopolyploids reveals multiple levels of reticulation. PLoS One. 2013;8(11):e78449. DOI 10.1371/journal.pone.0078449.

Okito P., Mott I.W., Wu Y., Wang R.R. A Y genome specific STS marker in *Pseudoroegneria* and *Elymus* species (Triticeae: Gramineae). Genome. 2009;52(4):391-400.

Shavrukov Y.N. CAPS markers in plant biology. Russ. J. Genet.: Appl. Res. 2016;6(3):279-287. DOI 10.1134/S2079059716030114.

Tsvelyov N.N. On the genus *Elymus* (Poaceae) in Russia. Botanicheskiy Zhurnal = Botanical Journal. 2008;93(10):1587-1596. (in Russian)

Tsvelyov N.N., Probatova N.S. The genera *Elymus* L., *Elytrigia* Desv., *Agropyron* Gaertn., *Psathyrostachys* Nevski, and *Leymus* Hochst. (Poaceae: Triticeae) in the flora of Russia. In: V.L. Komarov Memorial Lectures. Vladivostok: Dalnauka Publ., 2010;57:5-102. (in Russian)

Zhou Y.-H., Wu B.-H., Fu T.-H., Zheng Y.-L. Morphology, fertility and cytogenetics of intergeneric hybrid between *Roegneria kamoji* Ohwi and *Dasypyrum villosum* (L.) Candargy (Poaceae: Triticeae). J. Syst. Evol. 1999;37(2):125-130.

ORCID ID

A.V. Agafonov orcid.org/0000-0002-1403-5867
S.V. Asbaganov orcid.org/0000-0002-7482-7495

Acknowledgements. This work was supported by the state project "Estimation of the morphogenetic potential of the North Asian plant population by experimental methods" (state registration number: AAAA-A17-117012610051-5) for the Central Siberian Botanical Garden (CSBG) SB RAS and state project 0324-2019-0039-C-01 for the Institute of Cytology and Genetics SB RAS, the Russian Foundation for Basic Research (project No. 18-04-01030). Materials of the bioresource scientific collection of the CSBG SB RAS "Collections of living plants in open and closed ground", USU No. 440534 were used.

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

Received December 25, 2018. Revised August 8, 2019. Accepted October 30, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):123-130
DOI 10.18699/VJ20.607

Dynamics of the genetic diversity of oat varieties in the Tyumen region at avenin-coding loci

A.V. Lyubimova[1, 2] ✉, G.V. Tobolova[2], D.I. Eremin[2], I.G. Loskutov[3]

[1] Scientific Research Institute of Agriculture of the Northern Trans-Ural Region – Branch of the Tyumen Scientific Center of Siberian Branch of the Russian Academy of Sciences, Moskowsky village, Tyumen district, Tyumen region, Russia
[2] Northern Trans-Ural State Agricultural University, Tyumen, Russia
[3] Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
✉ e-mail: ostapenkoav88@yandex.ru

Abstract. Molecular and biochemical markers are used to analyze the intraspecific genetic diversity of crops. Prolamin-coding loci are highly effective for assessing this indicator. On the basis of the Laboratory of Varietal Seed Identification of the State Agrarian University of the Northern Trans-Urals, 18 varieties of common oat included in the State Register of Selection Achievements in the Tyumen Region from the 1930s to 2019 were studied by electrophoresis in 2018–2019. The aim of the work was to study the dynamics of the genetic diversity of oat varieties at avenin-coding loci. For the analysis, 100 grains of each variety were used. Electrophoresis was carried out in vertical plates of 13.2 % polyacrylamide gel at a constant voltage of 500 V for 4.0–4.5 h. It was found that 44.4 % of the varieties are heterogeneous, each consisting of two biotypes. For three loci, 20 alleles were identified, 10 of which were detected for the first time. The allele frequency of avenin-coding loci varied with time. In the process of variety exchange, alleles that are characteristic of varieties of non-Russian origin were replaced by alleles present in domestic varieties and then in the varieties developed by local breeding institutions. The following alleles had the highest frequency in Tyumen varieties: *Avn A4* (50.0 %), *A2* (25.0 %), *Avn B4* (50.0 %), *Bnew6* (37.5 %), *Avn C1* (37.5 %), *C2* and *C5* (25.0 %). These alleles are of great value as markers of agronomically and adaptively important characters for the region in question. The amount of genetic diversity of oats varied with time from 0.33 in 1929–1950 to up to 0.75 in 2019. The high value of genetic diversity in modern breeding varieties of the Scientific Research Institute of Agriculture of the Northern Trans-Urals and an increase in this indicator over the past 20 years are associated with the use of genetically heterogeneous source material in the breeding process. This allowed obtaining varieties with high adaptive potentials in the natural climatic conditions of the region.

Key words: oat; variety; electrophoresis; storage proteins; avenin; avenin-coding loci; alleles; genetic diversity.

For citation: Lyubimova A.V., Tobolova G.V., Eremin D.I., Loskutov I.G. Dynamics of genetic diversity of oat varieties in the Tyumen region at avenin-coding loci. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):123-130. DOI 10.18699/VJ20.607

Динамика генетического разнообразия сортов овса в Тюменской области по авенин-кодирующим локусам

А.В. Любимова[1, 2] ✉, Г.В. Тоболова[2], Д.И. Еремин[2], И.Г. Лоскутов[3]

[1] Научно-исследовательский институт сельского хозяйства Северного Зауралья – филиал Федерального исследовательского центра Тюменского научного центра Сибирского отделения Российской академии наук, пос. Московский, Тюменский район, Тюменская область, Россия
[2] Государственный аграрный университет Северного Зауралья, Тюмень, Россия
[3] Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
✉ e-mail: ostapenkoav88@yandex.ru

Аннотация. Для анализа внутривидового генетического разнообразия сельскохозяйственных культур применяются разнообразные молекулярные и биохимические маркеры. Высокой эффективностью при оценке этого показателя обладают проламин-кодирующие локусы. На базе лаборатории сортовой идентификации семян Государственного аграрного университета Северного Зауралья в 2018–2019 гг. методом электрофореза исследованы 18 сортов овса посевного, включенных в Государственный реестр селекционных достижений по Тюменской области с 1930-х гг. до 2019 г. Целью работы было изучить динамику генетического разнообразия сортов по авенин-кодирующим локусам. Для анализа использовали по 100 зерновок каждого сорта. Электрофорез проводили в вертикальных пластинах 13.2 % полиакриламидного геля при постоянном напряжении 500 В в течение 4.0–4.5 ч. Установлено, что 44.4 % сортов гетерогенны и состоят из двух биотипов. Для трех локусов идентифицировано 20 аллелей, 10 из которых выявлены впервые. Частота встречаемости аллелей авенин-кодирующих локусов изменялась с течением времени.

Аллели, характерные для сортов иностранного происхождения, в процессе сортосмены заместились аллелями, присутствующими в отечественных сортах, а затем в сортах местных селекционных учреждений. Наибольшую частоту встречаемости в сортах тюменской селекции имели аллели *Avn A4* (50.0 %), *A2* (25.0 %), *Avn B4* (50.0 %), *Bnew6* (37.5 %), *Avn C1* (37.5 %), *C2* и *C5* (25.0 %). Эти аллели имеют большую ценность как маркеры хозяйственно ценных и адаптивно значимых признаков. Величина генетического разнообразия в сортах овса изменялась с течением времени от 0.33 в 1929–1950 гг. до 0.75 в 2019 г. Высокое значение генетического разнообразия в современных сортах селекции Научно-исследовательского института сельского хозяйства Северного Зауралья, а также увеличение этого показателя на протяжении последних 20 лет связаны с использованием в селекционном процессе генетически разнородного исходного материала. Это позволило получить сорта, обладающие высоким адаптивным потенциалом в природно-климатических условиях данного региона.
Ключевые слова: овес; сорт; электрофорез; запасные белки; авенин; авенин-кодирующие локусы; аллель; генетическое разнообразие.

Introduction

Common oat (*Avena sativa* L.) is a valuable agricultural crop used both for food and animal feed (Barsila, 2018). An important factor in increasing the production of oat is the creation of new intensive type varieties characterized by high productivity and environmental sustainability (Goncharenko, 2016). In the Tyumen region, breeding work with this culture is very active. From the first half of the twentieth century to the present, 18 varieties of spring oat have been included in the State Register of Selection Achievements in the region. In 1993, the first variety of local breeding, Megion, was regionalized. The proportion of varieties created by the Scientific Research Institute of Agriculture of the Northern Trans-Urals in the region's crops has since been constantly increasing. Nowadays, only varieties of local breeding are included in the State Register of Selection Achievements in the region.

However, active breeding can lead to a decrease in the genetic diversity of the species. This is due to the frequent involvement of the same genotypes in the breeding process to enhance specific agronomic characters. A decrease in genetic diversity negatively affects the resistance of populations to diseases and the populations' ability to adapt to changing environmental and climatic conditions (Novoselskaya-Dragovich et al., 2007; Afanasenko, Novozhilov, 2009; Goncharenko, 2016).

A variety of molecular and biochemical markers are used to analyze intraspecific genetic diversity (Konarev et al., 2000; Montilla-Bascón et al., 2013; Shavrukov, 2016; Scheben et al., 2017). Prolamin-coding loci are very effective for assessing this indicator (Che, Li, 2007; Melnikova et al., 2010; Kudryavtsev et al., 2014; Lyalina et al., 2016; Lyubimova, Eremin, 2018a; Zobova et al., 2018; Utebayev et al., 2019). Prolamins of oat (avenins) are inherited as blocks and are controlled by three independent loci: *Avn A*, *Avn B* and *Avn C*, located in three homeologous chromosomes of group A (Portyanko et al., 1987, 1998). Due to the high level of avenin polymorphism, almost every oat variety, biotype, or line is characterized by a unique component composition of storage proteins (Loscutov, 2007; Lyubimova, Eremin, 2018b). This allows analyzing the individual allele frequency of avenin-coding loci, the dynamics of changes in their occurrence in time and space, and also assessing the genetic transformations that occur under the influence of prolonged artificial selection.

The aim of the work is to study the dynamics of genetic diversity at avenin-coding loci in common oat varieties included in the State Register of Selection Achievements in the Tyumen region from the 1930s to the present for assessing the effectiveness of selection work carried out in the region.

Materials and methods

The studies were carried out in the Laboratory of Varietal Seed Identification of the Agrobiotechnological Center of the Northern Trans-Urals State Agrarian University in 2018–2019. Eighteen varieties of common oat included in the State Register of Selection Achievements in the Tyumen Region since 1929 were studied (Table 1).

Plant material was provided from the collection of the Federal Research Center N.I. Vavilov All-Russian Institute of Plant Genetic Resources and the institution-originator of varieties, the Scientific Research Institute of Agriculture of the Northern Trans-Urals, a Branch of the Tyumen Scientific Center of Siberian Branch of the Russian Academy of Sciences.

For laboratory analysis, 100 grains of each variety selected by random sampling were used. For one-dimensional electrophoresis of avenins, a published technique (Portyanko et al., 1998) with modifications was used. Proteins were extracted from individual crushed grains by adding 90 µl of 70 % ethanol. The obtained extract was centrifuged, and 300 µl of methylene green dye was added to it. Protein extract (22 µl) was added to the polyacrylamide gel. Gel composition: 13.17 g of acrylamide, 0.66 g of N,N′-methylene-bis-acrylamide, 7.17 g of urea, 2.0 mg of iron sulfate (III), 80.0 mg of ascorbic acid, and 0.26 g of aluminum lactate. All reagents were dissolved in 100 ml aluminum-lactate buffer (pH 3.1). Acrylamide polymerization was initiated by adding 25 µl of 15 % hydrogen peroxide to 75 ml of

Table 1. Varieties of common oat included in the State Register of Selection Achievements in the Tyumen region (1929–2019)

VIR catalog number	Variety	Origin	Year of regionalization	Year of removal from regionalization	Total in regionalization (years)
7965	Seger	Sweden	1929	1963	34
7947	Golden Rain		1929	1976	47
8494	Omhafer		1939	1982	43
8256	Udarnik 883	Krasnoyarsk region	1957	1960	3
2874	Nidar	Norway	1957	1963	6
11132	Severyanin	Arkhangelsk region	1966	1974	8
11717	Skorospelyj	Kirov region	1974	1981	7
11122	Narymskij 943	Tomsk region	1975	1996	21
12245	Tayozhnik		1977	2001	24
11379	Astor	Netherlands	1978	2000	22
11584	Selma	Sweden	1981	1993	12
13478	Perona	Netherlands	1985	2018	33
14039	Megion	Tyumen region	1993	–	26
14031	Novosibirskij 88	Novosibirsk region	1994	2004	10
14784	Tyumenskij golozyornyj	Tyumen region	2000	–	19
14785	Talisman		2002	–	17
15380	Otrada		2014	–	5
15451	Foma		2015	–	4

a gel solution. Electrophoresis was carried out in vertical electrophoretic chambers with dimensions of the formed plates of 17.8 × 17.8 × 0.15 cm (VE-20, Helicon, Russia) for 4.0–4.5 h at a constant voltage of 500 V. To fix and stain the gel, a 10 % solution of trichloroacetic acid with the addition of 0.05 % Coomassie brilliant blue R-250 in ethanol was used. Identification of allelic variants of component blocks controlled by avenin-coding loci was carried out on the basis of a catalog developed by V.A. Portyanko et al. (1987). Astor common oat (*Avn A2 B4 C2*) were used as a standard. In case the detected block was not in the catalog, it was marked with a "new" mark.

In order to assess the dynamics of the change in the genetic diversity of oat varieties over time, all the studied samples were grouped. One group included varieties cultivated in the same ten-year period. The gene diversity at the locus (H) was calculated for each group of varieties separately according to the following formula:

$$H = \frac{n}{n-1} \times (1 - \Sigma_{i=1}^{k} p_i^2),$$

where p_i is the population frequency of the i-th allele; k is the number of locus alleles; n is the sample size (Nei, 1987). To calculate the average gene diversity ($\overline{H}$), the number of alleles per locus was averaged over all loci. The calculations were performed using the Arlequin Ver 3.5.2.2 program (Copyright 2015 L. Excoffier. CMPG, University of Berne).

Results

As a result of the studies, it was found that 8 (44.4 %) of the 18 analyzed varieties were heterogeneous in the composition of avenin. Seger, Golden Rain, Omhafer, Severyanin, Narymskij 943, Tayozhnik, Megion and Otrada varieties consisted of two biotypes. These varieties are characterized by the presence of several alleles at one or more avenin-coding loci. In the genetic formula, such states of loci were recorded with the "+" sign (Table 2). In subsequent calculations, each biotype was considered by us as a separate sample. A total of 26 samples were examined.

An analysis of the electrophoretic spectra of avenin allowed us to describe the genetic formulas for each of the studied varieties. Altogether, 8 alleles were detected for the *Avn A* locus; 7, for the *Avn B* locus; and 5, for the *Avn C* locus. It should be noted that some of the combinations of avenin components that we found were absent in the catalog of genetic nomenclature. To identify new blocks

Table 2. Alleles of avenin-coding loci of common oat varieties included in the State Register of Selection Achievements in the Tyumen region (1929–2019)

Variety	Number of biotypes	Alleles of the avenin-coding locus		
		Avn A	*Avn B*	*Avn C*
Seger	2	*2+new9*	*1*	*new8*
Golden Rain	2	*2*	*1*	*2+new8*
Omhafer	2	*new9*	*1*	*2+new8*
Udarnik 883	1	*new11*	*new9*	*3*
Nidar	1	*2*	*1*	*2*
Severyanin	2	*new11+new12*	*new9+new10*	*3*
Skorospelyj	1	*new12*	*new10*	*3*
Narymskij 943	2	*5+2*	*1+4*	*1+2*
Tayozhnik	2	*2+1*	*new8*	*2*
Astor	1	*2*	*4*	*2*
Selma	1	*new9*	*1*	*3*
Perona	1	*4*	*4*	*2*
Megion	2	*2+new11*	*new6*	*5*
Novosibirskij 88	1	*2*	*4*	*2*
Tyumenskij golozyornyj	1	*2*	*new6*	*3*
Talisman	1	*4*	*4*	*2*
Otrada	2	*new10+4*	*4*	*1*
Foma	1	*4*	*new7*	*1*

Table 3. The allele frequency of avenin-coding loci of common oat varieties, %

Locus	Allele	Years								
		1929–1930	1940–1950	1950–1960	1960–1970	1970–1980	1980–1990	1990–2000	2000–2010	2010–2019
Avn A	*1*	0	0	0	0	8.3	12.5	10.0	12.5	0
	2	60.0	60.0	50.0	44.4	41.7	37.5	50.0	50.0	25.0
	4	0	0	0	0	0	12.5	10.0	25.0	50.0
	5	0	0	0	0	8.3	12.5	10.0	0	0
	new9	40.0	40.0	37.5	33.3	16.7	12.5	10.0	0	0
	new10	0	0	0	0	0	0	0	0	12.5
	new11	0	0	12.5	11.1	8.3	0	10.0	12.5	12.5
	new12	0	0	0	11.1	16.7	12.5	0	0	0
Avn B	*1*	100.0	100.0	87.5	77.8	41.7	25.0	20.0	0	0
	4	0	0	0	0	16.7	37.5	40.0	37.5	50.0
	new6	0	0	0	0	0	0	20.0	37.5	37.5
	new7	0	0	0	0	0	0	0	0	12.5
	new8	0	0	0	0	16.7	25.0	20.0	25.0	0
	new9	0	0	12.5	11.1	8.3	0	0	0	0
	new10	0	0	0	11.1	16.7	12.5	0	0	0
Avn C	*1*	0	0	0	0	8.3	12.5	10	0	37.5
	2	33.3	33.3	37.5	33.3	50.0	62.5	60.0	62.5	25.0
	3	0	0	12.5	22.2	25.0	25.0	10.0	12.5	12.5
	5	0	0	0	0	0	0	20.0	25.0	25.0
	new8	66.7	66.7	50.0	44.4	16.7	0	0	0	0

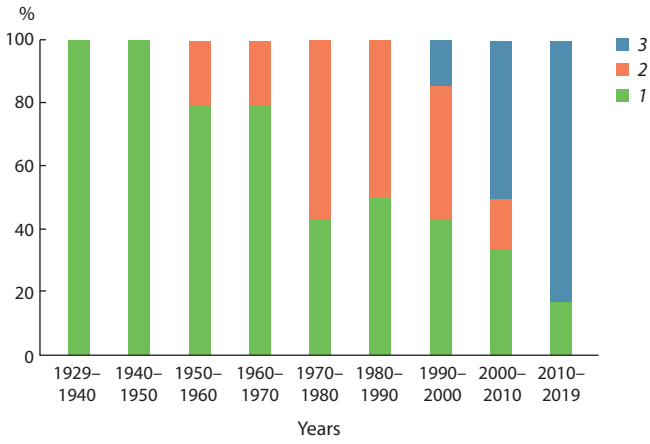


Fig. 1. The dynamics of the regionalized assortment of common oat in the Tyumen region (1929–2019).

Varieties: *1* – of foreign breeding; *2* – of domestic breeding; *3* – of local breeding institutions.

of components, it is necessary to conduct a hybridological analysis and assess the nature of the inheritance of avenin components. However, we highlighted the alleged blocks of components, each of which was assigned a number following the blocks previously described in the catalog. A "new" mark was added before the number of each of the proposed blocks.

To assess genetic diversity at different time intervals, we calculated the allele frequency of avenin-coding loci (Table 3).

Different alleles predominate in different groups of varieties. For the *Avn A* locus, only alleles *2* and *new9* were found before 1950. However, the frequency of their occurrence began to decrease with the appearance of domestic varieties and then varieties of local breeding in the crops of the region (Fig. 1). Alleles *1*, *5* and *new12* were characteristic of varieties cultivated from 1960 to 2010, and are no longer found today. Allele *4* (50.0 %) is currently the most widespread; allele *2* accounts for 25.0 %; *new11* and *new12*, for 12.5 % each.

For the *Avn B* locus of modern oat varieties, alleles *4* (50.0 %) and *new6* (37.5 %) predominate; *new7* is found with a frequency of 12.5 %. Alleles *1*, *new8*, *new9* and *new10*, which are characteristic of varieties of foreign and domestic breeding, but not found by us among the varieties of local breeding, are completely eliminated.

A similar situation is observed for the *Avn C* locus: allele *new8*, which occurred with a frequency of 66.7 % in 1929–1950, is currently replaced by alleles *1* (37.5 %), *5* (25.0 %) and *3* (12.5 %). It is necessary to pay attention to allele *2*, the presence of which in varieties has been noted at all periods of cultivation ever since 1929. This allele frequency ranged from 25.0 to 62.5 %. Nowadays, this allele is presented in 25.0 % of the varieties. The same feature was noted for allele *2* of the *Avn A* locus.

The value of genetic diversity, calculated on the basis of data on the allele frequency, also changed over time (Fig. 2).

This indicator was minimal before 1950 (0.38), when only three varieties of oat were cultivated in the region: Seger, Golden Rain and Omhafer. Subsequently, with the advent of new varieties in the region's crops, the value of genetic diversity increased, reaching its maximum in the period from 1970 to 1980 (0.78). During this period of time, an active variety exchange was carried out in the region – Seger, Udarnik 883 and Nidar were removed from regionalization, and they were replaced by Skorospelyj, Narymskij 943, Tayozhnik and Astor varieties bearing new alleles of avenin-coding loci. The period 1970–1980 was

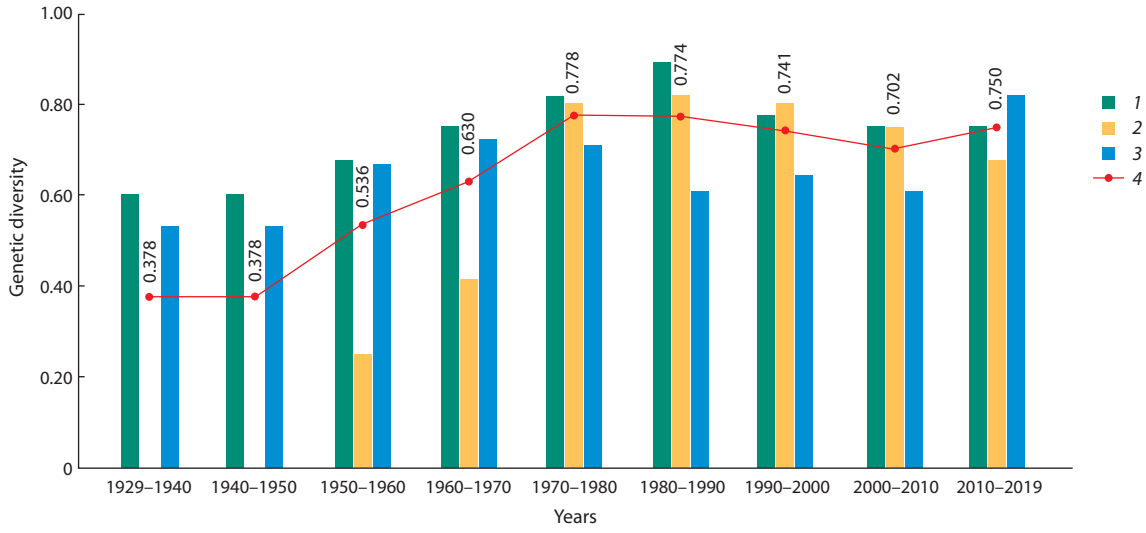


Fig. 2. Genetic diversity of oat varieties by avenin-coding loci.

Locuses: *1 – Avn A*; *2 – Avn B*; *3 – Avn C*; *4* – average gene diversity.

characterized by the largest variety of allelic variants in varieties – 15 alleles were found at three *Avn* loci (Table 3). Subsequently, in the process of replacing foreign varieties with domestic ones, the indicator of genetic diversity decreased to 0.70 by 2010. A decrease in diversity was caused by the exclusion from regionalization of a large number of varieties bearing alleles not found in varieties of local breeding. However, to date, there has been an increase in average gene diversity to 0.75.

Discussion

As a result of our analysis using multiple alleles of avenin-coding loci, we described the genetic formulas for 18 varieties of common oat included in the State Register of Selection Achievements in the Tyumen region. It was established that the heterogeneity of varieties is 44.4 %. The presence of several biotypes increases the adaptive potential of the variety (Metakovsky, 1990; Novoselskaya-Dragovich et al., 2013), which is extremely important in the natural climatic conditions of the Tyumen region, which is a risky farming zone.

In some varieties, identical prolamin spectra were found. Thus, the first and second biotypes of Golden Rain are identical to Seger (I biotype) (*2.1.new8*) and Nidar (*2.1.2*), respectively. The second biotype of Seger coincided with the second biotype of Omhafer (*new9.1.new8*). The first and second biotypes of Severyanin coincide with Udarnik (*new11.new9.3*) and Skorospelyj (*new12.new10.3*). The same types of spectra are characteristic of the second biotype of Narymskij 943 as well as Astor and Novosibirskij 88 (*2.4.2*); the spectra of Perona and Talisman (*4.4.2*) coincide. As a result of the analysis, it was found that only 10 (38.5 %) of the 26 studied genotypes are variety-specific. This is a fairly low rate.

The identity of alleles of prolamin-coding loci in varieties is associated with the involvement of the same genotypes in breeding programs (Portyanko et al., 1998; Melnikova et al., 2010; Novoselskaya-Dragovich et al., 2013). For example, Seger and Golden Rain were bred from the same oat variety Milton (=Propsteier), and Omhafer, too, originated from it. The old oat variety Milton appeared in northern Germany and was widespread in northern Europe (Portyanko et al., 1998). Apparently, this variety possessed outstanding economic characteristics, which led to its frequent inclusion in the breeding process. This was reflected in the matching set of alleles of avenin-coding loci in its descendants. The presence of varieties with the same genetic formulas of prolamins reduces the efficiency of using the method of electrophoresis for their differentiation. A number of authors in their studies concluded that the use of avenin-coding loci as the only marker system for distinguishing a large number of oat varieties is insufficient, since the allelic diversity of oat prolamin loci is characterized as low compared to wheat, barley and rye (Cliff, Cooke, 1984; Souza, Sorrels, 1990; Portyanko et al., 1998). In such cases, there is a need for an additional use of other marker systems (Wight et al., 2010). However, it should be noted that the modern varieties of oat created by the Scientific Research Institute of Agriculture of the Northern Trans-Urals have an individual allelic composition of avenin-encoding loci, which makes it possible to differentiate their genotypes with high accuracy.

An analysis of the frequencies of alleles of avenin-coding loci for all three loci allowed us to note the relationship between the frequency of alleles and a set of cultivated varieties, especially their origin. In the process of variety exchange, the alleles characteristic of varieties of foreign breeding were gradually replaced by alleles present in domestic varieties, and then in the varieties developed of local breeding istitutions. A similar replacement of one allele with another during breeding work was noted by many researchers in the study of prolamin-coding loci of wheat and barley (Novoselskaya-Dragovich et al., 2007; Lyalina et al., 2016). On a large number of examples, the adaptive nature of prolamin polymorphism has been proved. Their connection with adaptive gene complexes allows, based on the spectra of storage proteins, identifying genotypes that are most adapted to specific climatic conditions. A.Yu. Novoselskaya-Dragovich and the co-workers (2013) noted that genetic differences between varieties of different geographical origin are determined by natural selection. In this case, the reason for the rather rapid replacement of the "old" alleles with "new" ones is directed processes associated with new directions in breeding and the involvement of genetically different source material (Novoselskaya-Dragovich et al., 2007). Our data on the allele frequency of avenin-coding loci are in good agreement with this statement.

With the beginning of breeding work on oat in the Tyumen region, varieties appeared that possess a set of agronomically and adaptively significant characters for this region. This led to an increase in the frequency of certain alleles of avenin-coding loci, which can be considered markers of such genotypes or characters. At the same time, it caused a decrease in the frequency or even complete disappearance of alleles characteristic of foreign varieties. *A2* and *C2* alleles found in all groups of varieties, probably, mark highly competitive gene associations that give their carriers important advantages in the natural climatic conditions of the region.

Monitoring changes in the genetic diversity of varieties over time allows judging the presence or absence of genetic erosion. In the works devoted to the assessment of genetic diversity in varieties of other crops, its values were 0.62–0.76 for soft wheat varieties created in Serbia and Italy (Novoselskaya-Dragovich et al., 2007), 0.5–0.6 in soft

wheat varieties of Ukrainian selection (Zayka et al., 2014), and 0.42–0.64 in groups of durum wheat varieties originating from different countries of the world (Kudryavtsev et al., 2014). At the same time, a decrease in the value of this indicator in modern varieties is noted (Kudryavtsev et al., 2014; Lyalina et al., 2016).

The high values of genetic diversity identified as a result of our work and an increase in this indicator since 2000 indicate the absence of genetic erosion. It should be noted that, at different periods of time, the contribution of individual avenin-coding loci to the average gene diversity in varieties of oat in the region was not the same. In the period from 1970 to 2010, the *Avn A* and *Avn B* loci played an important role in the formation of genetic diversity. But currently, the maximum genetic diversity is observed at the *Avn C* locus. In our opinion, this suggests that the alleles of this locus may be important as markers of adaptively significant characters.

Conclusion

The allele frequency of avenin-coding loci in varieties of common oat included in the State Register of Selection Achievements in the Tyumen region from 1929 to 2019 changed over time. The alleles characteristic of the varieties of foreign selection were replaced by "new" ones, specific to the varieties of local selection: *Avn A4* (50.0 %), *A2* (25.0 %), *Avn B4* (50.0 %), *Bnew6* (37.5 %), *Avn C1* (37.5 %), *C2* and *C5* (25.0 %). These alleles are of great value as markers of agronomically and adaptively significant characters for the region in question.

Modern regionalized varieties of oat are characterized by high genetic diversity (0.75), which is associated with the use of heterogeneous source material in the breeding process. This allows obtaining varieties with high adaptive potentials in the climatic conditions of Western Siberia.

The high importance of genetic diversity in modern breeding varieties of the Scientific Research Institute of Agriculture of the Northern Trans-Urals and an increase in this indicator over the past 20 years indicate competently organized and effective breeding work with this crop in the Tyumen region.

References

Afanasenko O.S., Novozhilov K.V. Problems of rational use of genetic resources of plant resistance to diseases. Ekologicheskaya Genetika = Ecological Genetics. 2009;7(2):38-43. (in Russian)

Barsila S.R. The fodder oat (*Avena sativa*) mixed legume forage farming: Nutritional and ecological benefits. J. Agric. Nat. Resour. 2018;1(1):206-222. DOI 10.3126/janr.v1i1.22236.

Che Y.H., Li L.H. Genetic diversity of prolamines in *Agropyron mongolicum* Keng indigenous to northern China. Genet. Resour. Crop. Evol. 2007;54(5):1145-1151. DOI 10.1007/s10722-006-9006-7.

Cliff E.M., Cooke R.J. The characterisation of oat cultivars by electrophoresis. J. Natl. Inst. Agric. Bot. 1984;16:415-429.

Goncharenko A.A. Ecological stability of grain crop varieties and tasks of breedeng. Zernovoe Khozjaistvo Rossii = Grain Economy of Russia. 2016;3:31-37. (in Russian)

Konarev A.V., Konarev V.G., Gubareva N.K., Peneva T.I. Seed proteins as markers in solving problems of plant genetic resources, breeding, and seed production. Tsitologiya i Genetika = Cytology and Genetics. 2000;34(2):91-104. (in Russian)

Kudryavtsev A.M., Dedova L.V., Melnik V.A., Shishkina A.A., Upelniek V.P., Novoselskaya-Dragovich A.Yu. Genetic diversity of modern Russian durum wheat cultivars at the gliadin-coding loci. Russ. J. Genet. 2014;50(5):483-48. DOI 10.1134/S1022795414050093.

Loscutov I.G. Oat (*Avena* L.). Distribution, Taxonomy, Evolution and Breeding Value. Saint-Petersburg, 2007. (in Russian)

Lyalina E.V., Boldyrev S.V., Pomortsev A.A. Current state of the genetic polymorphism in spring barley (*Hordeum vulgare* L.) from Russia assessed by the alleles of hordein-coding loci. Russ. J. Genet. 2016;52(6):565-577. DOI 10.1134/S1022795416060077.

Lyubimova A.V., Eremin D.I. Peculiarities of the avenin component composition in oats cultivated in Western Siberia. Trudy po Prikladnoy Botanike, Genetike i Selektsii = Proceedings on Applied Botany, Genetics, and Breeding. 2018a;179(2):85-95. DOI 10.30901/ 2227-8834-2018-2-85-95. (in Russian)

Lyubimova A., Eremin D. Variety check as an element of precision farming in the modern agriculture. In: International Scientific and Practical Conference "AgroSMART – Smart Solutions for Agriculture" (AgroSMART 2018). 2018b. Vol. 151. DOI 10.2991/agrosmart-18.2018.91.

Melnikova N.V., Kudryavtsev A.M., Mitrofanova O.P., Liapunova O.A. Global diversity of durum wheat *Triticum durum* Desf. for alleles of gliadin-coding loci. Russ. J. Genet. 2010;46(1):43-49. DOI 10.1134/S1022795410010072.

Metakovsky E.V. The value of gliadin biotypes in commercial cultivars of wheat. In: Proc. 4th Int. Workshop on Gluten Proteins. Winnipeg, 1990.

Montilla-Bascón G., Sánchez-Martín J., Rispail N., Rubiales D., Mur L., Langdon T., Griffiths I., Howarth C., Prats E. Genetic diversity and population structure among oat cultivars and landraces. Plant Mol. Biol. Rep. 2013;31(1305):1305-1314. DOI 10.1007/s11105-013-0598-8.

Nei M. Molecular Evolutionary Genetics. New York: Columbia University Press, 1987.

Novoselskaya-Dragovich A.Yu., Fisenko A.V., Imasheva A.G., Pukhalskiy V.A. Comparative analysis of the genetic diversity dynamics at gliadin loci in the winter common wheat *Triticum aestivum* L. cultivars developed in Serbia and Italy over 40 years of scientific breeding. Russ. J. Genet. 2007;43(11):1236-1242. DOI 10.1134/S1022795407110051.

Novoselskaya-Dragovich A.Y., Fisenko A.V., Puhal'skii V.A. Genetic differentiation of common wheat cultivars using multiple alleles of gliadin-coding loci. Russ. J. Genet. 2013;49(5):487-496. DOI 10.1134/S1022795413020087.

Portyanko V.A., Pomortsev A.A., Kalashnik N.A., Bogachkov V.I., Sozinov A.A. The genetic control of avenins and the principles of classification. Genetika = Genetics (Moscow). 1987;23(5):845-853. (in Russian)

Portyanko V.A., Sharopova N.R., Sozinov A.A. Characterization of European oat germ plasm: allelic variation at complex avenin loci detected by acid polyacrylamide gel electrophoresis. Euphytica. 1998;102:15-27. DOI 10.1023/A:1018399919953.

Scheben A., Batley J., Edwards D. Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. Plant Biotechnol. J. 2017;15(2):149-161. DOI 10.1111/pbi.12645.

Shavrukov Y. Comparison of SNP and CAPS markers application in genetic research in wheat and barley. BMC Plant Biol. 2016;16:11. DOI 10.1186/s12870-015-0689-9.

Souza E., Sorrels M. Inheritance and distribution of variation at four avenin loci in North American oat germplasm. Genome. 1990;33:416-424. DOI 10.1139/g90-063.

Utebayev M., Dashkevich S., Bome N., Bulatova K., Shavrukov Y. Genetic diversity of gliadin-coding alleles in bread wheat (*Triticum aestivum* L.) from Northen Kazakhstan. PeerJ. 2019; 7:e7082. DOI 10.7717/peerj.7082.

Wight C.P., Yan W., Fetch J.M., Deyl J., Tinker N.A. A set of new simple sequence repeat and avenin DNA markers suitable for mapping and fingerprinting studies in oat (*Avena* spp.). Crop Sci. 2010;50:1207-1218. DOI 10.2135/cropsci2009.09.0474.

Zayka E.V., Kozub N.A., Sozinov I.A., Sozinov A.A., Starichenko V.N. Analysis of genotypes of winter soft wheat varieties of the Institute of Farming of the National Academy of Agricultural Science according to alleles of storage protein loci. Vestnik Belorusskoy Gosudarstvennoy Selskohozyaystvennoy Akademii = Bulletin of the Belarussian State Agricultural Academy. 2014;4:53-57. (in Russian)

Zobova N.V., Surin N.A., Gerasimov S.A., Chuslin A.A., Onufrienok T.V. Spectra of prolamines in agroecological evaluation of the collection barley. Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC. 2018; 32(5):45-47. DOI 10.24411/0235-2451-2018-10511. (in Russian)

ORCID ID
A.V. Lyubimova orcid.org/0000-0002-1570-9595
D.I. Eremin orcid.org/0000-0002-3672-6060
I.G. Loskutov orcid.org/0000-0002-9250-7225

Acknowledgements. The work was carried out according to the state task (Priority area X.10.4, Program X.10.4.148, Project X.10.4.148).
Conflict of interest. The authors declare no conflict of interest.
Recieved June 23, 2019. Revised October 30, 2019. Accepted October 30, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):131-138
DOI 10.18699/VJ20.608

Stem rust in Western Siberia – race composition and effective resistance genes

V.P. Shamanin[1] ✉, I.V. Pototskaya[1], S.S. Shepelev[1], V.E. Pozherukova[1], E.A. Salina[2], E.S. Skolotneva[2], D. Hodson[3], M. Hovmøller[4], M. Patpour[4], A.I. Morgounov[5]

[1] Omsk State Agrarian University named after P.A. Stolypin, Omsk, Russia
[2] Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] CIMMYT, Addis Ababa, Ethiopia
[4] Aarhus University, Slagelse, Denmark
[5] CIMMYT-Turkey, Ankara, Turkey
✉ e-mail: vp.shamanin@omgau.org

Abstract. Stem rust in recent years has acquired an epiphytotic character, causing significant economic damage for wheat production in some parts of Western Siberia. On the basis of a race composition study of the stem rust populations collected in 2016–2017 in Omsk region and Altai Krai, 13 pathotypes in Omsk population and 10 in Altai population were identified. The race differentiation of stem rust using a tester set of 20 North American *Sr* genes differentiator lines was carried out. The genes of stem rust pathotypes of the Omsk population are avirulent only to the resistance gene *Sr31*, Altai isolates are avirulent not only to *Sr31*, but also to *Sr24*, and *Sr30*. A low frequency of virulence (10–25 %) of the Omsk population pathotypes was found for *Sr11*, *Sr24*, *Sr30*, and for Altai population – *Sr7b*, *Sr9b*, *Sr11*, *SrTmp*, which are ineffective in Omsk region. Field evaluations of resistance to stem rust were made in 2016–2018 in Omsk region in the varieties and spring wheat lines from three different sources. The first set included 58 lines and spring bread wheat varieties with identified *Sr* genes – the so-called trap nursery (ISRTN – International Stem Rust Trap Nursery). The second set included spring wheat lines from the Arsenal collection, that were previously selected according to a complex of economically valuable traits, with genes for resistance to stem rust, including genes introgressed into the common wheat genome from wild cereal species. The third set included spring bread wheat varieties created in the Omsk State Agrarian University within the framework of a shuttle breeding program, with a synthetic wheat with the *Ae. tauschii* genome in their pedigrees. It was established that the resistance genes *Sr31*, *Sr40*, *Sr2 complex* are effective against stem rust in the conditions of Western Siberia. The following sources with effective *Sr* genes were selected: (Benno)/6*LMPG-6 DK42, Seri 82, Cham 10, Bacanora (*Sr31*), RL 6087 Dyck (*Sr40*), Amigo (*Sr24*, *1RS-Am*), Siouxland (*Sr24*, *Sr31*), Roughrider (*Sr6*, *Sr36*), Sisson (*Sr6*, *Sr31*, *Sr36*), and Fleming (*Sr6*, *Sr24*, *Sr36*, *1RS-Am*), Pavon 76 (*Sr2 complex*) from the ISRTN nursery; No. 1 BC_1F_2 (96×113)×145×113 (*Sr2*, *Sr36*, *Sr44*), No. 14a F_3 (96×113)×145 (*Sr36*, *Sr44*), No. 19 BC_2F_3 (96×113)×113 (*Sr2*, *Sr36*, *Sr44*), and No. 20 F_3 (96×113)×145 (*Sr2*, *Sr36*, *Sr40*, *Sr44*) from the Arsenal collection; and the Omsk State Agrarian University varieties Element 22 (*Sr31*, *Sr35*), Lutescens 27-12, Lutescens 87-12 (*Sr23*, *Sr36*), Lutescens 70-13, and Lutescens 87-13 (*Sr23*, *Sr31*, *Sr36*). These sources are recommended for inclusion in the breeding process for developing stem rust resistant varieties in the region.

Key words: bread wheat; stem rust; pathotype; effective resistance genes; breeding.

For citation: Shamanin V.P., Pototskaya I.V., Shepelev S.S., Pozherukova V.E., Salina E.A., Skolotneva E.S., Hodson D., Hovmøller M., Patpour M., Morgounov A.I. Stem rust in Western Siberia – race composition and effective resistance genes. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):131-138. DOI 10.18699/VJ20.608

Стеблевая ржавчина в Западной Сибири – расовый состав и эффективные гены устойчивости

В.П. Шаманин[1] ✉, И.В. Потоцкая[1], С.С. Шепелев[1], В.Е. Пожерукова[1], Е.А. Салина[2], Е.С. Сколотнева[2], Д. Ходсон[3], М. Хоумвёллер[4], М. Патпур[4], А.И. Моргунов[5]

[1] Омский государственный аграрный университет им. П.А. Столыпина, Омск, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] CIMMYT, Аддис-Абеба, Эфиопия
[4] Орхусский университет, Слагельсе, Дания
[5] Представительство CIMMYT, Анкара, Турция
✉ e-mail: vp.shamanin@omgau.org

Аннотация. Стеблевая ржавчина пшеницы в последние годы приобрела эпифитотийный характер, нанося значительный экономический ущерб производству зерна пшеницы в отдельных областях Западной Сибири. По результатам изучения расового состава популяций стеблевой ржавчины, собранной в 2016–2017 гг.

в Омской области и Алтайском крае, выявлено 13 патотипов в омской популяции и 10 – в алтайской. Дифференцирование рас стеблевой ржавчины проводили с помощью тестерного набора 20 североамериканских линий-дифференциаторов *Sr* генов. Гены патотипов стеблевой ржавчины омской популяции авирулентны только к гену устойчивости *Sr31*, алтайские изоляты авирулентны, помимо *Sr31*, к генам *Sr24*, *Sr30*. Низкая частота вирулентности (10–25 %) патотипов омской популяции установлена для *Sr11*, *Sr24*, *Sr30*, а патотипов алтайской – для *Sr7b*, *Sr9b*, *Sr11*, *SrTmp*, которые неэффективны в Омской области. Полевая оценка устойчивости к стеблевой ржавчине проводилась в 2016–2018 гг. в Омской области в динамике в течение вегетационного периода у сортов и линий мягкой пшеницы из трех различных источников. Первый набор включал 58 линий и сортов яровой мягкой пшеницы с идентифицированными генами *Sr*, условно называемыми «питомник-ловушка» (ISRTN – international stem rust trap nursery). Второй набор включал линии яровой пшеницы из коллекции «Арсенал», отобранные ранее по комплексу хозяйственно ценных признаков и несущие пирамиду генов устойчивости к стеблевой ржавчине, в том числе интрогрессированных в геном мягкой пшеницы от дикорастущих видов злаков. Третий набор включал сорта яровой мягкой пшеницы, созданные в Омском аграрном университете по программе челночной селекции, имеющие в родословной синтетическую пшеницу с геномом *Ae. tauschii*. Установлено, что линии с генами *Sr31*, *Sr40*, *Sr2 complex* невосприимчивы к стеблевой ржавчине в условиях Западно-Сибирского региона. Выделены источники с эффективными генами *Sr*: из питомника ISRTN – (Benno)/6*LMPG-6 DK42 (*Sr31*), Seri 82 (*Sr31*), Cham 10 (*Sr31*), Bacanora (*Sr31*), RL 6087 Dyck (*Sr40*), Amigo (*Sr24*, *1RS-Am*), Siouxland (*Sr24*, *Sr31*), Roughrider (*Sr6*, *Sr36*), Sisson (*Sr6*, *Sr31*, *Sr36*), Fleming (*Sr6*, *Sr24*, *Sr36*, *1RS-Am*), Pavon 76 (*Sr2 complex*); из коллекции «Арсенал» – № 1 BC_1F_2 (96×113)×145×113 (*Sr2*, *Sr36*, *Sr44*), № 14a F_3 (96×113)×145 (*Sr36*, *Sr44*), № 19 BC_2F_3 (96×113)×113 (*Sr2*, *Sr36*, *Sr44*), № 20 F_3 (96×113)×145 (*Sr2*, *Sr36*, *Sr40*, *Sr44*); сорта Омского аграрного университета – Элемент 22 (*Sr31*, *Sr35*), Лютесценс 27-12, Лютесценс 87-12 (*Sr23*, *Sr36*), Лютесценс 70-13, Лютесценс 87-13 (*Sr23*, *Sr31*, *Sr36*). Выделенные источники рекомендуются для включения в селекционный процесс при создании сортов, устойчивых к стеблевой ржавчине в условиях региона.

Ключевые слова: мягкая пшеница; стеблевая ржавчина; патотип; эффективные гены устойчивости; селекция.

Introduction

Stem rust of wheat caused by *Puccinia graminis* f. sp. *tritici* Erikss. for a long time had a weak manifestation in the territory of Western Siberia and only in the recent years acquired an epiphytotic nature, causing significant economic damage for wheat production in the region. First of all, this is due to the deterioration of the phytosanitary situation in the region, the general trend of climate warming and cultivation of susceptible wheat varieties on large area (Shamanin et al., 2015, 2016a). The threat of stem rust race *Ug99* appearance and the emergence of new pathotypes of this race, affecting varieties with genes *Sr24* and *Sr36* present a serious threat for wheat production in West Siberian region. Genetic diversity of cultivated wheat varieties for resistance to *Ug99* and stem rust in general is very limited (Shamanin et al., 2016b).

Enhancement of genetic resistance to pathogens can be solved germplasm exchange, and also cultivation of varieties with different level of resistance to diseases and to different races. Crop protection is necessary to restrain the evolution of pathogens and the emergence of new virulent races. Such programs are widely used in Europe and America. The duration of the variety cultivation in advanced countries is 3–4 years, while in Russia – 7–10 years (Sanin, 2016). In this regard, the breeding of spring wheat varieties, which have a diverse genetic basis of resistance to stem rust, is very relevant.

Since the 1950s, many resistance genes introduced into bread wheat have lost their effectiveness (Singh et al., 2008). The most significant genes for breeding practice are *Sr2*, *Sr23*, *Sr24*, *Sr25*, *Sr31*, *Sr33*, *Sr36*, *Sr38*, *Sr45*, *Sr50*, *SrTmp*, $Sr1RS^{Amigo}$ (Singh et al., 2015).

Introgression of resistance genes of wild and cultivated wheat relatives allows to expand the genetic diversity of varieties and contributes to their long-term protection (Leonova et al., 2014). To date, about 86 *Sr* genes have been identified, of which 26 stem rust resistance genes have been transferred into bread wheat from other cereal species (McIntosh et al., 2013). For example, *T. turgidum* was the source of the stem rust resistance genes *Sr2*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr11*, *Sr12*, *Sr13*, *Sr14*, and *Sr17*, of which the *Sr2*, *Sr13*, and *Sr14* genes are effective against *Ug99* race; *T. monococcum* was the source of *Sr21*, *Sr22*, and *Sr35* genes (Singh et al., 2011).

Genes that caused the resistance to stem rust have been introduced into wheat gene pool from the genome of various *Aegilops* L. species: *Ae. speltoides* – *Sr32*, *Sr39*, *Sr47*; *Ae. comosa* – *Sr34*; *Ae. ventricosa* – *Sr38* (Schneider et al., 2008). *Ae. tauschii* contributed genes *Sr33*, *Sr45*, *Sr46* (Kerber, Dyck, 1979). Direct hybridization of *T. aestivum* with *Ae. tauschii* and following backcrosses allowed introduction of new resistance genes *SrTA1662*, *SrTA1017*, and *SrTA10187* effective against *Ug99* race (Olson et al., 2013). The search of new resistance genes in wild wheat relatives continues, for example, G. Yu et al. (2017) identified two new *Sr* genes in *Ae. sharonesis*.

One of the objectives of Kazakh-Siberian Spring Wheat Improvement Network (KASIB) is expanding of the genetic polymorphism of new varieties, including resistance to harmful diseases (Gomez-Becerra et al., 2006). This is based on shuttle breeding with CIMMYT (Mexico). Varieties and breeding lines developed through shuttle breeding with participation of *Ae. tauschii* and *T. dicoccum*, as well as lines of the "Arsenal" collection, which have wild species in their pedigree are of interest for breeding for resistance to stem rust in the region.

The aim of the research was analysis of the racial composition of the Western-Siberian stem rust population, resistance assessment of spring bread wheat lines and varieties with

identified resistance genes and identification of the sources with effective *Sr* genes for breeding under Western Siberian conditions.

Material and methods

The racial composition of *Puccinia graminis* f. sp. *tritici* populations collected in 2016–2017 in Omsk region (15 entries of the nursery KASIB-16, Omsk State Agrarian University (SAU)) and Altai region (12 breeding samples, Altay Breeding Center) were analyzed in the Global Rust Reference Center (GRRC, Denmark; http://agro.au.dk/forskning/internationale-platforme/wheatrust).

Selection of single pustule isolates according to requirements of GRRC protocols (www.wheatrust.org) was carried out. Monopustule isolates were reproducted to identify race *Ug99* with usage of the test PCR-Stage 1. A total of 19 single pustule isolates were selected from Omsk population and 20 – from Altai population (Table 1).

Differentiation of stem rust races was performed with use of the set of 20 North American differentiator lines containing *Sr* genes: *Sr5* (ISr5-Ra), *Sr21* (CnS_*Triticum monoc.* Deriv.), *Sr9e* (Vernstein), *Sr7b* (ISr7b-Ra), *Sr11* (ISr11-Ra), *Sr6* (ISr6a-Ra), *Sr8a* (ISr8a-Ra), *Sr9g* (CnSr9g), *Sr36* (W2691SrTt-1), *Sr9b* (W2691Sr9b), *Sr30* (BtSr30Wst), *Sr17+13* (Combination VII), *Sr9a* (ISr9a-Ra), *Sr9d* (ISr9d-Ra), *Sr10* (W2691Sr10), *SrTmp* (CnsSrTmp), *Sr24* (LcSr24Ag), *Sr31* (Benno Sr31/6*LMPG), *Sr38* (VPM-1), *SrMcN* (McNair 701). Infected plants were evaluated in 14–16 days after inoculation according to modified E.C. Stakman scale (Roelfs, Martens, 1988). Virulence phenotypes were classified according to North American system (Jin et al., 2008).

The varieties and lines of bread wheat from three germplasm sets were evaluated in Omsk at least 4–5 times for reaction to stem rust on scales recommended by Koyshibaev et al. (2014). The type of reaction on E.B. Mains and H.S. Jackson scale (1926) and severity – on modified Peterson scale (Peterson et al.,1948) were considered: 0 – immunity, uredopustules not formed; R (Resistance – high resistance), 1 score, severity 5–10 %; MR (Moderately resistant – average resistance), 2 score, severity 10–25 %; M (heterogeneous type), pustules of different sizes, surrounded by chlorotic and necrotic spots or without them; MS (Moderately susceptible – average susceptibility), 3 score, severity 40–50 %; S (Susceptible – susceptibility), 4 score, severity more than 60 %.

In 2016–2018, International Stem Rust Trap Nursery with 58 genotypes with identified *Sr* genes was evaluated to Omsk stem rust population (Table 2). Varieties and lines of nursery-trap were sown manually in 100 cm-long rows with stem rust resistant (Element 22) and susceptible checks (Chernyava 13) alternating every entries.

In 2015, 9 spring wheat lines originating from wide crosses "Arsenal" collection were kindly provided by I.F. Lapochkina for evaluation in Omsk. These lines carry a pyramid of stem rust resistance genes (Lapochkina et al., 2017) – No. 1 [BC_1F_2 (96 × 113) × 145 × 113]; No. 13, 14a [F_3 (96 × 113) × 145]; No. 16, 17, 17a [BC_1F_4 (96 × 113) × 113]; No. 19 [BC_2F_3 (96 × 113) × 113]; No. 20, 22a [F_3 (96 × 113) × 145]. The lines were studied in 2016–2018 in un-replicated trial with the plot size of 2 m^2.

Nine spring wheat varieties and breeding lines from advanced yield trial at Omsk SAU developed through utilization of synthetic wheat with the *Ae. tauschii* genome (Lutescens 24-12 (Kasibovskaya), Lutescens 27-12, Lutescens 87-12, Lutescens 70-13, Lutescens 87-13, Lutescens 88-13 (Silantiy), Lutescens 124-13, Lutescens 53-15, Lutescens 128-15) were evaluated for stem rust resistance and other traits in 2016–2018. The plot size was 25 m^2 with four replications. The checks were Pamyati Azieva (early maturing), Duet (medium maturing), and Element 22 (late maturing).

Sr genes of Omsk SAU varieties were identified using molecular markers: Xsts638 – *Sr15*, Xcfa2123 – *Sr22*, Xgwm210 – *Sr23*, Xscs73 – *Sr24*, Xwmc221 – *Sr25*, BE518379 – *Sr26*, Xscm09 – *Sr31*, SCS421 – *Sr34*, Xcfa2170 – *Sr35*, Xstm773-2 – *Sr36*, Ventriup-LN2 – *Sr38*, Lr34plus – *Sr57*, according to established protocol (http://maswheat.ucdavis.edu/protocols/StemRust/index.htm). The

Table 1. Phenotypic composition and virulence of pathotypes of *Puccinia graminis* f. sp. *tritici* in Omsk and Altai regions (2016–2017)

Parameter	Experimental field of Omsk SAU, 2016	Experimental field of Altai Breeding Center, 2017
No. of samples	15	12
No. of single pustule isolates	19	20
No. of pathotypes	13	10
The indexes of pathotypes	RRGTF, TKRPF, RKRSP, RFRSF, THRTP, RHRTP, TKRTF, QHHSF, RCRTF, SHHSF, RCRTP, QFRSF, RFRTF	SFRSF, NFMSF, QKCSF, MPMTC, LHCSF, LFRSF, LKCSF, LKMSF, LTMSF, QHMSF
High frequency of virulence ≥25 %	*Sr5*, *Sr6*, *Sr7b*, *Sr8a*, *Sr9a*, *Sr9b*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr10*, *Sr17*, *Sr21*, *Sr36*, *Sr38*, *SrMcN*, *SrTmp*	*Sr5*, *Sr8a*, *Sr9a*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr10*, *Sr17*, *Sr21*, *Sr36*, *Sr38*, *SrMcN*
Low frequency of virulence 10–25 %	*Sr11*, *Sr24*, *Sr30*	*Sr7b*, *Sr9b*, *Sr11*, *SrTmp*
Avirulence	*Sr31*	*Sr24*, *Sr30*, *Sr31*

Table 2. Results of evaluation of lines and varieties with identified *Sr* genes on resistance/susceptibility to stem rust, experimental field of Omsk SAU, 2016–2018

No.	Variety, line	Genes	Infection response, %/type		
			2016	2017	2018
1	Element 22	Resistant check	R	5MR	5MR
2	Chernyava 13	Susceptible checks	80S	60S	80S
3	Morocco		40S	45S	40S
4	ISr5-Ra CI 14159	*Sr5*	70S	50S	40S
5	Na 101/6*Marquis	*Sr7a*	25S	40MS	30S
6	ISr7b-Ra CI 14165	*Sr7b*	–	50S	30S
7	CI 14167/9*LMPG-6 DK04	*Sr8a*	30MS	5M	25MS
8	Barleta Benvenuto (CI 14196)	*Sr8b*	–	50S	30S
9	ISr9a-Ra CI 14169	*Sr9a*	10MS	65S	40S
10	Prelude*4/2/Marquis*6/Kenya 117A	*Sr9b*	30M	10M	20MS
11	Vernstein PI 442914	*Sr9e*	20S	10M	20MS
12	Chinese Spring*7/Marquis 2B	*Sr9g*	20S	10M	40S
13	W2691Sr10 CI 17388	*Sr10*	10S	40MS	60S
14	Lee/6*LMPG-6 DK37	*Sr11*	10M	5M	20MS
15	Chinese Spring*5/Thatcher 3B	*Sr12*	10M	5M	40S
16	Preude*4/2Marquis*6/Khapstein	*Sr13*	5M	5M	10M
17	W2691*2/Khapstein	*Sr14*	5M	5M	30MS
18	Preiude*2/Norka	*Sr15*	30MS	10M	30MS
19	Thatcher/CS (CI14173)	*Sr16*	20S	5M	30S
20	Plelude/8*Marquis*2/2/Esp518/9	*Sr17*	60S	60S	50S
21	Little Club/Sr18Mq Marquis "A"	*Sr18*	20S	70S	40S
22	94A 236-1 Marquis "B"	*Sr19*	5MR	20MS	10M
23	94A 237-1 Marquis "C"	*Sr20*	40S	30S	5MS
24	T. monococcum/8*LMPG-6 DK13	*Sr21*	10M	10MR	20M
25	Einkorn		10MR	3MR	10M
26	Mq*6//Stewart*3/RL 5244	*Sr22*	20M	40M	50S
27	Exchange CI 12635	*Sr23*	10MR	5MR	10MR
28	LcSr24Ag + BTSr24Ag	*Sr24*	10MR	10MR	10M
29	Agatha (CI 14048)/9*LMPG-6 DK16	*Sr25*	25MR	15M	10M
30	Eagle Sr26 McIntosh	*Sr26*	15MR	3MR	10M
31	WRT 238-5 (1984) Roelfs	*Sr27*	–	20MS	10M
32	Kota RL471	*Sr28*	60S	15MS	10M
33	Prelude/8*Marquis/2/Etiole de Choisy	*Sr29*	25M	10M	15S
34	Selection from Webster F3:F4#6	*Sr30*	5M	10M	10M
35	Sr31 (Benno)/6*LMPG-6 DK42	*Sr31*	5MR	10MR	10MR
36	Seri 82		R	R	5MR
37	PBW343=Attila with Sr31		5MR	5MR	10MR
38	Cham 10=Kauz//Kauz/star		R	R	5MR
39	Bacanora=Kauz's'		R	R	5MR
40	ER5155 S-203 (1995)Roelfs	*Sr32*	–	10MR	10M
41	RL 5405 (1192) Kerber	*Sr33*	15MR	10MR	30S
42	RL 6098 (1997) Dyck	*Sr34*	–	40MS	50S
43	RL 6099 (1995) Dyck	*Sr35*	20M	40MS	30S
44	W2691SrTt-1 CI 17385	*Sr36*	–	10M	10M
45	Prelude*4/Line W (W3563)	*Sr37*	10M	5M	R
46	Trident Sr38	*Sr38*	5MR	R	R
47	Trident		5MR	R	R
48	RL 5711 Kerber	*Sr39*	10MR	5M	10M
49	RL 6087 Dyck	*Sr40*	5MR	10MR	10M
50	Amigo	*Sr24 + 1RS-Am*	R	R	R
51	Siouxland	*Sr24 + Sr31*	R	R	R
52	Roughrider	*Sr6 + Sr36*	R	5MR	R
53	Sisson	*Sr6 + Sr31 + Sr36*	R	R	R
54	Bt/Wld	*SrWld-1*	15MR	20M	10M
55	Fleming	*Sr6 + Sr24 + Sr36 + 1RS-Am*	10MR	5MR	10MR
56	Chris	*Sr7a + Sr12 + Sr6*	–	10MR	30S
57	CsSSrTmp	*SrTmp*	–	40MS	30S
58	Pavon 76	*Sr2 complex*	R	R	5M

resistance genes of spring bread wheat lines and varieties from nursery-trap and from collection "Arsenal" were identified earlier (McIntosh et al., 2013, 2017; Lapochkina et al., 2017).

In 2016, weather conditions in Omsk region were relatively dry, which contributed to moderate development of stem rust. In 2017, there was an intensive development of the disease, the degree of severity of susceptible accessions varied within 20S–80S. In 2018 high severity of stem rust was observed as the growing season was characterized by cool weather and more precipitation. The degree of severity of susceptible accessions was 30S–80S.

Results

The race composition analysis of stem rust populations identified a significant number of pathotypes: in the Omsk population – 13 and in Altai population – 10 (see Table 1). Unlike many regions of the world where stem rust is a harmful disease for decades, for example in Krasnodar region of Russia (Ablova et al., 2016), for Western Siberia this is surprising result considering a short period of time since its appearance. Most of the identified pathotypes of stem rust population in Omsk and Altai regions were not identical in virulence to the pathotypes, which were found in recent years in Asia and Africa (http://wheatrust.org/fileadmin/www). In all studied Western-Siberian populations of *P. graminis Ug99* and Sicilian races were not identified. Genes of stem rust pathotypes of Omsk population were avirulent only to *Sr31* gene, while Altai pathotypes were avirulent to *Sr31*, *Sr24*, and *Sr30*.

Low frequency of virulence (10–25 %) of Omsk population pathotypes was established for *Sr11*, *Sr24*, *Sr30* genes, for Altai population – for *Sr7b*, *Sr9b*, *Sr11*, *SrTmp* genes, which were ineffective in Omsk region. The results of laboratory evaluation of virulence of *P. graminis* pathotypes collected in Omsk region were confirmed by field of trap nursery with identified *Sr* genes (see Table 2).

Genotypes with *Sr31*: Sr31(Benno)/6*LMPG-6 DK42, Seri 82, PBW343=Attila with Sr31, Cham 10=Kauz//Kauz/star, Bacanora=Kauz's' showed high level of resistance to Omsk stem rust population in all years of study (2016–2018). Line 28 LcSr24Ag + BTSr24Ag with *Sr24* gene was characterized by moderate resistance. For some *Sr* genes, resistant type of reaction under epiphytotic conditions was observed on the stage of adult plants, and susceptible type – on the seedling stage in the laboratory conditions.

For example, variety Trident (entries 46 and 47) with *Sr38* gene had high resistance (R–5MR) in the field; variety Einkorn (entry 25) with *Sr21* gene, and line W2691SrTt-1 CI 17385 (entry 44) with *Sr36* gene had moderate resistance (10M) in the field conditions. In the laboratory conditions the seedlings plants with above mentioned genes were classified as susceptible. Genotypes of ISRTN nursery with a gene pyramid had high resistance to stem rust in all years of research: entry 50 Amigo (*Sr24+1RS-Am*), entry 51 Siouxland (*Sr24+Sr31*), entry 52 Roughrider (*Sr6+Sr36*), entry 53 Sisson (*Sr6+Sr31+Sr36*), entry 55 Fleming (*Sr6+Sr24+Sr36+ 1RS-Am*). The results of stem rust resistance evaluation of "Arsenal" collection and Omsk SAU germplasm are presented in Table 3.

Lines from "Arsenal" collection are of great interest as sources of resistance to pathogen since they possess the gene pyramid: *Sr2* (*T. turgidum*), *Sr36*, *Sr40* (*T. timopheevii*), *Sr44* (*Th. intermediate*). The pedigree of selected lines contains spring wheat line 13/00/i-4 with 7 resistance genes:

Table 3. Results of the assessment for resistance to stem rust of lines and the best varieties of spring bread wheat of Competitive Variety Trial, experimental field of Omsk SAU, 2016–2018

Variety, line	%/type			Resistance genes
	2016	2017	2018	
Pamyati Azieva, susceptible standard	80S	40S	70S	–
Element 22, resistant standard	R	5MR	5MR	*Sr31*, *Sr35*
Lines from "Arsenal" collection				
No. 1 BC_1F_2 (96×113)×145×113	R	R	10MR	*Sr2*, *Sr36*, *Sr44*
No. 14a F_3 (96×113)×145	R	R	R	*Sr36*, *Sr44*
No. 19 BC_2F_3 (96×113)×113	R	R	–	*Sr2*, *Sr36*, *Sr44*
No. 20 F_3 (96×113)×145	R	R	5MR	*Sr2*, *Sr36*, *Sr40*, *Sr44*
Omsk SAU germplasm				
Lutescens 27-12	R	R	25MR	*Sr23*, *Sr36*
Lutescens 87-12	R	R	40M	*Sr23*, *Sr36*
Lutescens 70-13	5MR	R	5MR	*Sr23*, *Sr31*, *Sr36*
Lutescens 87-13	5M	5MR	10MR	*Sr23*, *Sr31*, *Sr36*
Lutescens 88-13	5MR	R	25MR	*Sr23*

Sr2, *Sr36*, *Sr39*, *Sr40*, *Sr44*, *Sr47*, *Sr15*, and winter line GT 96/90 with genes *Sr15*, *Sr24*, *Sr31*, *Sr36*, *Sr40*, *Sr47* (Lapochkina et al., 2017).

In Omsk SAU varieties 3 resistance genes were identified: *Sr23*, *Sr31*, *Sr36*. Variety Element 22, which has winter wheat Aurora in its pedigree also possesses wheat-rye translocation 1BL.1RS with *Sr31* gene (Shamanin et al., 2016b). The combination of effective resistance genes *Sr31* and *Sr35* in this variety results a high level of resistance to stem rust. Element 22 is one of the few varieties with combined resistance to stem and leaf rust. It was included into State register of breeding achievements in Western Siberian region. This variety is the check of the late maturity group at the State Variety Trials in Omsk region.

Stem rust resistant breeding lines Lutescens 27-12, Lutescens 70-13, Lutescens 87-13, Lutescens 88-13 were selected from a cross Lutescens 30-94*2/3/*T. dicoccon* PI 94625/*Ae. squarrosa* (372)//3*Pastor involving Kazakhstan spring wheat line Lutescens 30-94 and CIMMYT line developed by hybridization of synthetic wheat with variety Pastor. The line Lutescens 87-12 originated from a cross Kazakhstanskaya 25/2*Attila/3/*T. dicoccon* PI 94625/*Ae. squarrosa* also involving synthetic wheat. Omsk SAU germplasm possessed different combinations of genes *Sr23*, *Sr31*, and *Sr36*.

Discussion

In modern conditions, stem rust is the most dangerous disease for grain production in Western Siberia. In the epiphytotic years the grain losses of wheat in the region were about 2 million tons. Unfortunately, stem rust resistant varieties included into the State register occupy about 10–15 % of the total wheat sowing area in the region. In 2015–2016, evaluation of spring wheat varieties at Moskalenskiy State Variety Trial of Omsk region (southern forest-steppe zone) demonstrated that out of 57 varieties tested only Element 22 (*Sr31*+*Sr35*), Omskaya 37, Sigma, Uralosibirskaya (*Sr31*), and Sigma 2 (*Sr31*+*Sr25*) were resistant to stem rust (5–15MR). The other varieties were affected by pathogen in medium and high degree requiring the use of chemical protection (Lapochkina et al., 2017). Previously, Shamanin et al. (2016b) identified the stem rust resistance genes in the germplasm developed by breeding institutions of Western Siberia. High frequency of genes *Sr25*, *Sr31*, and their combination was observed. High variability of the race composition of the pathogen population, as shown in our studies, and the uniformity of resistance genes to stem rust in cultivated varieties, threaten grain production stability in Western Siberia.

The breeding strategy should focus on limiting disease development in the region. The study of the populations of *P. graminis*, formed on wheat in the different regions, is very essential to guide the breeding efforts. There were no clones avirulent to *Sr24* gene in Omsk population of *P. graminis* while in Altai region there were no clones virulent to *Sr24*, which remains its effectiveness in Novosibirsk region (Skolotneva et al., 2018). The results of the population composition comparison suggest that Omsk and Altai subpopulations have relatively independent sources of genetic diversity and the contact zone. Western Siberian population of *P. graminis* has quite complex structure. Two subpopulations are assumed to exist: Omsk and Altai – with independent sources of genetic diversity, and zone of genotypic exchange on wheat crop in Novosibirsk region (Skolotneva et al., 2020).

Omsk stem rust population analysis showed that the spectrum of effective resistance genes has narrowed due to losses of some genes to the local population of *P. graminis*.

Highly resistant varieties and lines of ISRTN nursery were identified: Sr31 (Benno)/6*LMPG-6 DK42, Seri 82, Cham 10, Bacanora (*Sr31*), RL 6087 Dyck (*Sr40*), Amigo (*Sr24*, *1RS-Am*), Siouxland (*Sr24*, *Sr31*), Roughrider (*Sr6*, *Sr36*), Sisson (*Sr6*, *Sr31*, *Sr36*), Fleming (*Sr6*, *Sr24*, *Sr36*, *1RS-Am*), Pavon 76 (*Sr2 complex*). Selected varieties and lines are recommended for using as sources of resistance in breeding programs to create resistant wheat varieties to stem rust. Effective resistance genes *Sr31*, *Sr40*, *Sr2 complex*, and their combinations with ineffective genes are recommended for use in breeding, taking into account the constant rotation, combination of genes of nonspecific resistance, as well as the possibility of infection threat from neighboring territory.

The resistance gene *Sr2*, widely used in breeding for resistance to virulent stem rust races, is common in commercial varieties in a number of countries around the world, particularly in the United States, Australia, India, and Mexico. This gene is practically absent in the commercial varieties of Russian Federation, however, for effective protection against stem rust, its pyramiding with other resistance genes is recommended (Baranova et al., 2015).

For the development of varieties with long-term resistance, the strategy of combining genes responsible for different types of resistance in one genotype is used. Pyramiding of specific resistance genes (*Sr11*, *Sr24*, *Sr30*, and *Sr31*) with APR gene *Sr2*, which causes the slow development of the disease (slow rusting), will provide longer protection of wheat crops from stem rust in Western Siberia in the present phytosanitary situation.

In this regard, the lines from "Arsenal" collection – No. 1 BC_1F_2 (96 × 113) × 145 × 113 (*Sr2*, *Sr36*, *Sr44*); No. 14a F_3 (96 × 113) × 145 (*Sr36*, *Sr44*); No. 19 BC_2F_3 (96 × 113) × 113 (*Sr2*, *Sr36*, *Sr44*); No. 20 F_3 (96 × 113) × 145 (*Sr2*, *Sr36*, *Sr40*, *Sr44*) represent a promising starting material for breeding and creation of varieties with long-term resistance.

It is justified to include resistance sources to stem rust with minimum number of negative traits that reduce their breeding value. In this regard, stem rust resistant germplasm from Omsk SAU with identified effective genes Element 22 (*Sr31*, *Sr35*), Lutescence 27-12, Lutescence 87-12 (*Sr23*, *Sr36*), Lutescence 70-13, Lutescence 87-13 (*Sr23*, *Sr31*, *Sr36*), Lutescence 88-13 (*Sr23*) are valuable starting material for breeding in the region.

Conclusion

Thus, the genetic similarity of spring wheat varieties on stem rust resistance genes cultivated over large areas in Western Siberia, and the predominance of varieties with race specific

resistance genes contribute to spreading and high variability of the pathogen. The lines from collection "Arsenal" – No. 1 BC_1F_2 (96 × 113) × 145 × 113 (*Sr2*, *Sr36*, *Sr44*), No. 14a F_3 (96 × 113) × 145 (*Sr36*, *Sr44*), No. 19 BC_2F_3 (96 × 113) × 113 (*Sr2*, *Sr36*, *Sr44*), No. 20 F_3 (96 × 113) × 145 (*Sr2*, *Sr36*, *Sr40*, *Sr44*), varieties of Omsk Agrarian University – Element 22 (*Sr31*, *Sr35*), Lutescens 27-12, Lutescens 87-12 (*Sr23*, *Sr36*), Lutescens 70-13, Lutescens 87-13 (*Sr23*, *Sr31*, *Sr36*) are recommended for inclusion into breeding process of the creation of resistant to stem rust varieties in the region. Further monitoring of the virulence of stem rust pathogen and coordination strategy of breeding programs in Western Siberia, and neighboring regions of the Kazakhstan Republic is recommended. Incorporation of effective resistance genes, in particular *Sr2* and *Sr40*, will improve the phytosanitary situation and expand the segment of resistant varieties in the region.

References

Ablova I.B., Bespalova L.A., Kolesnikov F.A., Nabokov G.D., Kovtunenko V.Ya., Filobok V.A., Davoyan R.O., Khudokormova Zh.N., Mokhova L.M., Levchenko Yu.G., Tarkhov A.S. Principles and methods of wheat breeding on tolerance to diseases in KRIA named after P.P. Lukiyanenko. Zernovoe Khozjaistvo Rossii = Grain Economy of Russia. 2016;5:32-36. (in Russian)

Baranova O.A., Lapochkina I.F., Anisimova A.V., Gajnullin N.R., Iordanskaya I.V., Makarova I.Yu. Identification of *Sr* genes in new common wheat sources of resistance to stem rust race *Ug99* using molecular markers. Russ. J. Genet.: Appl. Res. 2016;6(3): 344-350.

Gomez-Becerra H., Morgounov A., Abugalieva A. Evaluation of yield grain stability, reliability and cultivar recommendation in spring wheat (*Triticum aestivum*) from Kazakhstan and Siberia. Central Eur. J. Agriculture. 2006;6:649-660.

Jin Y., Szabo L.J., Pretorius Z.A., Singh R.P., Ward R., Fetch T., Jr. Detection of virulence to resistance gene *Sr24* within race TTKS of *Puccinia graminis* f. sp. *tritici*. Plant Dis. 2008;92: 923-926.

Kerber E.R., Dyck P.L. Resistance to stem and leaf rust in *Aegilops squarrosa* and transfer of a gene for stem rust resistance to hexaploid wheat. In: Ramanujam S. (Ed.). Proceeding of the 5th International Wheat Genetics Symposium. New Delhi: Ind. Soc. of Genetics and Plant Breeding, Ind. Agric. Res. Institute, 1979;358-364.

Koyshybaev M., Shamanin V.P., Morgunov A.I. Screening of Wheat for Resistance to Major Diseases. Ankara: FAO-SEK, 2014. (in Russian)

Lapochkina I.F., Baranova O.A., Shamanin V.P., Volkova G.V., Gainullin N.R., Anisimova A.V., Galinger D.N., Lazareva E.N., Gladkova E.V., Vaganova O.F. The development of the initial material of spring common wheat for breeding for resistance to stem rust (*Puccinia graminis* Pers. f. sp. *tritici*), including the Ug99 race, in Russia. Russ. J. Genet.: Appl. Res. 2017;7(3): 308-317. DOI 10.1134/S207905971703008X.

Leonova I.N., Orlovskaya O.A., Röder M.S., Nesterov M.A., Budashkina E.B. Molecular diversity of common wheat introgression lines (*T. aestivum*/*T. timopheevii*). Russ. J. Genet.: Appl. Res. 2015; 5(3):191-197. DOI 10.1134/S2079059715030090.

Mains E.B., Jackson H.S. Physiologic specialization in the leaf rust of wheat, *Puccinia triticina* Erikss. Phytopathology. 1926;16(2): 89-120.

McIntosh R.A., Dubcovsky J., Rogers W.J., Morris C., Xia X.C. Catalogue of gene symbols for wheat: 2017 Supplement. Available at: http://shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement 2017.pdf

McIntosh R.A., Yamayaki Y., Dubcovsky J., Rogers W.J., Morris C., Appels R., Xia X. Catalogue of Gene Symbols for Wheat: 2013–2014 Supplement. Available at: http://www.shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement2013.pdf

Olson E.L., Rous M.N., Pumphrey M.O., Bowden R.L., Gill B.S., Poland J.A. Introgression of stem rust resistance genes *SrTA10187* and *SrTA10171* from *Aegilops tauschii* to wheat. Theor. Appl. Genet. 2013;126:2477-2484. DOI 10.1007/s00122-013-2148-z.

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. (Sect. C). 1948;26:496-500.

Roelfs A.P., Martens J.W. An international system of nomenclature for *Puccinia graminis* f. sp. *tritici*. Phytopathology. 1988;78: 526-533.

Sanin S.S. Problems of phytosanitary of grain production. In: Protection of Cereal Crops Against Diseases, Pests, and Weeds: Progress and Problems: Proc. of the Int. sci. and pract. conf. Bolshye Vyazemy, 5–9 Dec. 2016. Moscow, 2016;4-15. (in Russian)

Schneider A., Molnár I., Molnár-Láng M. Utilisation of *Aegilops* (goatgrass) species to widen the genetic diversity of cultivated wheat. Euphytica. 2008;163:1-19. DOI 10.1007/s10681-007-9624-y.

Shamanin V.P., Morgunov A.I., Petukhovskiy S.L., Likhenko I.E., Levshunov M.A., Salina E.A., Pototskaya I.V., Trushchenko A. Ju. Breeding of spring bread wheat for resistance to stem rust in West Siberia. Omsk: FGOU VPO OmGAU Publ., 2015. (in Russian)

Shamanin V.P., Pototskaya I.V., Klevakina M.V. Assessment of the Siberian collection of spring wheat on resistance to stem rust in the southern forest-steppe of Western Siberia. Vestnik Kazanskogo GAU = The Herald of Kazan State Agrarian University. 2016a;2(40):55-59. (in Russian)

Shamanin V., Salina E., Wanyera R., Zelenskiy Yu., Morgounov A. Genetic diversity of spring wheat from Kazakhstan and Russia for resistance to stem rust Ug99. Euphytica. 2016b;12:287-296. DOI 10.1007/s10681-016-1769-0.

Singh R.P., Hodson D.P., Huerta-Espino J., Jin Y., Njau P., Wanyera R., Herrera-Foessel S.A., Ward R.W. Will stem rust destroy the world's wheat crop? Adv. Agron. 2008;98:271-309. DOI 10.1016/S0065-2113(08)00205-8.

Singh R.P., Hodson D.P., Jin Y., Lagudah E.S., Ayliffe M.A., Bhavani S., Rouse M.N., Pretorius Z.A., Szabo L.J., Huerta-Espino J., Basnet B.R., Lan C., Hovmøller M.S. Emergence and spread of new races of wheat stem rust fungus: Continued threat to food security and prospects of genetic control. Phytopathology. 2015;105:872-884.

Singh R.P., Huerta-Espino J., Bhavani S., Herrera-Foessel S.A., Singh D., Singh P.K., Velu G., Mason R.E., Jin Y., Njau P., Crossa J. Race non-specific resistance to rust diseases in CIMMYT spring wheats: breeding and advances. Euphytica. 2011;179: 175-186. DOI 10.1007/s10681-010-0322-9.

Skolotneva E.S., Bukatich E.Ju., Bojko N.I., Piskarev V.V., Salina E.A. Screening of an international stem rust nursery trap for *Ug99* in the Priobie forest-steppe in 2017. In: Gene Pool and Plant Breeding: Proc. IV Int. sci. and pract. conf., 4–6 April 2018. Novosibirsk, 2018;313-318. (in Russian)

Skolotneva E.S., Kel'bin V.N., Morgunov A.I., Bojko N.I., Shamanin V.P., Salina E.A. Races composition of the Novosibirsk population of *Puccinia graminis* f. sp. *tritici*. Mikologiya i Fitopatologiya = Mycology and Phytopathology. 2020;54(1): 49-58. (in Russian)

Yu G., Champouret N., Steuernagel B., Olivera P.D., Simmons J., Williams C., Johnson R., Moscou M.J., Hernández-Pinzón I., Green P., Sela H., Millet E., Jones J.D.G., Ward E.R., Steffenson B.J., Wulff B.B.H. Discovery and characterization of two new stem rust resistance genes in *Aegilops sharonensis*. Theor. Appl. Genet. 2017;130:1207-1222.

ORCID ID

V.P. Shamanin orcid.org/0000-0003-4767-9957
I.V. Pototskaya orcid.org/0000-0003-3574-2875
S.S. Shepelev orcid.org/0000-0002-4282-8725
V.E. Pozherukova orcid.org/0000-0001-8429-2167
E.A. Salina orcid.org/0000-0001-8590-847X
E.S. Skolotneva orcid.org/0000-0001-8047-5695
M. Hovmøller orcid.org/0000-0002-4432-8898
A.I. Morgounov orcid.org/0000-0001-7082-5655

Acknowledgements. The present research has been carried out with the financial support of the Russian Science Foundation (project No. 16-16-10005) and Russian Foundation for Basic Research (project No. 17-29-08018).

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

Received March 24, 2019. Revised December 07, 2019. Accepted December 07, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):139-148
DOI 10.18699/VJ20.609

Septoria blotch epidemic process on spring wheat varieties

E.Yu. Toropova[1, 2] ✉, O.A. Kazakova[1, 2], V.V. Piskarev[3]

[1] Novosibirsk State Agrarian University, Novosibirsk, Russia
[2] All-Russian Research Institute of Phytopathology, r. p. Bolshie Vyazemy, Odintsovo district, Moscow region, Russia
[3] Siberian Research Institute of Plant Production and Breeding – Branch of the Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ e-mail: 89139148962@yandex.ru

Abstract. The *Septoria* blotch of spring wheat leaves and ears is one of the most economically significant infections in the Siberian region. In the control systems of *Septoria* blotch the main ecologically safe element is resistant varieties, which are designed to slow down the pathogens reproduction rate and slow down or stop the development of the epiphytotic process. The purpose of the work was to clarify the species composition of *Septoria* blotch pathogens for West Siberian regions and spring wheat varieties, to study the epiphytotic process of *Septoria* differentially on the leaves and ears of varieties, and to evaluate the activity of seed transmission of *Parastagonospora nodorum*. Studies were carried out in 2016–2018 according to generally accepted methods. *Septoria* leaf and ear blotch of spring wheat is widespread in West Siberia and the Trans-Urals, causing a decrease in yield by up to 50 % or more with the deterioration in grain quality. The causative agents of the disease are *P. nodorum*, *Septoria tritici*, and *P. avenae* f. sp. *triticae*, and the species ratio varied across the regions and varieties, and within plant organs. In Novosibirsk Region, *P. nodorum* completely dominated; *S. tritici* was 13.8 times less common; and *P. avenae* f. sp. *triticae* was a singleton. In Tyumen Region, the dominance of *P. nodorum* was disrupted in some geographic locations by *S. tritici* and *P. avenae* f. sp. *triticae*. In Altai Krai, *P. nodorum* predominated at all points studied; *S. tritici* and *P. avenae* f. sp. *triticae* were found everywhere, but 5.6 and 8.6 times less often, respectively. The study of spring wheat varieties of different origins has not revealed any samples immune to *Septoria* blotch. A differentiated manifestation of resistance to *Septoria* leaf and ear disease has been established. Some varieties show complex resistance, combining reduced susceptibility to *Septoria* leaf and ear disease. Seed infection with *P. nodorum* in the regions of Siberia reached 7 thresholds and was largely (52.5 %) determined by the August weather conditions. The study of the collection of spring wheat varieties from three Siberian regions has revealed the following trend. Transmission of *P. nodorum* with the seeds of varieties was the most active (7.6 %) in Novosibirsk Region and somewhat weaker in Omsk Region (5.7 %). The most favorable phytosanitary situation was in Kurgan Region, where varieties transmitted *P. nodorum* to a low degree (2.1 %), below the threshold.
Key words: *Septoria* leaf and ear blotch; spring wheat; monitoring; *Parastagonospora nodorum*; *Septoria tritici*; *P. avenae* f. sp. *triticae*; variety; resistance; seed transmission.

For citation: Toropova E.Yu., Kazakova O.A., Piskarev V.V. *Septoria* blotch epidemic process on spring wheat varieties. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):139-148. DOI 10.18699/VJ20.609

Эпифитотический процесс септориоза на сортах яровой пшеницы

Е.Ю. Торопова[1, 2] ✉, О.А. Казакова[1, 2], В.В. Пискарев[3]

[1] Новосибирский государственный аграрный университет, Новосибирск, Россия
[2] Всероссийский научно-исследовательский институт фитопатологии, р.п. Большие Вяземы, Одинцовский район, Московская область, Россия
[3] Сибирский научно-исследовательский институт растениеводства и селекции – филиал Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук, р. п. Краснообск, Новосибирская область, Россия
✉ e-mail: 89139148962@yandex.ru

Аннотация. Септориоз листьев и колоса яровой пшеницы – одна из наиболее экономически значимых инфекций в Сибирском регионе. В системах контроля септориоза основным экологически безопасным элементом являются устойчивые сорта, которые тормозят или останавливают развитие эпифитотического процесса путем замедления размножения возбудителей септориоза. Цель работы состояла в уточнении видового состава возбудителей септориоза по регионам Западной Сибири и сортам яровой пшеницы, исследовании эпифитотического процесса септориоза дифференцированно на листьях и колосьях сортов, а также в оценке активности семенной передачи *Parastagonospora nodorum*. Исследования проводили в 2016–2018 гг. по общепринятым методикам. В Западной Сибири септориоз листьев и колоса яровой пшеницы широко распространен: 35 % по показателю развития болезни и 90 % по распространенности. Видовой состав

возбудителей септориоза представлен *P. nodorum*, *Septoria tritici* и *P. avenae* f. sp. *triticae*, причем соотношение видов изменялось по регионам. В Новосибирской области отмечено полное доминирование *P. nodorum*, *S. tritici* встречался в 13.8 раза реже, а *P. avenae* f. sp. *triticae* единично. В Тюменской области доминирование *P. nodorum* не было таким безусловным и нарушалось в некоторых географических пунктах *S. tritici* и *P. avenae* f. sp. *triticae*. В Алтайском крае во всех точках учета преобладал *P. nodorum*; *S. tritici* и *P. avenae* f. sp. *triticae* встречались повсеместно, но в 5.6 и 8.6 раза реже соответственно. Изучение сортов яровой пшеницы разного происхождения не позволило выявить иммунных к септориозу листьев и колоса форм. Установлено дифференцированное проявление признаков устойчивости к септориозу листьев и колоса, выделены сорта, сочетающие пониженную восприимчивость к септориозу и листьев, и колоса. Инфицированность семян фитопатогеном в регионах Сибири достигала 7 экономических порогов вредоносности и в значительной степени (52.5 %) определялась погодными условиями августа. Изучение коллекций сортов из трех регионов Сибири позволило выявить следующую тенденцию: наиболее активно передача *P. nodorum* с семенами сортов шла в Новосибирской области (7.6 %), в меньшей степени – в Омской (5.7 %); самая благополучная фитосанитарная ситуация была в Курганской области: сорта передавали *P. nodorum* в слабой степени (2.1 %), ниже экологического порога вредоносности.

Ключевые слова: септориоз листьев и колоса; яровая пшеница; мониторинг; *Parastagonospora nodorum*; *Septoria tritici*; *P. avenae* f. sp. *triticae*; сорт; устойчивость; семенная передача.

Introduction

Septoria blotch of leaves and ears has long been one of the most common and damaging diseases of spring wheat in all areas of its cultivation (Eyal, 1999; Robert et al., 2004; Nazarova et al., 2010). When wheat is affected by *Septoria* blotch, the leaves dry out prematurely, and grain is poured only at the expense of the stem and spike green parts. Grain is formed hollow with a low grain unit and 1000-grain mass. The grain productivity of spring wheat falls by 25–60 %. The germination ability and germination energy of seeds are reduced by 7–12 % (Chulkina, 1991; Parker et al., 2004; Robert et al., 2004; Sanin et al., 2015).

The main causative agents of the disease in spring wheat are the fungi *Parastagonospora nodorum* (Berk.) Quaedvl., Verkley & Crous (syn. *Depazea nodorum* Berk.) and *Septoria tritici* Roberge ex Desm. (syn. *Zymoseptoria tritici* (Roberge ex Desm.) Quaedvl. & Crous.). Of the two species of fungi in spring wheat in Siberia, *P. nodorum* is predominant, characterized by faster (8–10 times) germination of pycniospores and colonization of the host plant tissue compared to *S. tritici* (Chulkina, 1991). Also, *Phaeosphaeria avenaria* f. sp. *triticae* Shoem. & C.E. Babc. (syn. *S. avenae* f. sp. *triticea*) is detected in spring wheat in Siberia (Toropova et al., 2019). The same species cause most damage to winter wheat crops (Kolomiets et al., 2018). Each of the plant pathogens has certain epidemiological features and requirements for environmental conditions, which ensures a greater ecological plasticity of the disease and difficulty in its control. Both species can exist together on the same plant. *S. tritici* mainly affects the leaves, developing more intensively on young than on old tissues. *P. nodorum* equally well affects both leaves and ears, and is able to live and multiply on dead tissues (Kolomiets et al., 2018). The highest susceptibility to *P. nodorum* was noted in the phase of heading and flowering (Cooke, Jones, 1970). The phase of the highest sensitivity of wheat to *S. tritici* occurs during the tillering–bumping period, which is mainly associated with increased air humidity in the crop folia (Adolf et al., 1993). The representation of the main pathogens in the pathogenic complex of winter wheat *Septoria* blotch is dependent on the weather. *S. tritici* predominates in years with low winter temperatures and warmer and rainy conditions in the first half of summer. *S. nodorum* prevails in years with wetter autumn, warm winters, and high rainfall in the second half of summer (Sanin et al., 2017).

Septoria leaf and ear blotch is recorded in more than fifty countries, mainly at latitudes with a temperate climate (Europe, North America, and Australia). In the territory of the former Soviet Union, *Septoria* is especially prevalent in the North Caucasus, the Urals, the Ukraine, and Belarus (Nazarova et al., 2010).

The germ tubes of plant pathogens are introduced into the leaf tissue of susceptible hosts after the germination of pycniospores in most cases through the stomatal cleft, less often through the epidermis. Several sprouts often penetrate one stoma. After the penetration of pathogens, weak branching and growth of fungi in the intercellular spaces of leaf mesophyll cells is noted, with the majority of hyphae growing along the leaf between the epidermal and mesophyll cells. This, apparently, explains why *Septoria* blotch spots are often elongated along the leaf veins. Branching of the hyphae gradually occurs, as well as their growth in various directions; the leaf thickness is penetrated several times; and the intercellular spaces are filled with mycelium (Robert et al., 2004; Nazarova et al., 2010). The toxin ochracin, which suppresses the growth of host plants, and septorin, which inhibits oxidative phosphorylation in plant cells, were detected in *S. tritici*, the *Septoria* leaf blotch pathogen (Eyal, 1999).

Plant infection with *Septoria* blotch is particularly successful if the period of drip wetting at the optimum temperature is at least 8 hours and the relative humidity is 98–100 %. Therefore, *Septoria* blotch most often develops in areas with sufficient moisture. However, there are cases when *Septoria* blotch is dangerous in dry areas. This is because pathogens can use an intermittent wet period, as a result of regular dewfall (Toropova et al., 2002).

The incubation period of *Septoria* blotch, depending on hydrothermal conditions, is from 6 to 49 days. The regression analysis showed that 45 % of the variability of the latent period of the pathogen is due to the influence of temperature, 12 % due to its population density, and only 3 % to the duration of hydration (Chulkina, 1991). This indicates that, once in the ecological niche, the pathogen is almost independent of

moisture. However, in the external environment throughout all the phases of the transmission mechanism (separation from the causative agent source, transmission of propagules with airborne droplets, and germination and incorporation into the tissues of susceptible plants), its life cycle largely depends on the presence of droplet-liquid moisture (Nolan et al., 1999; Toropova et al., 2002; Nazarova et al., 2010; Pakholkova, 2015).

The reproductive potential of *Septoria* blotch pathogens is quite high, amounting to 10–15 thousand spores in one pycnium. A close correlation was established between the numbers of pycnia and spores in them ($r = 0.901$) (Chulkina, 1991).

For the pycnia formation, high relative humidity (over 98 %) is required. The formation of mature *P. nodorum* pycnia takes 8–14 days, and *S. tritici* 14–20 days after inoculation (Pakholkova, 2015). At the end of the growing season of host plants, the pycnia number reaches its maximum value. From 6 to 12 fungi generations develop during the season. At the end of the growing season, pycnia and sacs (fruiting bodies) of different species of *Septoria* blotch pathogens are formed on the same leaves (Kolomiets et al., 2018).

The *Septoria* blotch pathogens winter on infected plant debris as mycelium, fruiting bodies, and pycnia; *P. nodorum* also on or inside seeds as mycelium and pycnia (Toropova et al., 2016; Sanin et al., 2018). The pathogens can survive for 6–18 months on infected plant debris in the surface soil layer or on its surface and until the June–July end in the soil at the depth of the arable layer. At the same time, 1 g of plant residues contains 1.5–6 million pathogen spores in the soil and 52–63 million on the soil surface. The viability of both spores in pycnia and especially ascospores in sacs is high, reaching 100 % in the spring wheat earing phase, when mass plant infection occurs (Chulkina, 1991; Toropova et al., 2002; Sanin et al., 2018).

The ability of pycniospores to spread after release from pycnia is associated with rain. Wind without rain cannot take away spores, as they are covered with an adhesive mass, which in the absence of moisture sticks them to the substrate. As you move away from the infection source, the spore population decreases. Further than 500 m away, spores are usually not detected. In the vertical direction, spores in the mass rise to 75 cm and are absent at a height of 150 cm (Eyal, 1999; Toropova et al., 2002; Robert et al., 2004; Sanin et al., 2018).

Ascospores can be released from perithecia only during rain. This process is extended and can last several months. Ascospores are carried over a few (2–3) kilometers.

The *Septoria* blotch development is largely dependent on hydrothermal conditions. At low humidity, spores are not released from the pycnia and do not spread. Therefore, the disease outbreaks occur in years with significant rainfall, with a maximum air temperature of no higher than 30 °C and an average daily temperature between 14 and 21 °C. The damage by root rot even in a slight degree (3–10 %) enhances the *Septoria* blotch severity (Toropova et al., 2002; Toropova, 2005; Nazarova et al., 2010; Sanin et al., 2015).

A decrease in soil cultivation intensity and the accumulation of infected plant residues on the soil surface have led to a 2–2.5-fold increase in the frequency of *Septoria* blotch epiphytotic in the forest-steppe of West Siberia over the past 10 years. Since the seed transmission of the main causative agent *P. nodorum* has intensified in spring and winter wheat in the last decade, this annually created the prerequisites for the formation of early foci of the disease (Toropova et al., 2018). The epiphytotic development of *Septoria* blotch foci, in which the infection on the upper leaves reaches the economic threshold (15–20 %), occurs when 3 times more rain falls over the ten-day period than the long-term average, at a temperature of 14–22 °C; the disease develops at a rate of up to 2–3 % per day, which necessitates the use of fungicides (Toropova, 2005; Sanin et al., 2015).

The seed transmission of *P. nodorum* causes the early appearance of *Septoria* blotch on the coleoptile and basal leaves in the seedling–tillering phase. There is no linear relationship between the infection of seeds and seedlings. The seed infection by 5–10 % can already lead to an epiphytotic of *Septoria* blotch in favorable weather conditions (Chulkina, 1991; Sanin et al., 2015; Toropova et al., 2018).

The *Septoria* blotch pathogens have an r life cycle strategy. The r-strategy characteristic features are: numerous (6–12) generations of conidial sporulation (pycnia with spores) under favorable conditions; high pathogen transmission rate; and a polycyclic, variable type of the epiphytotic process dynamics. The strategy of phytosanitary measures against *Septoria* blotch is to reduce the rate of pathogen propagation and the development of epiphytotic process to a level below the economic threshold. This is achieved due to the genetic and physiological plant resistance and to the prevention of vertical transmission of *P. nodorum* with seeds.

The breeding of resistant varieties is the most promising and environmentally friendly component of systems of integrated protection from *Septoria* leaf and ear blotch of spring wheat. In practice, selection of wheat for resistance to *Septoria* blotch is difficult, because this trait is unstable, varies in time and space, and is controlled by many mechanisms (Kolomiets et al., 2018). Genotypes with complex resistance are rare: varieties can be resistant to one pathogen type and susceptible to another (Jenkins, Jones, 1981). At one time, it was believed that wheat was generally not resistant to *P. nodorum* (Scharen, Krupinsky, 1970; Broennimann, 1975). However, further studies showed that the situation is not so obvious (Mullaney et al., 1981; Du et al., 1999).

Stable progress is observed in the selection of wheat varieties resistant to *S. tritici*. Russian researchers have identified a number of varieties that are recommended for inclusion into the breeding programs as sources and donors of resistance to the pathogen (Kolomiets et al., 2018).

Resistance to *Septoria* blotch can be either quantitative (horizontal) or isolate-specific (vertical) (Tyryshkin, Ershova, 2004; Kolomiets et al., 2017). Currently, 17 genes of resistance to *S. tritici* (*Stb1–Stb17*) have been identified in wheat. Due to genetic analysis in the "wheat–*M. graminicola*" pathosystem, gene-for-gene interaction has been proven (Kolomiets et al., 2017). Recent studies have established some biochemical mechanisms of resistance of common wheat to *Septoria* blotch (Veselova et al., 2018, 2019). Also, *Septoria*-resistant common wheat forms have been isolated abroad (Van Ginkel, Rajaram, 1999; Simón et al., 2003; Robert et al., 2004). One of the aspects that impede the search for plant forms resistant to *Septoria* blotch is the underestimation of the multicomponent

species composition of *Septoria* blotch causative agents and the insufficient knowledge of the regional species representation in the disease pathogenic complex. In addition, when assessing plant resistance, they often do not carry out a differential account of leaf and spike lesions, though they might be determined by different mechanisms.

The purpose of the work was to clarify the species composition of *Septoria* blotch pathogens for regions of West Siberia and spring wheat varieties, to study the epiphytotic process of *Septoria* blotch differentially on the leaves and ears of varieties, and to evaluate the activity of seed transmission of *P. nodorum*.

Materials and methods

The studies were carried out in 2016–2018 in the West Siberian forest-steppe zone. *Septoria* leaf and ear blotch was recorded using the international scale (Chulkina et al., 2017) in the technological conditions of the region's farms. To clarify the species composition of *Septoria* blotch pathogens, we collected samples of infected wheat plants and plant residues in the fields at the end of the wheat growing season, taking 10–20 infected leaves at each point. To determine the species of fungi, fragments of the diseased tissue with fruiting bodies (pycnia) were placed on glass slides in a water drop, and after 10–15 minutes they were viewed at low magnification. The shape and size of spores emerging from the pycnia determined the species and its percentage in the total number of the pycnia studied (Pyzhikova et al., 1988). The study of *Septoria* leaf and ear blotch on the spring wheat varieties and variety specimens was carried out on the natural infection background using a collection from the Institute of Cytology and Genetics of SB RAS. The collection consisted of 10 varieties from 5 Russian regions and 13 foreign samples from 8 countries. The area under each variety (variety specimen) ranged from 3 to 10 m^2 in triplicate.

According to the degree of damage, the varieties were divided into the following groups: 0–5, highly resistant; 6–20 %, resistant; 21–40 %, poorly susceptible; 41–65 %, susceptible; and 66–100 %, highly susceptible (Sanin et al., 2015).

The seed samples for the analysis were taken from the farms of Novosibirsk, Tomsk, and Tyumen Regions as well as Altai Krai and Krasnoyarsk Krai. Seed analysis for *P. nodorum* infection was carried out by an original method (Chulkina et al., 2017). Over the years, 258 seed samples of 53 spring wheat varieties have been analyzed totally.

In the northern forest-steppe of the Novosibirsk Region, 2016 was dry (Hydrothermal coefficient = 0.81), whereas 2017 and 2018 were wet (Hydrothermal coefficient = 1.26 and 1.33, respectively), which significantly influenced the intensity of the natural infection background.

Results and discussion

The *Septoria* blotch monitoring in agrocenoses of winter and spring wheat in Novosibirsk, Tomsk, Kemerovo, Kurgan, and Tyumen Regions and Altai Krai, conducted in 2016–2018, established a widespread distribution of the disease in spring wheat varieties. The disease incidence was from 5 to 35 %, and its severity reached 90 %. By the start of the earing phase, a critical situation aroused in most agrocenoses, which required prompt measures to protect spring wheat from *Septoria* blotch of leaves and ears, despite a significant diversity of weather conditions, varieties, and technologies for spring wheat cultivation (Toropova et al., 2019).

The first single foci of *Septoria* blotch on the lower leaves of spring wheat during transmission of the pathogen from infected plant residues were observed in 2016 and 2017 in the last two ten-day periods of June; in 2018, due to late sowing, in the first two ten-day periods of July. Moreover, *P. nodorum* appeared earlier (June–early July) than *S. tritici* (late July–August).

A comparison of weather conditions over the years contrasting in the incidence and severity of *Septoria* blotch shows that moderate and significant intensity epidemics began when 76.0 to 111.0 mm of precipitation fell at an average air temperature of 16.7 °C. The years favorable to *Septoria* blotch were distinguished by an increase in precipitation during critical periods for plant infection by an average of 6.7 times and a decrease in air temperature by an average of 2.5 °C.

The climatic trend characterized by warming and increased contrast in the weather conditions during the growing season turned out to be favorable to plant pathogens, leading to an increase in the frequency of *Septoria* blotch epidemics in spring wheat distribution regions, including West Siberia (Levitin, 2015; Toropova et al., 2016). The results of our studies are consistent with published data on the increase in the spread of *Septoria* blotch on the winter wheat in the European part of Russia (Sanin et al., 2017; Gultyaeva et al., 2019).

***Septoria* leaf and ear blotch of spring wheat was encountered** as *P. nodorum*, *S. tritici*, and *P. avenae* f. sp. *triticae*, and the ratio of species varied across the regions (Table 1). The table shows that *P. nodorum*, *S. tritici*, and *P. avenae* f. sp. *triticae* pycnia were present on the infected leaves of spring wheat varieties cultivated in Siberia; however, their ratio in the regions varied significantly. Thus, according to the averaged data, at 6 sampling points in Novosibirsk Region, an overwhelming dominance of *P. nodorum* was revealed. *S. tritici* was 13.8 times less common; *P. avenae* f. sp. *triticae* in the *Septoria* blotch pathogenic complex in Novosibirsk Region was a singleton.

A study of the species composition of *Septoria* blotch pathogens in Tyumen Region showed a significant diversity at the sampling points. Two sampling points in Tyumen Region were under the dominance of *P. nodorum*; the second position belonged to *P. avenae* f. sp. *triticae*, not to *S. tritici*. On the wheat leaves from the third point, only *P. nodorum* was detected. The dominance of *S. tritici* was revealed on wheat leaves from the fourth point in Tyumen Region, and at 3 points out of five *P. avenae* f. sp. *triticae* made a significant contribution to the pathogenic complex of *Septoria* blotch in spring wheat, which was not observed in Novosibirsk Region.

P. nodorum dominated at all the sampling points in Altai Krai. *S. tritici* and *P. avenae* f. sp. *triticae* were found everywhere, but 5.6 and 8.6 times less often than the main causative agent of the disease, respectively. *P. avenae* f. sp. *triticae* was found on spring wheat leaves in Altai agrocenoses 11.3 times more often than in Novosibirsk Region; that is, its contribution to the *Septoria* blotch pathogenic complex was much more significant.

Thus, significant differences were found in the species composition of *Septoria* blotch of spring wheat in regions

Table 1. The species composition of *Septoria* blotch pathogens on spring wheat leaves in the Siberian regions, 2016–2018, %

Region	*P. nodorum*	*S. tritici*	*P. avenae* f. sp. *triticae*
Altai Krai	77.2 ± 6.6	13.8 ± 1.9	9.0 ± 1.1
Novosibirsk Region	92.5 ± 8.9	6.7 ± 0.9	0.8 ± 0.1
Tyumen Region	68.0 ± 7.1	24.0 ± 3.2	8.0 ± 1.2

of West Siberia, and this should be taken into account when creating pathogen populations for artificial infection of plants during selection for resistant varieties. A comparison of the data presented above with the results of similar studies in the 1980s (Chulkina, 1991) indicates that the species composition of *Septoria* blotch has undergone some changes and has become more diverse in the regions. Note the appearance of *P. avenae* f. sp. *triticae* in the pathogenic complex of *Septoria* blotch at all the sampling points, which was not mentioned 40 years ago. The change in the species composition is probably associated with both climatic variations and a change in the spring wheat cultivating technology.

Table 2 shows the differences in the species composition of *Septoria* blotch on the spring wheat varieties from the Institute of Cytology and Genetics collection in Novosibirsky District of Novosibirsk Region. The table shows that three *Septoria* blotch pathogens were present on the spring wheat leaves: *P. nodorum*, *S. tritici*, and *P. avenae* f. sp. *triticae*. The main causative agent of *Septoria* leaf and ear blotch was *P. nodorum*, the occurrence of which averaged 85.4 % of the

Table 2. The species composition of causative agents of *Septoria* blotch in the spring wheat varieties in the full ripeness phase, %

Origin	Variety	*P. nodorum*	*S. tritici*	*P. avenae* f. sp. *triticae*
Russia, Novosibirsk Region	Novosibirskaya 15	85	10	5
	Novosibirskaya 31	90	10	0
	Sibirskaya 17	95	5	0
	Obskaya 2	85	15	0
Russia, Orenburg Region	Orenburgskaya 23	90	10	0
Russia, Kirov Region	Vyatchanka	95	5	0
Russia, Tyumen Region	Tyumenochka	80	10	10
Russia, Kurgan Region	Ariya	85	10	5
	Fora	80	10	10
	Zauralochka	90	10	0
Canada	NIL Thatcher Lr13	90	10	0
	NIL Thatcher Lr2c	85	10	5
China	Long Chun 7 Hao	85	15	0
	Ke Zhuang	90	10	0
USA	UI Alta Blanca	80	5	10
	UI Pettit	75	15	10
Kazakhstan	Kaiyr	80	20	0
	Dostyk	85	15	0
Switzerland	Quarna	90	10	0
Syria	Mayon 1	85	15	0
Germany	KWS Akvilon	70	20	10
Tajikistan	K65835	90	10	0
	K65834	85	15	0
Average		85.4	11.8	2.8

Table 3. The *Septoria* blotch incidence in the spring wheat varieties at the beginning of ripening phase by year, %

Origin	Variety	Flag leaf			Ear		
		2017	2018	Average	2017	2018	Average
		Russia					
Novosibirsk Region	Novosibirskaya 15	10.0	1.0	5.5	5.0	30.0	17.5
	Novosibirskaya 31	10.0	1.0	5.5	0	20.0	10.0
	Sibirskaya 17	1.0	0	0.5	1.0	20.0	10.5
	Obskaya 2	1.0	1.0	1.0	0	20.0	10.0
Orenburg Region	Orenburgskaya 23	15.0	5.0	10.0	0	0	0
Kirov Region	Vyatchanka	5.0	1.0	3.0	1.0	0	0.5
Tyumen Region	Tyumenochka	5.0	5.0	5.0	0	0	0
Kurgan Region	Ariya	20.0	5.0	12.5	1.0	0	0.5
	Fora	20.0	5.0	12.5	1.0	10.0	5.5
	Zauralochka	5.0	5.0	5.0	0	0	0
		Foreign countries					
Canada	NIL Thatcher Lr13	1.0	15.0	8.0	1.0	1.0	1.0
	NIL Thatcher Lr2c	5.0	5.0	5.0	1.0	0	0.5
China	Long Chun 7 Hao	5.0	5.0	5.0	1.0	1.0	1.0
	Ke Zhuang	0	5.0	2.5	0	5.0	2.5
USA	UI Alta Blanca	20.0	5.0	12.5	1.0	5.0	3.0
	UI Pettit	15.0	5.0	10.0	5.0	0	2.5
Kazakhstan	Kaiyr	10.0	5.0	7.5	1.0	0	0.5
	Dostyk	5.0	5.0	5.0	1.0	0	0.5
Switzerland	Quarna	0	1.0	0.5	1.0	1.0	1.0
Syria	Mayon 1	20.0	1.0	10.5	1.0	0	0.5
Germany	KWS Akvilon	0	10.0	5.0	1.0	1.0	1.0
Tajikistan	K65835	20.0	5.0	12.5	10.0	1.0	5.5
	K65834	40.0	10.0	25.0	10.0	5.0	7.5
Average		9.9	4.6	7.2	1.9	5.2	3.6

pathogenic complex in all the varieties. The second place in distribution on the wheat leaves was taken by *S. tritici* pycnia, 11.8 %, which reached a maximum of 20 % in the varieties 'Kaiyr' and 'KWS Akvilon'. The pathogen of the most limited occurrence was *P. avenae* f. sp. *triticae*: it was detected in only 8 varieties, and the average occurrence was 2.8 %. The data obtained indicate a predominantly regional confinement of the pathogen species composition. The varieties of different origins were infected with plant pathogens according to the regional type that is characteristic of Novosibirsk Region.

Assessment of the resistance of 23 spring wheat varieties to *Septoria* leaf and ear blotch in the northern forest-steppe of the Ob (Novosibirsk) region showed the absence of plant forms immune to *Septoria* blotch (Table 3).

The formation of *Septoria* blotch foci each year began on cereal grasses, such as common meadow-grass and cock's-foot, in which the disease severity by the first detecting date (July 5–7) had already reached 60 %. By that time, on winter wheat, the *Septoria* blotch symptoms had been detected on the second and third leaves from above and averaged 10 %. On the spring wheat, the *Septoria* blotch signs were noted only on the lower leaves and reached 3 % in the first ten-day period of July.

The years of research were wet, and the weather was favorable to the disease development. At the beginning of the loading phase, all the studied varieties showed signs of *Septoria* blotch infection; however, the intensity of the epiphytotic process varied significantly, both by variety and by plant organs in both years of research. For example, in 2017, the disease incidence ranged from 0 to 40 % on the flag leaves of spring wheat varieties and from 0 to 10 % on the spikes. In 2018, flag leaves were affected in the same varieties more

Table 4. The spring wheat seeds infection with *P. nodorum* across Siberia regions and by years of seed production, %

Region	2015	2016	2017	2018	Average
Tomsk Region	15.0 ± 2.64	3.3 ± 0.61	4.0 ± 0.53	14.3 ± 2.56	9.2 ± 1.65
Kemerovo Region	6.0 ± 0.92	4.6 ± 0.69	1.6 ± 0.22	5.3 ± 0.82	4.4 ± 0.68
Novosibirsk Region	5.2 ± 0.85	5.2 ± 0.81	1.7 ± 0.20	10.2 ± 2.21	5.6 ± 1.12
Krasnoyarsk Krai	5.0 ± 0.82	3.5 ± 0.52	1.8 ± 0.22	–	–
Tyumen Region	4.2 ± 0.71	1.8 ± 0.23	1.8 ± 0.24	12.2 ± 2.40	5.0 ± 0.95
Altai Krai	2.9 ± 0.48	2.7 ± 0.46	0.6 ± 0.13	7.3 ± 1.21	3.4 ± 0.62
Average	6.4 ± 1.12	3.5 ± 0.70	1.9 ± 0.21	9.9 ± 1.89	

Note: Influence of the factors: "region" – 15.1 % (level of significance 5 %); "year" – 52.5 % (level of significance 1 %).

evenly, from 0 to 15 %, and the wheat ears, in comparison, showed stark contrast, from 0 to 30 %. The tendency of differentiated manifestation of resistance to *Septoria* blotch on the leaves and ears of varieties was revealed. The correlation coefficient of *Septoria* blotch of leaves and spikes by variety was $r = 0.414 \pm 0.280$.

'Sibirskaya 17' and 'Obskaya 2' (Novosibirsk) were the most resistant to *Septoria* leaf blotch, with moderate spike damage. In both years of research, the flag leaf was affected at a sporadic level, providing grain loading. However, the resulting grain could become infected with *P. nodorum* and lead to the appearance of early *Septoria* blotch foci when sowing seeds in the following year.

Resistance to *Septoria* spike blotch was shown by 'Orenburgskaya 23' (Orenburg Region) and 'Ariya' (Kurgan Region), as well as the foreign varieties 'NIL Thatcher Lr13' (Canada), 'Kaiyr' (Kazakhstan), 'Mayon 1' (Syria), and 'KWS Akvilon' (Germany), with either unaffected or sporadically affected ears during the loading phase in both years of research. Flag leaves in these varieties were affected by *Septoria* blotch at the level of 10–20 %.

The domestic varieties 'Vyatchanka' (Kirov Region), 'Tyumenochka' (Tyumen Region), and 'Zauralochka' (Kurgan Region) and the foreign varieties 'NIL Thatcher Lr2c' (Canada), 'Long Chun 7 Hao' and 'Ke Zhuang' (China), 'Dostyk' (Kazakhstan), and 'Quarna' (Switzerland) showed complex resistance to both *Septoria* diseases, leaf blotch and ear blotch. This group of varieties was slightly affected by *Septoria* leaf and ear blotch, and the domestic varieties 'Tyumenochka' and 'Zauralochka' had a completely healthy spike at the beginning of the filling phase with a weak damage to the flag leaves.

The survey carried out in the phase of milk ripeness showed that the *Septoria* blotch severity reached 100 % in all varieties. The domestic varieties 'Orenburgskaya 23' and 'Vyatchanka' as well as 'Long Chun 7 Hao' from China showed a complex decreased susceptibility in the phase of milk ripeness. They had moderate, at the level of economic threshold (20 %), leaf and spike damage at the end of the growing season.

The variance analysis showed that the influence of the year conditions on the incidence of *Septoria* leaf blotch was 17.9–25.4 % (1 % significance level). The influence of the variety factor was 3.5–10 times lower and not always statistically significant.

Considering the ability of *P. nodorum* to use seeds as a transmission factor in time and to create early disease foci, we evaluated the intensity of seed infection in spring wheat varieties. The results of monitoring seed infection with *P. nodorum* in the Siberian regions are presented in Table 4.

The maximum infection of seed samples with *Septoria* blotch in the spring of 2016 (the year of seed production 2015) was noted in Tomsk Region, where it reached 36 %, which is more than 7 economic thresholds (Chulkina et al., 2017). In Krasnoyarsk Krai, seed infection reached 2.5 thresholds; in the remaining regions it was about 2 thresholds. The exception was Altai Krai, where in 2015 there were favorable conditions for obtaining high-quality seeds and the maximum infection rate reached 7 %.

The analysis of spring wheat seeds in the spring of 2017 showed that the infection of individual samples with *Septoria* blotch reached 4.4 thresholds, while the average seed infection only in Novosibirsk Region reached the threshold level. The remaining regions provided moderately infected seeds for analysis, on average below the economic threshold.

According to spring studies in 2018, the infection of spring wheat seeds with *P. nodorum* reached the economic threshold only in individual samples, while the average in all regions studied did not reach the economic threshold. The highest seed infection was noted in the more humid Tomsk Region. The most favorable situation for *Septoria* blotch was revealed in Altai Krai, where grain ripening in summer and September, 2017, in most areas took place under the dry weather. In general, the infection of the spring wheat seeds for sowing in 2018 was insignificant, lower than that in the previous years.

P. nodorum was detected in significant quantities on the spring wheat seeds produced in 2018 in all the regions. The average seed infection exceeded the economic threshold. On the seeds from Tomsk and Tyumen Regions the *P. nodorum* infection reached 7 thresholds, which should be considered a strong epidemic (Chulkina, 1991; Toropova et al., 2002). In 70 % of the seed samples the *Septoria* blotch threshold was exceeded, and by 38.5 % it was exceeded more than 2 times. In most West Siberian regions the 2018 growing season was quite moist, characterized by the epidemic severity of *Septoria* leaf and ear blotch, which provided *P. nodorum* with favorable conditions for seeds infection. In general, the 2019 spring analysis revealed the highest seed infection in recent

Table 5. The spring wheat seed infection with *P. nodorum* in collections from State variety plots by production regions

Variety	*Septoria* blotch incidence, %
Novosibirsk Region	
Novosibirskaya 47	9.0
Novosibirskaya 16	6.0
Novosibirskaya 41	6.0
Novosibirskaya 14	5.0
Novosibirskaya 31	7.0
Novosibirskaya 18	12.0
Novosibirskaya 29	6.0
Novosibirskaya 15	10.0
Obskaya 2	10.0
Average	7.9
Kurgan Region	
Zauralochka	3.0
Iset′	2.0
Tertsiya	2.0
Desyatka	1.0
Ariya	1.0
Raduga	3.0
Arka	3.0
Average	2.1
Omsk Region	
Uralosibirskaya	9.0
Omskaya 36	8.0
Sigma 2	4.0
Element 22	3.0
Omskaya 42	7.0
Sigma 2	5.0
Stolypinskaya	4.0
Average	5.7

years, which created the prerequisites for the early occurrence of a *Septoria* blotch epiphytotic process in all the regions of Siberia. The variance analysis showed that the influence of the factor "region" on the infection of spring wheat seeds with *P. nodorum*, reflecting the climate and cultivation technologies, is 3.4 times weaker than the influence of the factor "year weather conditions". In the more humid Tomsk Region the infection of spring wheat seeds with the water-depending pathogen *P. nodorum* has been 2 times higher on average over the years, compared to the least moistened Altai Krai, where the spring wheat grain production is concentrated mainly in arid warm zones. The correlation coefficients between *P. nodorum* seed infection and the total precipitation in August were $r = 0.746 \pm 0.135$ to 0.872 ± 0.126 by year and region (5 % significance level). The data presented indicate that it has relevance to the control of vertical transmission of *P. nodorum* with seeds of spring wheat varieties.

The analysis of spring wheat seeds from the breeding plots in Novosibirsk, Kurgan, and Omsk Regions (Table 5) indicates some differences in the activity of seed transmission of *P. nodorum* in years favorable for *Septoria* blotch. All the varieties from the Institute of Cytology and Genetics collection (Novosibirsk Region) ensured the transmission of pathogenic fungus at the level of economic threshold or 2.4 times higher; no varieties resistant to vertical transmission were detected. The varietal difference in the activity of vertical transmission of *P. nodorum* reached 2.4 times. In Kurgan Region the varieties transmitted *P. nodorum* 3.8 times more weakly; in none of the varieties did the seed infection reach the threshold. The varietal differences in seed infection reached 3 times. In Omsk Region the situation was intermediate: the transmission of the plant pathogen with seeds was on average 25 % less active than that in Novosibirsk Region and 2.7 times more active than that in Kurgan Region. In the collection of the Omsk State Agrarian University 4 varieties were identified in which the transmission of *Septoria* blotch was at or up to 1.8 times above the threshold. The varietal differences in the studied parameter reached 2.3 times.

Conclusion

Many years of studies have shown that *Septoria* leaf and ear blotch of spring wheat is widespread in the Siberian regions, reaching 35 % in terms of the disease incidence and 90 % in severity, which indicates the relevance of breeding resistant varieties. When developing breeding programs the species composition of *Septoria* blotch pathogens (*P. nodorum*, *S. tritici*, and *P. avenae* f. sp. *triticae*) should be taken into account – all the more so because it is characterized by significant regional differences. In Novosibirsk Region, *P. nodorum* completely dominated; *S. tritici* was 13.8 times less common; and *P. avenae* f. sp. *triticae* was a singleton. In Tyumen Region, the dominance of *P. nodorum* was disrupted in some geographic locations by *S. tritici* and *P. avenae* f. sp. *triticae*. In Altai Krai, *P. nodorum* predominated at all points studied; *S. tritici* and *P. avenae* f. sp. *triticae* were found everywhere, but 5.6 and 8.6 times less often, respectively.

Modern spring wheat varieties of different origins do not have complete immunity to *Septoria* blotch, but are characterized only by resistance or poor susceptibility to the disease. An independent manifestation of resistance to *Septoria* leaf blotch and *Septoria* ear blotch has been established. The correlation coefficients for the incidence of *Septoria* leaf and ear blotch have been $r = 0.323 \pm 0.241$ to 0.414 ± 0.280 over the years and varieties. Some varieties show relative resistance to *Septoria* leaf blotch with severe damage to the ear; others, in contrast, are resistant to *Septoria* ear blotch with severe damage to the leaf apparatus. Based on these data a cautious assumption can be made about the different genetics of resistance to *Septoria* leaf and ear blotch. The varieties 'Orenburgskaya 23' (Orenburg Region) and 'Ariya' (Kurgan Region), as well as the foreign varieties 'NIL Thatcher Lr13' (Canada), 'Kaiyr' (Kazakhstan), 'Mayon 1' (Syria), and

'KWS Akvilon' (Germany), show complex decreased susceptibility to the disease. They are weakly affected by *Septoria* leaf and ear blotch. The domestic varieties 'Tyumenochka' (Tyumen Region) and 'Zauralochka' (Kurgan Region) have a completely healthy ear at the beginning of the filling phase with weak damage to the flag leaf and can be considered the most promising sources of resistance.

The activity of vertical transmission of *P. nodorum* with seeds should also be monitored in the selection process, since seed transmission of *Septoria* blotch increases the infectious load on the plants significantly. Seed infection with the pathogen in the Siberian regions reached 7 thresholds and was largely determined by the weather conditions in August. The study of the collection of spring wheat varieties from three Siberian regions has revealed the following trend: transmission of *P. nodorum* with the seeds of varieties was the most active (7.6 %) in Novosibirsk Region and somewhat weaker in Omsk Region (5.7 %). The most favorable phytosanitary situation was in Kurgan Region, where varieties transmitted *P. nodorum* in a weak degree (2.1 %), below the threshold.

References

Adolf B., Schofl U., Verreet J.A. Effect of infection with *Septoria tritici* of different growth stages of wheat (GS 25 to GS 59) an dry matter production, nitrogen uptake and yield. Meded. Jc. Landbauw wetensch. 1993;53(3):1167-1174.

Broennimann A. Contributions to genetics of tolerance towards *Septoria nodorum* Berk. in wheat (*Triticum aestivum* L.). J. Plant Breed. 1975;75:138-160.

Chulkina V.A. Biological Fundamentals of Plant Epidemiology. Moscow, 1991. (in Russian)

Chulkina V.A., Toropova E.Yu., Stetsov G.Ya., Kirichenko A.A., Marmuleva E.Yu., Grishin V.M., Kazakova O.A., Selyuk M.P. Phytosanitary Diagnosis of Agroecosystems. Toropova E.Yu. (Ed.). Barnaul, 2017. (in Russian)

Cooke B.M., Jones D.G. The epidemiology of *S. tritici* and *S. nodorum*. II. Comparative studies of head infection by *S. tritici* and *S. nodorum* on spring wheat. Trans. Br. Mycol. Soc. 1970;56:121-125.

Du C.G., Nelson L.R., McDaniel M.E. Diallel analysis of gene effects conditioning resistance to *Stagonospora nodorum* (Berk.) in wheat. Crop Sci. 1999;39(3):686-690.

Eyal Z. The *Septoria tritici* and *Stagonospora nodorum* blotch diseases of wheat. Eur. J. Plant Pathol. 1999;105(7):629-641.

Gultyaeva E.I., Shaidayuk E.L., Levitin M.M., Maslova I.V., Kolesnikova O.A., Vusatyuk M.P. Phytosanitary state of grain crops in the North-West of Russia in 2018. Zashchita i Karantin Rasteniy = Plant Protection and Quarantine. 2019;5:30-31. (in Russian)

Jenkins P.D., Jones D.G. The effect of dual inoculation of wheat cultivars with *S. tritici* and *S. nodorum*. Phytopathologische Zeitschrift. 1981;101(3):210-221.

Kolomiets T.M., Pankratova L.F., Pakholkova E.V. Wheat (*Triticum* L.) cultivars from GRIN collection (USA) selected for durable resistance to *Septoria tritici* and *Stagonospora nodorum* Blotch. Selskokhozyaystvennaya Biologiya = Agricultural Biology. 2017;52(3): 561-569. DOI 10.15389/agrobiology.2017.3.561eng.

Kolomiets T.M., Shamanin V.P., Pakholkova E.V., Pankratova L.F., Salnikova N.N., Shepelev S.S., Pototskaya I.V., Abugalieva A., Morgunov A.I. Screening of hexaploid synthetic wheat lines and spring bread wheat variety for resistance to *Septoria* blotch. Vestnik Omskogo GAU = Bulletin of the Omsk State Agrarian University. 2018;3(31):13-26. (in Russian)

Levitin M.M. Microorganisms and global climate change. Selskokhozyaystvennaya Biologiya = Agricultural Biology. 2015;50(5): 641-647. DOI 10.15389/agrobiology.2015.5.641eng.

Mullaney E.J., Martin J.M., Scharen A.L. Generation mean analysis to identify and partition the components of genetic resistance to *Septoria nodorum* in wheat. Euphytica. 1981;31:539-545.

Nazarova L.N., Korneva L.G., Zhokhova T.P., Polyakova T.M., Sanin S.S. Epidemiological situation with *Septoria* disease on wheat in 2001–2009. Zashchita i Karantin Rasteniy = Plant Protection and Quarantine. 2010;10:18-20. (in Russian)

Nolan S., Cooke B.M., Monahan F.J. Studies on the interaction between *Septoria tritici* and *Stagonospora nodorum* in wheat. Eur. J. Plant Pathol. 1999;105(9):917-925.

Pakholkova E.V. The rate of leaf-stem infections development on cereal crops. Zashchita i Karantin Rasteniy = Plant Protection and Quarantine. 2015;3:39-40. (in Russian)

Parker S.R., Welham S., Paveley N.D., Foulkes J., Scott R.K. Tolerance of *Septoria* leaf blotch in winter wheat. Plant Pathol. 2004;53(1): 1-10.

Pyzhikova G.V., Sanina A.A., Kurakhtanova T.I. Septoria Blotch on Cereal Crops. Moscow, 1988. (in Russian)

Robert C., Bancal M.-O., Nicolas P., Lannou Ch., Ney B. Analysis and modelling of effects of leaf rust and *Septoria tritici* blotch on wheat growth. J. Exp. Bot. 2004;55(399):1079-1094.

Sanin S.S., Korneva L.G., Polyakov T.M. Forecasting the risk of epiphytotics developing of *Septoria* leaf and ear on wheat. Zashchita i Karantin Rasteniy = Plant Protection and Quarantine. 2015;3: 33-36. (in Russian)

Sanin S.S., Pakholkova E.V., Karlova L.V., Rulyova O.M., Franchuk L.I. Influence of weather factors on *S. tritici* and *S. nodorum* representation in a winter wheat Septoria blotch agent population. In: Epidemics of Plant Diseases: Monitoring, Prognosis, Control: Proc. of the Int. conf. Bolshie Vyazemy, 2017;1:81-84. (in Russian)

Sanin S.S., Pakholkova E.V., Sanina A.A., Karlova L.V., Korneeva L.G. Epidemiology of wheat leaf blotch: formation of infectious potentials of pathogens. Zashchita i Karantin Rasteniy = Plant Protection and Quarantine. 2018;5:25-29. (in Russian)

Scharen A.L., Krupinsky J.M. Cutural and inoculation studies of *Septoria nodorum*, cause of glume blotch of wheat. Phytopathology. 1970;60(10):1480-1485.

Simón M.R., Cordo C.A., Perelló A.E., Struik P.C. Influence of nitrogen supply on the susceptibility of wheat to *Septoria tritici*. J. Phytopathol. 2003;151(5):283-289.

Toropova E.Yu. Ecological Fundamentals of Plant Protection Against Diseases in Siberia. Novosibirsk, 2005. (in Russian)

Toropova E.Yu., Kazakova O.A., Piskarev V.V., Porsev I.N., Khristov Yu.A. The role of varieties and fungicides in the control of spring wheat septoriosis. Agrokhimiya = Agrochemistry. 2019;5:66-75. DOI 10.1134/S0002188119050107. (in Russian)

Toropova E.Yu., Kazakova O.A., Selyuk M.P. Monitoring of *Septoria* blight on spring wheat in the forest-steppe of Western Siberia. Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC. 2016;30(12):33-35. (in Russian)

Toropova E.Yu., Kazakova O.A., Selyuk M.P., Insebaeva M.K., Kirichenko A.A., Filipchuk O.D., Kvitko A.V. Contamination of wheat seeds with *Parastagonospora nodorum* Berk. Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC. 2018;32(12):15-19. DOI 10.24411/0235-2451-2018-11204. (in Russian)

Toropova E.Yu., Stetsov G.Ya., Chulkina V.A. Epiphytotiological Grounds of Plant Protection Systems. Chulkina V.A. (Ed.). Novosibirsk, 2002. (in Russian)

Tyryshkin L.G., Ershova M.A. Inheritance of resistance to Septoria blotch in common wheat sample MN81330. Russ. J. Genet. 2004; 40(4):454-457.

Van Ginkel M., Rajaram S. Breeding for resistance to the Septoria/ Stagonospora blights of wheat. In: Van Ginkel M., McNab A., Krupinsky J.M. (Eds.). Septoria and Stagonospora Diseases of Cereals: a Compilation of Global Research: Proc. of the Fifth Int. Septorial Workshop. Mexico, 1999:117-126.

Veselova S.V., Burkhanova G.F., Nuzhnaya T.V., Maksimov I.V. Effect of NADPH oxidases cascade on mechanisms regulating defense common wheat reactions under the disease *Stagonospora nodorum* blotch. Izvestiya Ufimskogo Nauchnogo Tsentra Rossiyskoy Akademii Nauk = Proceedings of the RAS Ufa Scientific Center. 2018; 3(1):66-74. (in Russian)

Veselova S.V., Burkhanova G.F., Nuzhnaya T.V., Rumyantsev S.D., Maksimov I.V. Effect of the host-specific toxin SnTOX3 from *Stagonospora nodorum* on ethylene signaling pathway regulation and redox-state in common wheat. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(7): 856-864.

ORCID ID

V.V. Piskarev orcid.org/0000-0001-9225-5227

Acknowledgements. This work was supported by state budget projects Nos. 0598-2015-0001 (All-Russian Research Institute of Phytopathology) and 0324-2019-0039 (Federal Research Center Institute of Cytology and Genetics SB RAS).

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

Received July 16, 2019. Revised January 18, 2020. Accepted January 21, 2020.

Вавиловский журнал генетики и селекции. 2020;24(2):149-157
DOI 10.18699/VJ20.602

Diversity and occurrence of methylotrophic yeasts used in genetic engineering

A.S. Rozanov[1] ✉, E.G. Pershina[1], N.V. Bogacheva[1], V. Shlyakhtun[1], A.A. Sychev[2], S.E. Peltek[1]

[1] Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Innovation Centre Biruch-NT, Belgorod region, Russia
✉ e-mail: sibiryak.n@gmail.com

Abstract. Methylotrophic yeasts have been used as the platform for expression of heterologous proteins since the 1980's. They are highly productive and allow producing eukaryotic proteins with an acceptable glycosylation level. The first *Pichia pastoris*-based system for expression of recombinant protein was developed on the basis of the tree-exudate-derived strain obtained in the US southwest. Being distributed free of charge for scientific purposes, this system has become popular around the world. As methylotrophic yeasts were classified in accordance with biomolecular markers, strains used for production of recombinant protein were reclassified as *Komagataella phaffii*. Although patent legislation suggests free access to these yeasts, they have been distributed on a contract basis. Whereas their status for commercial use is undetermined, the search for alternative stains for expression of recombinant protein continues. Strains of other species of methylotrophic yeasts have been adapted, among which the genus *Ogataea* representatives prevail. Despite the phylogenetic gap between the genus *Ogataea* and the genus *Komagataella* representatives, it turned out possible to use classic vectors and promoters for expression of recombinant protein in all cases. There exist expression systems based on other strains of the genus *Komagataella* as well as the genus *Candida*. The potential of these microorganisms for genetic engineering is far from exhausted. Both improvement of existing expression systems and development of new ones on the basis of strains obtained from nature are advantageous. Historically, strains obtained on the southwest of the USA were used as expression systems up to 2009. Currently, expression systems based on strains obtained in Thailand are gaining popularity. Since this group of microorganisms is widely represented around the world both in nature and in urban environments, it may reasonably be expected that new expression systems for recombinant proteins based on strains obtained in other regions of the globe will appear.
Key words: methylotrophic yeasts; *Pichia pastoris*; *Ogataea*; *Komagataella*; recombinant enzymes.

For citation: Rozanov A.S., Pershina E.G., Bogacheva N.V., Shlyakhtun V., Sychev A.A., Peltek S.E. Diversity and occurrence of methylotrophic yeasts used in genetic engineering. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):149-157. DOI 10.18699/VJ20.602

Разнообразие и распространение метилотрофных дрожжей, используемых в генной инженерии

А.С. Розанов[1] ✉, Е.Г. Першина[1], Н.В. Богачева[1], В. Шляхтун[1], А.А. Сычев[2], С.Е. Пельтек[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] ООО «Инновационный центр «Бирюч-НТ», Белгородская область, Россия
✉ e-mail: sibiryak.n@gmail.com

Аннотация. Метилотрофные дрожжи используются в качестве платформы для экспрессии гетерологичных белков с 1980-х гг. Они имеют высокий уровень продукции и позволяют получить белки эукариот с приемлемым уровнем гликозилирования. Первая система для экспрессии рекомбинантного белка на основе *Pichia pastoris* была разработана на основе штамма, выделенного из сокотечения деревьев на юго-западе США. Система распространялась бесплатно в научных целях и применяется во всем мире. В ходе классификации метилотрофных дрожжей по молекулярно-биологическим маркерам штаммы, используемые для получения рекомбинантного белка, были реклассифицированы как *Komagataella phaffii*. Они находятся в свободном доступе согласно патентному законодательству, однако распространялись на договорной основе. Это делает неопределенным их статус для коммерческого использования и, соответственно, стимулирует поиск альтернативных штаммов для экспрессии рекомбинантного белка. Были адаптированы штаммы других видов метилотрофных дрожжей, среди которых преобладают представители рода *Ogataea*. Несмотря на филогенетическую удаленность представителей рода *Ogataea* и *Komagataella*, во всех случаях оказалось возможным использовать классические векторы и промоторы для экспрессии рекомбинантного белка. Существуют системы экспрессии на основе других штаммов рода *Komagataella*, а также рода *Candida*. Потенциал этих микроорганизмов для генной инженерии далеко не исчерпан. Перспективно как усовершенствование имеющихся систем экспрессии, так и создание новых на основе штаммов, выделенных из природных источников. Исторически до 2009 г. в качестве систем экспрессии исполь-

зовались штаммы, выделенные на юго-западе США. В настоящее время начали развиваться системы экспрессии на основе штаммов, полученных в Таиланде. Поскольку эта группа микроорганизмов широко представлена по всему миру как в природной, так и в городской среде, можно ожидать появления систем экспрессии рекомбинантных белков, созданных на основе штаммов, выделенных и в других регионах планеты.
Ключевые слова: метилотрофные дрожжи; *Pichia pastoris*; *Ogataea*; *Komagataella*; рекомбинантные ферменты.

Introduction

Methylotrophs are a group of microorganisms that can use single-carbon methane compounds, such as methanol, methylamine, etc., as the sole source of carbon and energy. The requirement that all C–C bonds are formed enzymatically during cellular metabolism poses a challenge for the cell. Only some microorganisms are capable of doing so, such as Gram-negative proteobacteria and Gram-positive bacteria (Antony, 1986), as well as yeasts (Wegner, Harder, 1987), which employ metabolic pathways for methanol oxidation to produce energy and form C–C bonds. Both in yeasts and bacteria, C–C bonds are formed through the formation of formaldehyde (Yurimoto et al., 2005), a toxic intermediate product that subsequently is either dissimilated into CO_2 or assimilated into biomass.

Methylotrophic yeasts were discovered in the late 1960s, when methylotrophic bacteria had already been well known. The difference in times of their discovery was mainly caused by the challenges related to yeast isolation and significant bacterial contamination of samples (Trotsenko, Torgonskaya, 2011). The habitats of methylotrophic yeasts are those where the biomass is degraded to give rise to methoxy groups (soil, fallen trees, rotten fruit, etc.). Methanol is produced naturally during methane oxidation (e. g., by methane-oxidizing bacteria in the plant rhizosphere) and pectin or lignin degradation (MacDonald, Fall, 1993; Nakagawa et al., 2000). Most natural isolates of methylotrophic yeasts were detected in tree sap (exudates) or rotting wood (Kurtzman, Robnett, 1998; Kurtzman, 2005).

The interest in methylotrophic organisms for bioengineering applications arose in the early 1970s, when methanol was an inexpensive raw material and was considered a virtually inexhaustible fossil feedstock. However, after the 1973 oil crisis its price never dropped back to the pre-crisis level, and Western countries chose to lower their dependence on hydrocarbons. Plant-based proteins (and soybean protein in particular) came to fore and replaces fodder protein derived from unicellular methylotrophs.

In the 1980s, methylotrophic yeasts became widely used again, although in new areas. Today, they are utilized as a platform for genetic engineering and commercial-scale production of recombinant proteins. Furthermore, methylotrophic yeasts are a convenient object for studying the features of eukaryotic cell organization. Methods for utilizing methylotrophic yeasts in applied areas (as markers of coastal pollution, for treatment of discharge water at sulfate-cellulose and alcohol-and-liquor manufacturing enterprises, for treatment of formaldehyde-contaminated air, etc.) are currently being elaborated (Kutty, Philip, 2008; Trotsenko, Torgonskaya, 2012; dos Reis et al., 2018).

Pichia pastoris (*Komagataella phaffii*) is a methylotrophic yeast species that is most frequently used for scientific research and commercial purposes. These yeasts can consume both sugars and methanol to produce proteins with a high yield. Initially, back in the 1970s, Phillips Petroleum suggested using *P. pastoris* as a single-cell protein producer due to their ability to form high-density cultures both on glucose and methanol substrates (Mishra, Baranwal, 2009). The maximum cell density achieved during fermentation is higher than 100 g/L (dry basis) (Wegner, 1981). In the 1980s, the *P. pastoris*-based heterologous expression system was developed with the use of a strong and strictly-regulated alcohol oxidase 1 (*AOX1*) promoter (Cregg et al., 1985). In combination with the existing fermentation technologies for animal feed protein manufacturing, *AOX1* promoter was ensuring an exceptionally high recombinant proteins expression level. An important benefit of using it in recombinant protein manufacturing is the high expression level in the presence of methanol and robust inhibition of the process in the absence of methanol, which makes it possible to regulate the production of target proteins, including the autotoxic ones (Kurtzman, 2009). Production of biomass-derived hydroxynitrile lyase enzyme (the yield being 20 g of recombinant protein per liter of the culture) was one of the first large-scale commercial manufacturing processes established in the 1990s (Hasslacher et al., 1997).

Another advantage of *P. pastoris* is that it ensures the low glycosylation level. Thus, the *S. cerevisae* yeast species is no longer used in production of recombinant proteins, since it often hyperglycosylates proteins (up to their complete inactivation) (Darby et al., 2012).

The *P. pastoris*-based expression system has become widely used in basic research. Among methylotrophic yeasts, the *P. pastoris* species was the first to be developed. It turned out to be a convenient bioengineering system for production of eukaryotic proteins that are not properly expressed in bacteria. Phillips Petroleum made a forward-looking decision to make this expression system available to the scientific community for research purposes, which was a major driver for common application of this yeast species as a platform for recombinant protein expression. Histidine-auxotrophic strain GS115, reconstituted prototrophic strain X-33, aox1 knockout strains KM71 and KM71H, as well as protease-deficient strains SMD1168 and SMD1168H, and ade2 auxotrophic PichiaPink™ strain, are the most frequently utilized commercially available strains. However, their use for commercial purposes is restricted by material distribution policy. Therefore, searching for alternative options that could be used instead of the licensed *P. pastoris* strains is quite relevant today.

Phylogenic analysis

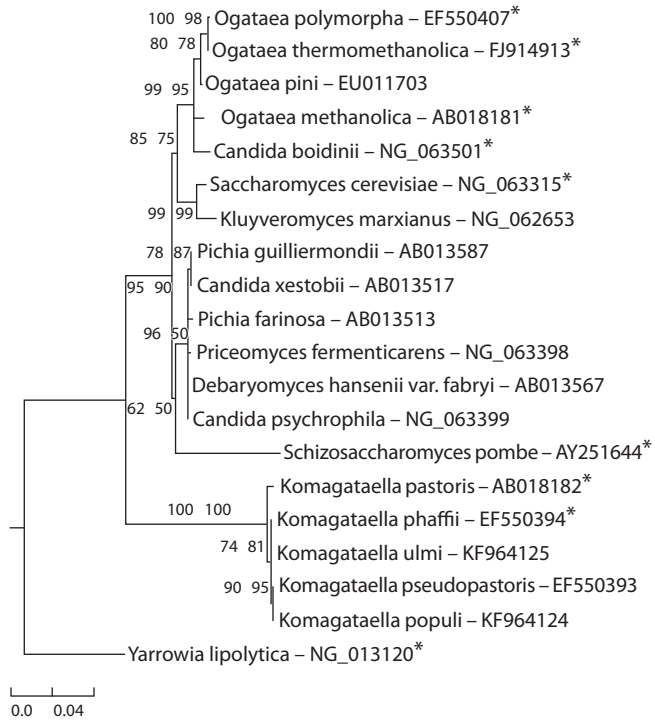
Progress in molecular biology methods as well as mounting evidence on gene sequences for various species made it possible to investigate into phylogeny of methylotrophic yeasts, which was accompanied by a number of surprises.

First of all, gene sequence analysis revealed that yeasts known as *Pichia* are not monophyletic despite the phenotypic similarity of the majority of species of this genus. First research works discovered that methanol-assimilating species *P. pastoris*, *P. angusta* and *Hansenula polymorpha* are remotely related with one another as well as with *Pichia membranifaciens*, the typical *Pichia* species. Yamada et al. (1994, 1995) suggested assigning *P. pastoris* to a new genus *Komagataella* and classifying *P. angusta* as *Ogataea polymorpha* into a freshly described genus *Ogataea*. New genera were suggested on the basis of analysis of divergence in partial large subunit (LSU) and small subunit (SSU) ribosomal RNA (rRNA) sequences. Since each analysis was limited to a relatively small number of species, it was unclear how close *P. pastoris* and *P. angusta* (*H. polymorpha*) were related to numerous unstudied species, and therefore suggestions were not accepted (Kurtzman, 1998). However, further analysis of D1/D2 regions in the LSU of the rRNA gene sequences for all now-known ascomycetes confirmed that the *Komagataella*, *Ogataea* and *Pichia* classification is correct (Kurtzman, Robnett, 1998), while phylogenetic divergence demonstrated as a result of single-gene sequence analysis was confirmed during the analysis of multi-gene sequences (Kurtzman et al., 2008). Therefore, *Komagataella*, *Ogataea* and *Pichia* are three separate genera. Thus, *Komagataella pastoris* is the most correct name from the viewpoint of taxonomy. However, due to historical reasons, *P. pastoris* remains most wide-spread today.

Secondly, it turned out that the name *P. pastoris* covers at least two yeast species, i. e. *Komagataella phaffii* and *K. pastoris* (Cregg et al., 1993; Kurtzman, 2009). The gene sequence analysis also revealed that these are not the only *Komagataella* representatives. *K. pseudopastoris* initially described as *P. pseudopastoris* (Dlauchy et al., 2003) was also classified as part of *Komagataella* on the basis of LSU and SSU rRNA sequence analysis. Similarly, *K. phaffii* was classified as *Komagataella* on the basis of the analysis of the D1/D2 fragment in the LSU rRNA gene (Kurtzman et al., 2008).

Evolution of parallel sequencing methods helped to build up the database on genome sequences of a great number of microorganisms (including methylotrophic yeasts), which allows for a more detailed phylogenetic analysis. 18S rRNA gene sequence is the most wide-spread and conservative marker for analysis of phylogenetic relations among various fungi representatives. Relevant genes were extracted from genomes of methylotrophic yeasts part of the NCBI database and used for the phylogenetic tree design (see the Figure).

Nucleotide sequences were evened with the help of MAFFT algorithm in MAFFT v7.312 program (Katoh, Standley, 2013) with –localpair and –maxiterate 1,000 parameters. Phylogenetic tree was designed on the basis of the maximum likelihood method with the help of IQ-TREE program (Trifinopoulos et al., 2016). –Auto and +R (FreeRate heterogeneity) parameters were used to determine the best patterns of nucleotide replacement models, minimum value of the Akaike criterion (AICc) being the key parameter. Two ratios, i. e. SH-like aLRT (1,000 replications) and ultrafast bootstrap (UfBoot, 1,000 replications), were used for statistical support of maximum likelihood in the IQ-TREE program.



Phylogenetic tree built on the basis of sequences of ribosomal 18S gene of Saccharomycetaceae species with the help of maximum likelihood method.

SH-Like aLTR and UfBoot support ratios are expressed in percentage on tree branch joints. Figures to the right of species names are identification numbers for relevant sequences in the NCBI data base.

As one can see, methylotrophic yeast used for production of recombinant proteins split into three groups: *Pichia*, *Ogataea* and *Komagataella* (see the Figure). Also, one can derive at a conclusion that *Ogataea* is phylogenetically more heterogeneous in comparison to *Pichia* and *Komagataella.*

Methylotrophic yeasts used for production of recombinant proteins as alternative strains, and their occurrence in nature

***P. pastoris* (*K. phaffii* and *K. kurtzmanii*).** The *P. pastoris* yeasts group, the majority of representatives of which have been reclassified as *Komagataella* representatives, is most well-studied among methylotrophic yeasts. Its typical habitat is tree sap (exudate) in regions from moderate to tropical. Initially, *P. pastoris* was isolated from chestnut exudates in France. Later, it was discovered to occur widely in Hungary and USA (Spencer et al., 1996; Negruta et al., 2010; Kurtzman, 2011a). Besides, *P. pastoris* yeasts were discovered in exudates of oil palm in Nigeria (Faparusi, 1974), sap of white algarrobo in Argentina (Spencer et al., 1995), and sap of red oak in Canada (Bowles, Lachance, 2007). The zoning of their occurrence is very curious: this is a predominant species of methylotrophic yeasts in the woods of the Pacific coast of the American Northwest; they are present (although not predominant) in Europe, Africa and South America, and are completely missing in Japan (Lachance et al., 1982).

Although the *P. pastoris*-based expression system patented by Philips Petroleum is wide-spread, alternative strains of

methylotrophic yeasts are being searched for. Those are required for invention of new unlicensed and thus applicable for commercial use platforms for recombinant proteins production.

P. pastoris (*K. phaffii*) CBS7435 strain is the closest to those previously suggested by Philips Petroleum. Strains derived from it are unlicensed (Ahmad et al., 2014). At the same time, this strain is a predecessor of patented strains most widely used for production of recombinant proteins these days. On the one hand, this paves the way for use of CBS7435 compounds thanks to the vast knowledge base built up with their help, and on the other hand, not all genome modifications for this strain can be patented taking into consideration their description in academic literature and in patents.

Strain *K. kurtzmanii* Y-727/KPB 2878/Starmer 75-208.2/CBS 12817/NRRL Y-63667 was patented in the Russian Federation. It was isolated by Prof. Starmer from fir sap in Arizona mountains, USA (Naumov et al., 2013), and is one of the closest relatives of the *P. pastoris* (*K. phaffii*) CBS7435 strain.

***P. guilliermondii* (*Meyerozyma guilliermondii*).** *P. guilliermondii* sporogenous species can be isolated from a variety of sources, i. e. plants, lake water, cow rumen, or oil-contaminated soil. Besides, these yeasts were discovered in elm-dwelling insects, in uncontaminated oil and water (Negruta et al., 2010a), in shrimp and other invertebrates, and in low-salinity sea water (Kutty, Philip, 2008). Before, this yeast species was used in gene engineering as a source of genes that were among other things expressed in *P. pastoris* (Handumrongkul et al., 1998; Zhang et al., 2009), rather than as an expression system.

Since the majority of methylotrophic yeasts have similar methanol-inducible promoter in methanol utilization paths (Hartner, Glieder, 2006), the research group from Malaysia tested the assumption that expression constructs developed on the basis of *K. phaffii* could be used for expression of recombinant proteins in other methylotrophic yeasts (Oslan et al., 2015). They isolated the strain called *Pichia sp. strain SO* from a rotten orange; its SSU sequence demonstrated its 100 % similarity to *P. guilliermondii*. Then, authors discovered that zeocin can be used as a marker for strain SO (Oslan et al., 2012), and conducted work on cloning the recombinant lipase expression construct. The work continued, and in 2017 an article on optimization of expression of T1 lipase isolated earlier from *Geobacillus zalihae* with use of *P. guilliermondii* was published. As the result of this work, T1 lipase yielded a 3-fold increase over medium (Abu et al., 2017).

P.* (*O.*) *methanolica was suggested by Invitrogen (USA) as a platform for production of recombinant proteins a little later than the same company suggested *P. pastoris*. Initially, it was isolated from soil sample in Japan in 1974; strains of this species were also isolated in the USA (Sibirny, 1996; Kurtzman, 2011b). In 2008, yeasts of this group (heterogeneous as *P. pastoris*) were also isolated on the territory of Russia from willow galls created by slug (Glushakova et al., 2010). *P.* (*O*). *methanolica* didn't become popular as a platform for production of recombinant proteins although there are single messages about its use, e. g., for expression of human glutamic acid decarboxylase (Raymond et al., 1998). Probably, the popularity of *P.* (*O*). *methanolica* as a platform for recombinant proteins is low because Invitrogen has a different, *P. pastoris*-based platform (*K. phaffii*).

Hansenula* (*O.*) *polymorpha is a thermotolerant methylotrophic yeast able to grow at temperatures below 50 °C wide-spread in nature. Besides such media as rotting fruit and other plants typical for methylotrophic yeasts, one of typical habitats for *H. polymorpha* is organism of insects, including *Drosophila melanogaster* typically used in research (Spencer J., Spencer D., 1997). Since *H. polymorpha* demonstrates good growth at high temperatures, it could possibly be found around hot springs and in tropical areas.

In science, these yeasts have been used as a model organism for studying peroxisome biogenesis and degradation mechanisms, methanol metabolism control, assimilation of nitrates and reaction to stress (van der Klei et al., 2006). *H. polymorpha* turned out rather effective for production of recombinant proteins as well (Gellissen, 2005). Recombinant antigen of the hepatitis B virus (HBsAg), that was successfully commercialized under HepaVax-Gene and AgB trademarks (Seo et al., 2008), is the most significant therapeutic protein produced with the help of *H. polymorpha*. Producers of recombinant proteins with high potential for pharmaceutical purposes were developed on the basis of *H. polymorpha*: hirudin from leech *Hirudinaria manillensis* (Weydemann et al., 1995) and some human proteins including α1-antitrypsin (Kang et al., 1998), IFNα-2a (Degelmann et al., 2002), serum albumin (Kang et al., 2001), epidermal growth factor (Heo et al., 2002) and parathyroid hormone (Sohn et al., 2012). Besides medical proteins, there were developed producers of food and commercial enzymes: hexose oxidases (Cook, Thygesen, 2003), phytases (Mayer et al., 1999), levansucrase from *Zymomonas mobilis* (Park et al., 2004), and glucose oxydases from *Aspergillus niger* (Kim et al., 2004).

P.* (*O.*) *thermomethanolica. Besides strains utilized by Invitrogen, there is *O. thermomethanolica* BCC16875 strain that can be considered one of best-studied methylotrophic yeasts alternatives for production of recombinant proteins. Knowledge of occurrence of *O. thermomethanolica* is very scarce as this yeast species was discovered only recently, in 2005, in soil samples in Thailand (Limtong et al., 2013).

For the first time, information about strain BCC16875 was published by a research group from Thailand in 2012. The research focuses on testing the possibility of using biomolecular tools for accumulation of protein in this strain. Classical methanol-inducible alcohol oxidase (*AOX1*) promoters and constitutive glyceraldehyde 3-phosphate dehydrogenase (*GAP*) promoters utilized for working with *P. pastoris* were shown to drive efficient gene expression in this new strain. Recombinant phytase and xylanase were expressed from both promoters as secreted proteins, with the former demonstrating different patterns of N-glycosylation dependent on the promoter and culture medium used. The major glycoprotein oligosaccharide species produced from *O. thermomethanolica* BCC16875 is Man8-12GlcNAc2 that is similar to that of other methylotrophs. Moreover, mannosylphosphate and α-1,6- and α-1,2-linked mannose modifications of heterologous secreted protein were also detected. The level of expression of recombinant protein turned out to be equal to the level of expression of commercial strains, which

makes the suggested platform a good alternative to widely used Invitrogen's strains (Tanapongpipat et al., 2012).

Studies of the suggested strain *O. thermomethanolica* BCC16875 continued during next following years. Promoters typical for this train were studied (Harnpicharnchai et al., 2014; Promdonkoy et al., 2014), as well as methods of high-density cultivation for expression of recombinant proteins (Charoenrat et al., 2016). Besides, work on optimization of strain's metabolism for increase of target products' output was started. Thus, the level of expression of auxiliary proteins of endoplastic reticulum was increased for this purpose (Roongsawang et al., 2016). Studies of the strain became especially active after 2016, and today there is a significant number of works on updating it to the present-day level as an expression system. In 2018, the CRISPR-Cas9 system was adapted for this strain (Phithakrotchanakoon et al., 2018b), a new sucrose-induction-based expression system was developed (Puseenam et al., 2018; Boonchoo et al., 2019), and studies of metabolism at proteomic and transcriptomic levels continued (Phithakrotchanakoon et al., 2018a).

Candida boidinii is the first described species of methylotrophic yeasts. It is also apparently most wide-spread in natural habitat (Ogata et al., 1969). Mainly it's various plant substrates (tree sap, rotten fruit, some flowers). These yeasts are also abundantly present in naturally fermented olives (Coton et al., 2006). Methylotrophic yeasts are also found in cacti. Their spoiled parts happen to host both *C. boidinii*, and *O. polymorpha*, although these species do not prevail among yeasts discovered in these yeast samples. In addition to regular occurrences of methylotrophic yeasts, *C. boidinii* is also an important marker of seashore contamination. Their lines are predominant in many water and sand samples in Brazil (Kutty, Philip, 2008).

According to the phylogenetic tree of 18S rRNA gene sequences presented in this review, *C. boidinii* could be classified as *Ogataea* (see the Figure). Up to 2009, *C. boidinii* was developed as a platform for production of recombinant proteins alternative to *P. pastor*is by a group of Japanese scientists (Yurimoto, Sakai, 2009). *C. boidinii*-based recombinant protein expression system has some characteristics that can be useful in comparison to other methylotrophic yeasts. Level of expression in *C. boidinii* varies depending on the source of carbon: *AOD1* promoter demonstrates high level of expression in methanol-grown or methanol-glycerol-grown cells, medium level of expression in glycerol-grown cells and zero expression in case glucose- or ethanol-grown cells are used as source of carbon (Sakai et al., 1995; Yurimoto et al., 2000). The level of expression is significantly higher for *C. boidinii* than for *O. polymorpha*. Besides, high level of expression can be ensured in *C. boidinii* in case of methanol + glycerol medium, which allows shortening the time for high-density cell cultivation. In case of *P. pastoris*, glycerol suppresses expression of methanol-induced genes, and therefore control over complete eating of glycerol in the culture prior to methanol induction is required. A strain with knocked-out vacuolar proteinase A (PEP4) and proteinase B (PRB1) is available for both *P. pastoris* and *C. boidinii* (Komeda et al., 2002). During studies with use of *C. boidinii* genes expression system, there were developed strains for production of toxic proteins, i. e. membrane-bound peroxisome allowing to cumulate toxic proteins (Nishikawa et al., 2000; Yurimoto et al., 2001), as well as effective secretion system for production of active transglutaminase (Yurimoto et al., 2004).

Obtainment of methylotrophic yeasts from natural sources

Yeasts are isolated from water, seawater, atmosphere and ground habitats. They dwell in rotting vegetables and fruit, in moulds, exudates of trees and their barks, in xylophage insects, pig's intestine, milk of cows suffering from mastitis, in forest, garden and swampy soils, especially drenched with sewage waters, in sea weed and so on (Negruta et al., 2010; Trotsenko, Torgonskaya, 2011). These environmental preferences are most likely due to discharge of methoxyl groups during degradation of lignin and pectin (Nakagawa et al., 2005). In addition to fruit juices and soil samples, methylotrophic yeasts are also found in food (Mu et al., 2012; Kozhakhmetov et al., 2016; Syromyatnikov et al., 2018).

Many types of yeast are wide-spread while some are limited only to a certain narrow habitat. They rarely occur in nature in absence of micellar fungi and bacteria. Therefore, to obtain them one must use selective methods allowing yeasts to have advantages in growth speed. When media for selective isolation of yeasts are developed, low pH is usually used as in the majority of cases yeasts prevail over bacteria in such conditions. Media could also include antibiotics for suppression of bacteria and fungistatic agents for suppression of moulds (Kurtzman et al., 2011).

When yeasts are present in great amounts, they can be isolated by direct application of the material or its suspension on sour agarized medium that can also be enriched with antibiotics or have other selecting properties. Agar hydrolyzes in low-pH medium during autoclaving. Therefore agar and medium are sterilized separately, cooled to around 45 °C, mixed and distributed among Petrie dishes. The majority of yeast species can be isolated at 3.7 pH, but some species such as *Schizosaccharomyces* species require higher pH ranging from 4.5 to 5.0. If yeasts are present in the sample in low quantities, their population could be increased by preliminary incubation of the sample in liquid medium at pH up to 3.8 (Kurtzman et al., 2011).

To isolate specific physiological groups of yeasts, it's necessary to find additional selecting parameters. Methylotrophic yeasts can be selected with use of methanol as the sole source of carbon and energy in the medium. Thus, when methylotrophic yeasts were isolated from grape leaves in Thailand, YNB medium with extra 0.5 % of methanol was used to get the enrichment culture. Cultivation continued for 4–5 days at 27 °C, following which enrichment cultures were spread on 0.5 % v/v methanol-YNB agar. As a result, 2 new species were isolated that classified as *Ogataea* (Limtong et al., 2013).

Conclusion

Methylotrophic yeasts are wide-spread as a platform for production of recombinant protein. Initially, they became of interest to biotech companies as single-cell protein producers. However, due to the 1973 oil crisis, methanol grew in price and isolation of feed protein from it became irrelevant. At

the same time, due to discovery of proteins key to molecular biology such as thermotolerant polymerases, lygases and restrictases, various microorganisms modification methods started being developed, including those for production of recombinant proteins.

Production of proteins with use of microorganisms for various purposes, in the first place, for food and animal feed as well as technological purposes, has been actively developing since 1940s. In the first place, this was connected with the use of natural producers, upgraded in many cases with the help of undirected mutagenesis methods. As molecular biology methods developed, it became possible to develop protein producers untypical for a specific organism (heterologous or recombinant proteins), including with use of yeasts.

Following this trend, Phillips Petroleum developed their own *P. pastoris*-based expression system for production of recombinant proteins (later renamed into *Komagataella phaffii*). Thanks to its outstanding properties as well as Phillips Petroleum's decision to allow its wide-spread utilization for scientific purposes around the world, it got widely popular as a recombinant protein production platform.

Understanding of key properties that made *P. pastoris*-based expression systems popular is important for further development of protein-expressing platforms. Number one is high protein production level. It's lower than that of bacteria and micellar fungi but higher than that of other systems, i. e. cells of mammals, plants and insects. Number two: yeasts are eukaryotes, and they have all cell compartments necessary for synthesis and assembly of eukaryotic proteins, which allows them to synthesize proteins that cannot be synthesized with use of bacterial expression systems. Number three is extracellular protein's ability to synthesize, which brings the costs of the production process down. Number four is the low level of glycosylation of proteins in comparison to many other types of yeast.

Other methylotrophic yeasts, except *K. phaffii*, have also been tried out as a platform for expression of recombinant protein. In the first place, it should be noted that *P. pastoris* strain CBS7435 initial for the Invitrogen's system was suggested to be used as a patent-independent platform. In Russia, *Komagataella kurtzmanii* strain of American origin was patented. Alternative *Ogataea*-based expression platforms were developed in the USA: *Hansenula* (*Ogataea*) *polymorpha* and *Pichia* (*Ogataea*) *methanolica*, the latter belongs to Invitrogen. In Japan, there was developed an alternative platform on the basis of *Candida boidinii*. This system was being developed until 2009, but since then this organism has not been mentioned as a platform used for production of recombinant protein. In Malaysia, strain *P. guilliermondii* (*Meyerozyma guilliermondii*) was suggested for production of protein.

One can notice that all yeast strains used for production of heterologous proteins up to 2009 have American origin. This sets a question, whether or not American isolates have unique properties allowing for production of protein in yeasts. More recent works witness that it's not accurate. First strains were developed by American researchers working in one state. Actually, rapid development of bioscience, including discovery of new methods of molecular biology, took place in the US Southwest. It seems that this fact was the key reason for using strains of this region specifically. The Russian patent is actually the copy of known systems, which requires utilization of the closest species that in the majority of cases have geographically similar habitats.

Overall, alternative methylotrophic-yeasts-based systems haven't become popular. Reasons are multiple. Lack of interest on part of recombinant proteins market players and disability of new players to enter this market with a new system along with lack of acute need for it is the main reason. For example, Chinese manufacturers actively use strains obtained on the basis of Invitrogen's strains. It shall also be noted that modern strains may be significantly genetically modified to ensure higher yields of recombinant proteins.

References

Abu M.L., Nooh H.M., Oslan S.N., Salleh A.B. Optimization of physical conditions for the production of thermostable T1 lipase in *Pichia guilliermondii* strain SO using response surface methodology. BMC Biotechnol. 2017;17:78. DOI 10.1186/s12896-017-0397-7.

Ahmad M., Hirz M., Pichler H., Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl. Microbiol. Biotechnol. 2014;98:5301-5317. DOI 10.1007/s00253-014-5732-5.

Anthony C. Bacterial oxidation of methane and methanol. Adv. Microbial Physiol. 1986;27:113-210. DOI 10.1016/S0065-2911(08)60305-7.

Boonchoo K., Puseenam A., Kocharin K., Tanapongpipat S., Roongsawang N. Sucrose-inducible heterologous expression of phytase in high cell density cultivation of the thermotolerant methylotrophic yeast *Ogataea thermomethanolica*. FEMS Microbiol. Lett. 2019;366(5). DOI 10.1093/femsle/fnz052.

Bowles J.M., Lachance M.-A. Patterns of variation in the yeast florae of exudates in an oak community. Can. J. Bot. 2007;61(12):2984-2995. DOI 10.1139/b83-335.

Charoenrat T., Antimanon S., Kocharin K., Tanapongpipat S., Roongsawang N. High cell density process for constitutive production of a recombinant phytase in thermotolerant methylotrophic yeast *Ogataea thermomethanolica* using table sugar as carbon source. Appl. Biochem. Biotechnol. 2016;180:1618-1634. DOI 10.1007/s12010-016-2191-8.

Cook M.W., Thygesen H.V. Safety evaluation of a hexose oxidase expressed in *Hansenula polymorpha*. Food Chem. Toxicol. 2003;41:523-529. DOI 10.1016/S0278-6915(02)00280-6.

Coton E., Coton M., Levert D., Casaregola S., Sohier D. Yeast ecology in French cider and black olive natural fermentations. Int. J. Food Microbiol. 2006;108:130-135. DOI 10.1016/j.ijfoodmicro.2005.10.016.

Cregg J.M., Barringer K.J., Hessler A.Y., Madden K.R. *Pichia pastoris* as a host system for transformations. Mol. Cell. Biol. 1985;(12):3376-3385. DOI 10.1128/mcb.5.12.3376.

Cregg J.M., Vedvick T.S., Raschke W.C. Recent advances in the expression of foreign genes in *Pichia pastoris*. Biotechnology (NY). 1993;11:905-910. DOI 10.1038/nbt0893-905.

Darby R.A.J., Cartwright S.P., Dilworth M.V., Bill R.M. Which yeast species shall I choose? *Saccharomyces cerevisiae* versus *Pichia pastoris* (Review). In: Bill R.M. (Ed.). Recombinant Protein Production in Yeast: Methods and Protocols. Humana Press, 2012;11-23. DOI 10.1007/978-1-61779-770-5_2.

Degelmann A., Müller F., Sieber H., Jenzelewski V., Suckow M., Strasser A.W.M., Gellissen G. Strain and process development for the production of human cytokines in *Hansenula polymorpha*. FEMS Yeast Res. 2002;2:349-361. DOI 10.1016/S1567-1356(02)00096-X.

Dlauchy D., Tornai-Lehoczki J., Fülöp L., Péter G. *Pichia* (*Komagataella*) *pseudopastoris* sp. nov., a new yeast species from Hun-

gary. Antonie van Leeuwenhoek. 2003;83:327-332. DOI 10.1023/A:1023318829389.

dos Reis K.C., Arrizon J., Amaya-Delgado L., Gschaedler A., Schwan R.F., Silva C.F. Volatile compounds flavoring obtained from Brazilian and Mexican spirit wastes by yeasts. World J. Microbiol. Biotechnol. 2018;34(152). DOI 10.1007/s11274-018-2535-3.

Faparusi S.I. Microorganisms from oil palm tree (*Elaeis guineensis*) tap holes. J. Food Sci. 1974;39(4):755-757. DOI 10.1111/j.1365-2621.1974.tb17972.x.

Gellissen G. (Ed.). Production of Recombinant Proteins: Novel Microbial and Eukaryotic Expression Systems. John Wiley & Sons, 2005.

Glushakova A.M., Maximova I.A., Kachalkin A.V., Yurkov A.M. *Ogataea cecidiorum* sp. nov., a methanol-assimilating yeast isolated from galls on willow leaves. Antonie van Leeuwenhoek. 2010;98(1):93-101. DOI 10.1007/s10482-010-9433-5.

Handumrongkul C., Ma D.P., Silva J.L. Cloning and expression of *Candida guilliermondii* xylose reductase gene (*xyl1*) in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 1998;49:399-404. DOI 10.1007/s002530051189.

Harnpicharnchai P., Promdonkoy P., Sae-Tang K., Roongsawang N., Tanapongpipat S. Use of the glyceraldehyde-3-phosphate dehydrogenase promoter from a thermotolerant yeast, *Pichia thermomethanolica*, for heterologous gene expression, especially at elevated temperature. Ann. Microbiol. 2014;64:1457-1462. DOI 10.1007/s13213-013-0765-z.

Hartner F.S., Glieder A. Regulation of methanol utilisation pathway genes in yeasts. Microb. Cell Fact. 2006;5(39). DOI 10.1186/1475-2859-5-39.

Hasslacher M., Schall M., Hayn M., Bona R., Rumbold K., Lückl J., Griengl H., Kohlwein S.D., Schwab H. High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts. Protein Expr. Purif. 1997;11:61-71. DOI 10.1006/prep.1997.0765.

Heo J.H., Won H.S., Kang H.A., Rhee S.K., Chung B.H. Purification of recombinant human epidermal growth factor secreted from the methylotrophic yeast *Hansenula polymorpha*. Protein Expr. Purif. 2002;24(1):117-122. DOI 10.1006/prep.2001.1527.

Kang H.A., Kang W., Hong W.K., Kim M.W., Kim J.Y., Sohn J.-H., Choi E.S., Choe K.B., Rhee S.K. Development of expression systems for the production of recombinant human serum albumin using the MOX promoter in *Hansenula polymorpha* DL-1. Biotechnol. Bioeng. 2001;76(2):175-185. DOI 10.1002/bit.1157.

Kang H.A., Sohn J.-H., Choi E.-S., Chung B.H., Yu M.-H., Rhee S.-K. Glycosylation of human α1-antitrypsin in *Saccharomyces cerevisiae* and methylotrophic yeasts. Yeast. 1998;14(4):371-381. DOI 10.1002/(SICI)1097-0061(19980315)14:4<371::AID-YEA231>3.0.CO;2-1.

Katoh K., Standley D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 2013;30(4):772-780. DOI 10.1093/molbev/mst010.

Kim M.W., Rhee S.K., Kim J.Y., Shimma Y.I., Chiba Y., Jigami Y., Kang H.A. Characterization of N-linked oligosaccharides assembled on secretory recombinant glucose oxidase and cell wall mannoproteins from the methylotrophic yeast *Hansenula polymorpha*. Glycobiology. 2004;14(3):243-251. DOI 10.1093/glycob/cwh030.

Komeda T., Sakai Y., Kato N., Kondo K. Construction of protease-deficient *Candida boidinii* strains useful for recombinant protein production: cloning and disruption of proteinase A gene (*PEP4*) and proteinase B gene… Biosci. Biotechnol. Biochem. 2002;66(3):628-631. DOI 10.1271/bbb.66.628.

Kozhakhmetov S., Tynybayeva I., Baikhanova D., Saduakhasova S., Saduakhasova G., Saduakhasova A., Nurgozhin T., Zhumadilov Z. Metagenomic analysis of koumiss in Kazakhstan. Cent. Asian J. Glob. Health. 2016;3(Suppl.). DOI 10.5195/cajgh.2014.163.

Kurtzman C.P. Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella*. Int. J. Syst. Evol. Microbiol. 2005;55:973-976. DOI 10.1099/ijs.0.63491-0.

Kurtzman C.P. *Pichia* E.C. Hansen emend. Kurtzman. In: Kurtzman C.P., Fell J.W. (Eds.). The Yeasts: A Taxonomical Study. Elsevier, Bv, 1998;273-352. DOI 10.1016/b978-044481312-1/50046-0.

Kurtzman C.P. Biotechnological strains of *Komagataella* (*Pichia*) *pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. J. Ind. Microbiol. Biotechnol. 2009;36:1435-1438. DOI 10.1007/s10295-009-0638-4.

Kurtzman C.P. *Komagataella* Y. Yamada, Matsuda, Maeda & Mikata (1995). In: Kurtzman C.P., Fell J.W., Boekhout T. (Eds.). The Yeasts: A Taxonomical Study. 5th edn. 2011a;491-495. DOI 10.1016/B978-0-444-52149-1.00037-9.

Kurtzman C.P. *Ogataea* Y. Yamada, K. Maeda & Mikata (1994). In: Kurtzman C.P., Fell J.W., Boekhout T. (Eds.). The Yeasts: A Taxonomical Study. 5th edn. 2011b;645-671. DOI 10.1016/B978-0-444-52149-1.00053-7.

Kurtzman C.P., Fell J.W., Boekhout T., Robert V. Methods for isolation, phenotypic characterization and maintenance of yeasts. In: Kurtzman C.P., Fell J.W., Boekhout T. (Eds.). The Yeasts: A Taxonomical Study. 5th edn. 2011;87-110. DOI 10.1016/B978-0-444-52149-1.00007-0.

Kurtzman C.P., Robnett C.J. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie van Leeuwenhoek. 1998;73:331-371. DOI 10.1023/A:1001761008817.

Kurtzman C.P., Robnett C.J., Basehoar-Powers E. Phylogenetic relationships among species of *Pichia*, *Issatchenkia* and *Williopsis* determined from multigene phylogenetic analysis and the proposal of *Barnettozyma* gen. nov., *Lindnera* gen. nov. and *Wickerhamomyces* gen. nov. FEMS Yeast Res. 2008;8(6):939-954. DOI 10.1111/j.1567-1364.2008.00419.x.

Kutty S.N., Philip R. Marine yeasts – a review. Yeast. 2008;25:465-483. DOI 10.1002/yea.1599.

Lachance M.A., Metcalf B.J., Starmer W.T. Yeasts from exudates of *Quercus*, *Ulmus*, *Populus*, and *Pseudotsuga*: new isolations and elucidation of some factors affecting ecological specificity. Microb. Ecol. 1982;8:191-198. DOI 10.1007/BF02010452.

Limtong S., Kaewwichian R., Groenewald M. *Ogataea kanchanaburiensis* sp. nov. and *Ogataea wangdongensis* sp. nov., two novel methylotrophic yeast species from phylloplane in Thailand. Antonie van Leeuwenhoek. 2013;103:551-558. DOI 10.1007/s10482-012-9837-5.

Limtong S., Srisuk N., Yongmanitchai W., Yurimoto H., Nakase T., Kato N. *Pichia thermomethanolica* sp. nov., a novel thermotolerant, methylotrophic yeast isolated in Thailand. Int. J. Syst. Evol. Microbiol. 2005;55:2225-2229. DOI 10.1099/ijs.0.63712-0.

MacDonald R.C., Fall R. Detection of substantial emissions of methanol from plants to the atmosphere. Atmos. Environ. 1993;27(11):1709-1713. DOI 10.1016/0960-1686(93)90233-O.

Mayer A.F., Hellmuth K., Schlieker H., Lopez-Ulibarri R., Oertel S., Dahlems U., Strasser A.W.M., Van Loon A.P.G.M. An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. Biotechnol. Bioeng. 1999;63(3):373-381. DOI 10.1002/(SICI)1097-0290(19990505)63:3<373::AID-BIT14>3.0.CO;2-T.

Mishra S., Baranwal R. Yeast genetics and biotechnological applications. In: Satyanarayana T., Kunze G. (Eds). Yeast Biotechnology: Diversity and Applications. Springer, 2009;323-355. DOI 10.1007/978-1-4020-8292-4_16.

Mu Z., Yang X., Yuan H. Detection and identification of wild yeast in koumiss. Food Microbiol. 2012;31:301-308. DOI 10.1016/j.fm.2012.04.004.

Nakagawa T., Miyaji T., Yurimoto H., Sakai Y., Kato N., Tomizuka N. A methylotrophic pathway participates in pectin utilization by *Candida boidinii*. Appl. Environ. Microbiol. 2000;66(10):4253-4257. DOI 10.1128/AEM.66.10.4253-4257.2000.

Nakagawa T., Yamada K., Fujimura S., Ito T., Miyaji T., Tomizuka N. Pectin utilization by the methylotrophic yeast *Pichia methanolica*. Microbiology. 2005;151(6):2047-2052. DOI 10.1099/mic.0.27895-0.

Naumov G.I., Naumova E.S., Tyurin O.V., Kozlov D.G. *Komagataella kurtzmanii* sp. nov., a new sibling species of *Komagataella* (*Pichia*) *pastoris* based on multigene sequence analysis. Antonie van Leeuwenhoek. 2013;104:339-347. DOI 10.1007/s10482-013-9956-7.

Negruta O., Csutak O., Stoica I., Rusu E., Vassu T. Methylotrophic yeasts: diversity and methanol metabolism. Rom. Biotechnol. Lett. 2010;15(4):5369-5375.

Nishikawa M., Hagishita T., Yurimoto H., Kato N., Sakai Y., Hatanaka T. Primary structure and expression of peroxisomal acetylspermidine oxidase in the methylotrophic yeast *Candida boidinii*. FEBS Lett. 2000;476:150-154. DOI 10.1016/S0014-5793(00)01708-7.

Ogata K., Nishikawa H., Ohsugi M. A yeast capable of utilizing methanol. Agr. Biol. Chem. 1969;33(10):1519-1520. DOI 10.1080/00021369.1969.10859497.

Oslan S.N., Salleh A.B., Rahman R.N.Z.R.A., Basri M., Chor A.L.T. Locally isolated yeasts from Malaysia: identification, phylogenetic study and characterization. Acta Biochim. Pol. 2012;59(2):225-229. DOI 10.18388/abp.2012_2142.

Oslan S.N., Salleh A.B., Rahman R.N.Z.R.A., Leow T.C., Sukamat H., Basri M. A newly isolated yeast as an expression host for recombinant lipase. Cell. Mol. Biol. Lett. 2015;20(2):279-293. DOI 10.1515/cmble-2015-0015.

Park B.S., Vladimir A., Kim C.H., Rhee S.K., Kang H.A. Secretory production of *Zymomonas mobilis* levansucrase by the methylotrophic yeast *Hansenula polymorpha*. Enzyme Microb. Technol. 2004;34:132-138. DOI 10.1016/j.enzmictec.2003.09.005.

Phithakrotchanakoon C., Puseenam A., Phaonakrop N., Roytrakul S., Tanapongpipat S., Roongsawang N. Hac1 function revealed by the protein expression profile of a Ot*HAC1* mutant of thermotolerant methylotrophic yeast *Ogataea thermomethanolica*. Mol. Biol. Rep. 2018a;45:1311-1319. DOI 10.1007/s11033-018-4287-4.

Phithakrotchanakoon C., Puseenam A., Wongwisansri S., Eurwilaichitr L., Ingsriswang S., Tanapongpipat S., Roongsawang N. CRISPR-Cas9 enabled targeted mutagenesis in the thermotolerant methylotrophic yeast *Ogataea thermomethanolica*. FEMS Microbiol. Lett. 2018b;365(11). DOI 10.1093/femsle/fny105.

Promdonkoy P., Tirasophon W., Roongsawang N., Eurwilaichitr L., Tanapongpipat S. Methanol-inducible promoter of thermotolerant methylotrophic yeast *Ogataea thermomethanolica* BCC16875 potential for production of heterologous protein at high temperatures. Curr. Microbiol. 2014;69:143-148. DOI 10.1007/s00284-014-0568-x.

Puseenam A., Kocharin K., Tanapongpipat S., Eurwilaichitr L., Ingsriswang S., Roongsawang N. A novel sucrose-based expression system for heterologous proteins expression in thermotolerant methylotrophic yeast *Ogataea thermomethanolica*. FEMS Microbiol. Lett. 2018;365(20). DOI 10.1093/femsle/fny238.

Raymond C.K., Bukowski T., Holderman S.D., Ching A.F.T., Vanaja E., Stamm M.R. Development of the methylotrophic yeast *Pichia methanolica* for the expression of the 65 kilodalton isoform of human glutamate decarboxylase. Yeast. 1998;14:11-23. DOI 10.1002/(SICI)1097-0061(19980115)14:1<11::AID-YEA196>3.0.CO;2-S.

Roongsawang N., Puseenam A., Kitikhun S., Sae-Tang K., Harnpicharnchai P., Ohashi T., Fujiyama K., Tirasophon W., Tanapongpipat S. A novel potential signal peptide sequence and overexpression of ER-resident chaperones enhance heterologous protein secretion in thermotolerant methylotrophic yeast *Ogataea thermomethanolica*. Appl. Biochem. Biotechnol. 2016;178:710-724. DOI 10.1007/s12010-015-1904-8.

Sakai Y., Rogi T., Takeuchi R., Kato N., Tani Y. Expression of *Saccharomyces* adenylate kinase gene in *Candida boidinii* under the regulation of its alcohol oxidase promoter. Appl. Microbiol. Biotechnol. 1995;42:860-864. DOI 10.1007/BF00191182.

Seo H.S., Park J.S., Han K.Y., Bae K.D., Ahn S.J., Kang H.A., Lee J. Analysis and characterization of hepatitis B vaccine particles synthesized from *Hansenula polymorpha*. Vaccine. 2008;26:4138-4144. DOI 10.1016/j.vaccine.2008.05.070.

Sibirny A.A. *Pichia methanolica* (*Pichia pinus* MH_4). In: Wolf K. (Ed.). Nonconventional Yeasts in Biotechnology. Springer, 1996;277-291. DOI 10.1007/978-3-642-79856-6_8.

Sohn M.J., Oh D.B., Kim E.J., Cheon S.A., Kwon O., Kim J.Y., Lee S.Y., Kang H.A.H. *HpYPS1* and *HpYPS7* encode functional aspartyl proteases localized at the cell surface in the thermotolerant methylotrophic yeast *Hansenula polymorpha*. Yeast. 2012;29:1-16. DOI 10.1002/yea.1912.

Spencer D.M., Spencer J.F.T., Fengler E., de Figueroa L.I. Yeasts associated with algarrobo trees (*Prosopis* spp.) in northwest Argentina: a preliminary report. J. Ind. Microbiol. 1995;14:472-474. DOI 10.1007/BF01573960.

Spencer D.M., Spencer J.F.T., de Figueroa L.I., Garro O., Fengler E. Yeasts associated with pods and exudates of algarrobo trees (*Prosopis* spp.) and species of columnar cacti in northwest Argentina. Appl. Microbiol. Biotechnol. 1996;44(6):736-739. DOI 10.1007/BF00178611.

Spencer J.F.T., Spencer D.M. Ecology: where yeasts live. In: Spencer J.F.T., Spencer D.M. (Eds.). Yeasts in Natural and Artificial Habitats. Springer, 1997;33-58. DOI 10.1007/978-3-662-03370-8_4.

Syromyatnikov M.Y., Kiryanova S.V., Popov V.N. Development and validation of a TaqMan RT-PCR method for identification of mayonnaise spoilage yeast *Pichia kudriavzevii*. AMB Express. 2018;8(186). DOI 10.1186/s13568-018-0716-y.

Tanapongpipat S., Promdonkoy P., Watanabe T., Tirasophon W., Roongsawang N., Chiba Y., Eurwilaichitr L. Heterologous protein expression in *Pichia thermomethanolica* BCC16875, a thermotolerant methylotrophic yeast and characterization of N-linked glycosylation in secreted protein. FEMS Microbiol. Lett. 2012;334:127-134. DOI 10.1111/j.1574-6968.2012.02628.x.

Trifinopoulos J., Nguyen L.T., von Haeseler A., Minh B.Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 2016;(44):W232-W235. DOI 10.1093/nar/gkw256.

Trotsenko Yu.A., Torgonskaya M.L. Methylotrophic Yeasts. Moscow: TR-Print Publ., 2011. (in Russian)

Trotsenko Yu.A., Torgonskaya M.L. Aerobic methylotrophs – promising objects of modern biotechnology. J. Siberian Federal University. Biology. 2012;5(3):243-279. (in Russian)

van der Klei I.J., Yurimoto H., Sakai Y., Veenhuis M. The significance of peroxisomes in methanol metabolism in methylotrophic yeast. Biochim. Biophys. Acta. 2006;1763(12):1453-1462. DOI 10.1016/j.bbamcr.2006.07.016.

Wegner E.H. Biochemical Conversions by Yeast Fermentation at High Cell Densites. U.S. Pat. 4,414,329. Oct. 29, 1981.

Wegner G.H., Harder W. Methylotrophic Yeasts – 1986. In: van Verseveld H.W., Duine J.A. (Eds.). Microbial Growth on C_1 Compounds: Proc. of the 5th Int. Symp. Dordrecht: Springer Netherlands, 1987;131-138. DOI 10.1007/978-94-009-3539-6_17.

Weydemann U., Keup P., Piontek M., Strasser A.W.M., Schweden J., Gellissen G., Janowicz Z.A. High-level secretion of hirudin by *Hansenula polymorpha* – authentic processing of three different preprohirudins. Appl. Microbiol. Biotechnol. 1995;4:377-385. DOI 10.1007/BF00169932.

Yamada Y., Maeda K., Mikata K. The phylogenetic relationships of the hat-shaped ascospore-forming, nitrate-assimilating *Pichia* species, formerly classified in the genus *Hansenula* Sydow et Sydow, based on the partial sequences of 18S and 26S ribosomal RNAs (Saccharomycetaceae): the proposals of three new genera, *Ogataea*, *Kuraishia*, and *Nakazawaea.* Biosci. Biotechnol. Biochem. 1994; 58(7):1245-1257. DOI 10.1271/bbb.58.1245.

Yamada Y., Matsuda M., Maeda K., Mikata K. The phylogenetic relationships of methanol-assimilating yeasts based on the partial sequences of 18S and 26S ribosomal RNAs: the proposal of *Komagataella* gen. nov. (Saccharomycetaceae). Biosci. Biotechnol. Biochem. 1995;59(3):439-444. DOI 10.1271/bbb.59.439.

Yurimoto H., Hasegawa T., Sakai Y., Kato. N. Characterization and high-level production of D-amino acid oxidase in *Candida boidinii*. Biosci. Biotechnol. Biochem. 2001;65(3):627-633. DOI 10.1271/bbb.65.627.

Yurimoto H., Kato N., Sakai Y. Assimilation, dissimilation, and detoxification of formaldehyde, a central metabolic intermediate of methylotrophic metabolism. Chem. Rec. 2005;5:367-375. DOI 10.1002/tcr.20056.

Yurimoto H., Komeda T., Lim C.R., Nakagawa T., Kondo K., Kato N., Sakai Y. Regulation and evaluation of five methanol-inducible promoters in the methylotrophic yeast *Candida boidinii*. Biochim. Biophys. Acta. 2000;1493:56-63. DOI 10.1016/S0167-4781(00)00157-3.

Yurimoto H., Sakai Y. Methanol-inducible gene expression and heterologous protein production in the methylotrophic yeast *Candida boidinii*. Biotechnol. Appl. Biochem. 2009;53:85-92. DOI 10.1042/ba20090030.

Yurimoto H., Yamane M., Kikuchi Y., Matsui H., Kato N., Sakai Y. The pro-peptide of *Streptomyces mobaraensis* transglutaminase functions in *cis* and in *trans* to mediate efficient secretion of active enzyme from methylotrophic yeasts. Biosci. Biotechnol. Biochem. 2004;68(10):2058-2069. DOI 10.1271/bbb.68.2058.

Zhang T., Gong F., Chi Z., Liu G., Chi Z., Sheng J., Li J., Wang X. Cloning and characterization of the inulinase gene from a marine yeast *Pichia guilliermondii* and its expression in *Pichia pastoris*. Antonie van Leeuwenhoek. 2009;95:13-22. DOI 10.1007/s10482-008-9281-8.

ORCID ID

E.G. Pershina orcid.org/0000-0003-2658-7906
S.E. Peltek orcid.org/0000-0002-3524-0456

Acknowledgements. This research is supported by the publically funded project No. 0259-2019-0005 of the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences.

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

Recieved October 08, 2019. Revised November 15, 2019. Accepted December 19, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):158-167
DOI 10.18699/VJ19.38-o

Perspectives of using Illumina MiSeq for identification of arbuscular mycorrhizal fungi

A.A. Kryukov[1] ✉, A.O. Gorbunova[1, 2], E.M. Machs[3], Y.V. Mikhaylova[3], A.V. Rodionov[2, 3], P.M. Zhurbenko[3], A.P. Yurkov[1]

[1] All-Russian Research Institute for Agricultural Microbiology, St. Petersburg, Russia
[2] Saint Petersburg State University, Biological Faculty, St. Petersburg, Russia
[3] Komarov Botanical Institute of the Russian Academy of Sciences, St. Petersburg, Russia
✉ e-mail: rainniar@rambler.ru

Abstract. Arbuscular mycorrhiza fungi (AMF) form one of the most common symbiosis with the majority of land plants. AMF supply the plant with various mineral elements, primarily phosphorus, and improve the water supply. The search for the most effective AMF strains for symbiosis and the creation of microbial preparations on that basis is an important task for modern biology. Owing to the difficulties of cultivation without a host plant and their high genetic polymorphism, identifying AMF is very difficult. A high number of cryptic species often makes morphological identification unreliable. Recent years have seen a growth in the number of AMF biodiversity studies performed by modern NGS-based methods, Illumina MiSeq in particular. Currently, there are still many questions that remain for the identification of AMF. The most important are whether conservative or variable sequences should be used to select a marker for barcoding and whether universal primers or those specific to AMF should be used. In our work, we have successfully used universal primers ITS3 and ITS4 for the sequencing in Illumina MiSeq of the 5.8S rDNA – ITS2 region of the 35S rRNA genes, which contain both a conservative and variable regions. The molecular genetic approach for AMF identification was quite effective and allowed us to reliably identify eight of nine isolates to the species level: five isolates of *Rhizophagus irregularis*, and one isolate of *R. invermaius*, *Paraglomus laccatum*, and *Claroideoglomus etunicatum*, respectively. For all five *R. irregularis* isolates, high variability in the ITS region and the absence of ecotopic-related molecular characters in the ITS2 region were demonstrated. The NCBI data is still insufficient for accurate AMF identification of *Acaulospora* sp. isolates from the genus to the species level.

Key words: Glomeromycotina; arbuscular mycorrhiza; Illumina MiSeq; *Rhizophagus irregularis*; *R. invermaius*; *Paraglomus laccatum*; *Claroideoglomus etunicatum*; *Acaulospora*.

For citation: Kryukov A.A., Gorbunova A.O., Machs E.M., Mikhaylova Y.V., Rodionov A.V., Zhurbenko P.M., Yurkov A.P. Perspectives of using Illumina MiSeq for identification of arbuscular mycorrhizal fungi. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):158-167. DOI 10.18699/VJ19.38-o

Перспективы использования Illumina MiSeq для идентификации грибов арбускулярной микоризы

А.А. Крюков[1] ✉, А.О. Горбунова[1, 2], Э.М. Мачс[3], Ю.В. Михайлова[3], А.В. Родионов[2, 3], П.М. Журбенко[3], А.П. Юрков[1]

[1] Всероссийский научно исследовательский институт сельскохозяйственной микробиологии, Санкт-Петербург, Россия
[2] Санкт-Петербургский государственный университет, биологический факультет, Санкт-Петербург, Россия
[3] Ботанический институт им. В.Л. Комарова Российской академии наук, Санкт-Петербург, Россия
✉ e-mail: rainniar@rambler.ru

Аннотация. Грибы арбускулярной микоризы (АМГ) образуют один из наиболее распространенных симбиозов с большинством наземных растений. АМГ снабжают растение различными минеральными элементами, в первую очередь, фосфором, а также улучшают водоснабжение. Поиск наиболее симбиотически эффективных штаммов АМГ и создание на их основе микробных препаратов – важная задача современной биологии. Идентификация АМГ очень сложна. Это связано, прежде всего, с высоким генетическим полиморфизмом АМГ, а также с трудностями их выращивания без растения-хозяина. Морфологическая идентификация АМГ часто ненадежна из-за большого числа криптических видов. В последние годы увеличивается число работ по изучению биологического разнообразия АМГ, проводимых современными методами на основе NGS (Next Generation Sequencing), в частности Illumina MiSeq. В настоящее время остается много вопросов по идентификации АМГ. К наиболее важным из них относятся: выбор маркера для генетического штрих-кодирования АМГ – консервативных или вариабельных последовательностей, а также выбор праймеров – специфичных для АМГ или универсальных. В настоящей работе мы успешно использовали универсальные праймеры ITS3 и ITS4 для секвенирования с Illumina MiSeq региона 5.8S рДНК – ITS2, содержащего как консервативные, так и вариабельные участки. Этот подход для идентификации АМГ оказался достаточно эффективным и позволил достоверно идентифицировать 8 из 9 изолятов до уровня вида: 5 – *Rhizophagus irregularis*, 1 – *R. invermaius*, 1 – *Paraglomus laccatum*, 1 – *Claroideoglomus etunicatum*. Для всех

изолятов *R. irregularis* показаны высокая вариабельность в области ITS и отсутствие связи с экотопом клад, образуемых ITS на филогенетических деревьях. Для изолята *Acaulospora* sp., определенного до рода, данные NCBI все еще недостаточны для точной идентификации АМГ рода *Acaulospora* до вида.
Ключевые слова: Glomeromycotina; арбускулярная микориза; Illumina MiSeq; *Rhizophagus irregularis*; *R. invermaius*; *Paraglomus laccatum*; *Claroideoglomus etunicatum*; *Acaulospora*.

Introduction

Arbuscular mycorrhizal fungi (AMF) belonging to the subdivision Glomeromycotina are a relatively small yet diverse group. Various estimations report from 240 (Stockinger et al., 2014) to 348 species (Öpik et al., 2014). At the same time, AMF form mycorrhizal relationships with more than 200,000 species of land plants (Lee et al., 2013). Identification the AMF strains with the highest symbiosis efficiency is of great value for agricultural applications. To do it, we isolated, identified (Kryukov et al., 2017; Kryukov, Yurkov, 2018; Yurkov et al., 2018b), and evaluated the symbiotic efficiency (Yurkov et al., 2017a, b; Yurkov et al., 2018a) for a number of strains from the collection.

AMF are traditionally identified by more than 20 morphological characters (Schenk, Perez, 1990; Blaszkowski, 2019; INVAM, 2019; Schüßler, 2019). However, in some cases morphological identification fails to discriminate closely related species. A number of AMF species are morphologically indistinguishable. Furthermore, interspecific genetic polymorphism can sometimes be mistakenly treated as intraspecific polymorphism (Savary et al., 2017). Everything mentioned above may result in wrong estimations of AMF species. AMF usually cannot be cultivated on artificial media, thus making the identification very hard. As a result many AMF species are not studied at all, and so have not received the proper morphological description essential to their identification (Bruns et al., 2017). The type genus *Glomus* is a striking example of the taxonomic problems concerning species and genera identification. In the last twenty years, *Glomus* taxon has been revised a number of times, so that many Glomus species put under 14 other genera instead: *Ambispora*, *Claroideoglomus*, *Corymbiglomus*, *Diversispora*, *Dominikia*, *Entrophospora*, *Funneliformis*, *Kamienskia*, *Pacispora*, *Paraglomus*, *Redeckera*, *Rhizophagus*, *Sclerocystis*, *Septoglomus* (Schüßler, 2019). In this key, molecular-genetic identification of Glomeromycotina is very significant as the exponential increase in the number of AMF DNA sequences deposited in databases over the last ten years attests (NCBI, 2018).

The forementioned problems make the choice of the best method to identify Glomeromycotina genera and species very topical (Kryukov et al., 2017; Kryukov, Yurkov, 2018; Yurkov et al., 2018a).

The study of Glomeromycotina by molecular genetic methods is associated with a number of difficulties. One of the main problems is the high degree of genetic polymorphism, including polymorphism at the intraspecific level in the SSU–ITS1–5.8SrRNA–ITS2–LSU region commonly used for molecular-genetic identification (Stockinger et al., 2010; Yurkov et al., 2018a). The reasons for this variability may be: 1) the ability of AMF to form anastomoses and exchange genetic material (Daubois et al., 2016); 2) the formation of a very large number of nuclei – from 576 to 35,000 in one spore (Hosny et al., 1998). Copies of DNA markers independently evolve in different nuclei (Lin et al., 2014).

The ITS region is the main marker for AMF barcoding, it is well represented in different databases, among which are the UNITE (User-friendly Nordic ITS Ectomycorrhiza) (UNITE, 2017) and MaarjAM (Öpik et al., 2010). It should be noted that the ITS is also often used for the barcoding of vascular plants and the construction of phylogeny (Rodionov et al., 2016). But the main advantage of the ITS region is the possibility of identifying AMF up to the species level. A majority of papers where less variable markers have been used report on identification up to the genus level or even to the order level (Schoch et al., 2012; Tedersoo et al., 2015).

In the case of Illumina MiSeq it is recommended that ITS2 or the full ITS region be used as the fungal barcode (Tedersoo et al., 2015). ITS2 provides a higher taxonomic resolution than SSU or LSU genes, which are suitable for identifying genera and higher-level taxa. ITS1 in fungi is usually shorter that ITS2, also ITS1 is considered as a hypervariable region and thus less suitable for barcoding fungi (Tedersoo et al., 2015). To increase efficiency of molecular-genetic identification, modifications of the universal primers ITS3 and ITS4 specific for various fungal divisions were proposed. In comparison with other primer combinations, the primer pair ITS3tagmix and ITS4ngs gave a significantly larger number of sequencing reads as well as OTUs (Operational Taxonomic Units) (Tedersoo et al., 2015). At the same time, the contribution of AMF in the total OTUs pool was less than 3 %, whereas Agaricomycetes represented half of all OTUs. For AMF we proposed using the slightly modified primer ITS3 – CATC GATGAAGAACG***T***AG (the modification is in bold) as a direct primer and the primer ITS4 without changes as reverse.

Various aspects of using specific primers for the AMF identification were reviewed earlier (Kryukov et al., 2017; Kryukov, Yurkov, 2018; Yurkov et al., 2018a). Identification using universal primers allows us to investigate the maximal broad range of AMF species and genera. But this then introduces the problem of foreign DNA admixture. Another problem with molecular genetic identification of AMF is the generation of chimeric sequences during the sequencing process (Senés-Guerrero et al., 2014). Phusion DNA-polymerase, which generates high accuracy PCR-products, serves to reduce the chances of introducing chimera (Senés-Guerrero et al., 2014). At the same time, special software has been developed for detecting chimeric sequences and excluding them from analysis, USEARCH for example (Edgar, 2010).

Before 2015, the main method of AMF molecular-genetic identification was cloning followed by Sanger sequencing. This method applied to AMF demands careful selection of efficient and highly specific primers, as well as the use of nested PCR. In addition, due to the high variability of marker sequences, only multiple sequence cloning can be used (Krüger et al., 2009), which is very time-consuming and labor-intensive.

NGS has turned into a powerful and attractive method of AM fungi identification since it can overcome the weak points in the Sanger-based identification. One of the earli-

est techniques of NGS was 454 pyrosequencing, which has been employed for AMF identification since 2009. Using the universal fungal primers NS31 and AM1 179,279 sequences were obtained, of which 77.5 % belonged to 47 taxa of AMF isolated from the roots of 10 plant species (Öpik et al., 2009). However, these primers (NS31 and AM1) are not suitable for the analysis of the SSU region in Archaeosporaceae and Paraglomeraceae families (Helgason et al., 1998). Nonetheless, 454 pyrosequencing revealed in the roots of the *Hepatica nobilis* Mill. 1.5 times more fungal taxa than Sanger sequencing (Öpik et al., 2009). This clearly showed the advantage of NGS methods over cloning-Sanger sequencing methods. 454 pyrosequencing following nested PCR was successfully used for AMF identification with a barcode in the LSU region (Senés-Guerrero, Schüßler, 2015). An interesting result of this work is that about 60 % of the studied plants each formed a symbiosis with at least 10 AMF taxa, and 2 % of plants had more than 25 AMF species in their root system. The authors used the LSU-D1f modified primer (Senés-Guerrero, Schüßler, 2015) together with the LSUmBr primer (Krüger et al., 2009) in the second round of nested PCR. This modification allowed us to obtain 698,297 sequences, of which 0.17 % were the target AMF sequences, 41 taxa were detected, of which 15 are unknown, not registered in the databases (Senés-Guerrero, Schüßler, 2015). 454 pyrosequencing has its advantages over cloning-Sanger sequencing, but this technique is more expensive than the Illumina MiSeq technology that replaced pyrosequencing.

With the development of Illumina technology, Illumina MiSeq has been becoming more and more widely used due to its relative low cost. Illumina MiSeq compared to HiSeq 2000 allows for processing sequences of reads 2.5 times longer, and each sequencing in this case is cheaper. The advantage of HiSeq is more deep sequencing, which allows us to obtain reads at a rate of three orders greater than MiSeq. But this is less significant for fungi identification than sequence read length (Razzauti et al., 2015).

A comparative study of the efficiency of 454 pyrosequencing and Illumina MiSeq showed a difference of five times in the diversity of sequences (in favor of the second method), but both approaches revealed the same species composition (Vasar et al., 2017). A comparative study of the Illumina MiSeq and Ion Torrent Personal Genome Machine (PGM) showed that the second method generated a 2–5 fold greater rate of error than Illumina MiSeq (Salipante et al., 2014).

The objectives of this study included the identification of nine strains of AM fungi from the collection of the All Russian Research Institute for Agricultural Microbiology using the Illumina MiSeq approach and universal primers for the ITS2 region, supplemented by the morphological characteristics of the analyzed AMF spores.

Materials and methods

AMF isolation. AMF were isolated in 2015 from samples, collected in various habitats in two different regions (author of the analyzed AMF isolates: A.P. Yurkov). Only isolates with spores with an unambiguously identifiable morphology were used for molecular-genetic identification. Four strains were isolated from samples taken in the Rostov region (1.4 km NW from Zernograd): 46°52′2″ N, 40°16′8″ E, a tree belt area with oaks, maples and alders: 1) isolate (strain) 01-053 was isolated from *Ambrosia artemisiifolia* roots, 2) isolate (strain) 01-056a and, 3) isolate 01-056b both were isolated from one ordinary chernozem soil sample, but differentiated via reinoculations of spores with a different morphology, 4) at 46°52′7″ N, 40°16′8″ E, a maize field, the isolate (strain) 02-060 was isolated from a *Zea mays* 282MB Zelenogradckij hybrid roots. Five isolates were obtained from samples taken in the Moscow region (Lobnya town, academic village in Lugovaya): 1) 56°02′33.80″ N, 37°29′13.70″ E, natural meadow, the isolate (strain) 03-097 was isolated from *Vicia sepium* roots, 2) 56°02′24.30″ N, 37°29′20.00″ E, a *Festuca rubra* field, isolate (strain) 04-067 was isolated from *Festuca rubra* roots, 3) isolate (strain) 04-068 was isolated from *Agrostis vulgaris* roots, 4) 56°02′31.40″ N, 37°29′17.60″ E, *Medicago* × *varia* field, isolate (strain) 05-077 was isolated from *Trifolium pratense* roots, 5) isolate 05-104 was isolated from a sod-podsol gleyic soil sample.

AMF cultivation. AMF collection at the All-Russian Research Institute for Agricultural Microbiology is cultivated in the *Plectranthus australis* R. Br. (taxonomical synonyms *P. verticillatus* (L.f.) Druce, *P. nummularius* Briq.) line of Swedish ivy. For AMF-inoculated plants a soil-sand mix was used, described earlier (Yurkov et al., 2015). The substrate had a low level of plant-available phosphorus (3.0 mg P_2O_5/100 g). *Plectranthus* cuttings 12–15 cm length with two leaves were sterilized in 0.1 % sodium hypochlorite, then germinated in water. On the 7th day *Plectranthus* plants were inoculated by root fragments, containing AMF vesicles and abruscules. Root fragments of mycorrhizal *Plectranthus* were selected by visual analysis in stereomicroscope MBS-10 (LZOS, Russia). For further reinoculations (each 6–8 months) sporocarps, or an arrangement of spores in an out-root area, or 5 mm length root fragments with large quantities of observable vesicles were used. Each inoculated plant was cultivated in an individual container with 350 g of sterile substrate at +24–26 °C, 18 h light day. Two luminescent lamp LB-40 (Russia) and OSRAM L36/77 Fluora (Germany) on a 1:1 ratio with output lumen ~4000 were used. Plants were watered every other day by 60 % of soil full water capacity. For the culture purification AMF spores were reinoculated at least three times.

Morphological identification of AMF. More that 20 features were used for morphological identification (Schenck, Pérez, 1990; Blaszkowski, 2003; INVAM, 2019): color, transparency, size and shape of extraradical (out-root) spores; shape of the place of attachment of spores to subtending hypha; number, thickness, density, elasticity or fragility, color in Melzer's reagent of layers of spore walls and subtending hypha; presence/absence/disappearance/appearance of spore wall layers and subtending hypha during ontogenesis (from juvenile to mature spore); presence/absence of a septum in the place of attachment of the spores to subtending hypha; structural characteristics of AMF and intraradical spores.

In order to assess mycorrhization parameters and the type of mycorrhiza, the roots were macerated and stained by trypan blue according to the method developed by J.M. Phillips and D.S. Hayman (1970). Mycorrhization parameters were determined by light microscopy (Trouvelot et al., 1986) using a special computer program, developed earlier in our research group (Vorob'ev et al., 2016).

DNA extraction, PCR and sequencing. DNA extractions were carried out using the method of J.J. Doyle and J.L. Doyle (1987), with modifications. Micorrhizated roots of *P. australis* were washed twice in distilled water, placed in 2 ml tubes, dried at +45 °C, and mechanically homogenized with glass beads 2–4 mm in the FastPrep24 (MP Biomedicals, USA), followed by CTAB-protocol. The target region ITS2 was amplified with universal primers ITS3 (5′-GCATCGATGAAGAACG CAGC-3′) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3′) (White et al., 1990). Ready-mix ScreenMix (Evrogen, Russia) was used for PCR. Amplicons were sliced from agarose gel and purified by the silica approach. Illumina library preparation was made according to MiSeq Reagent Kit Preparation Guide in the Core Center of "Genomic Technologies, Proteomics and Cell Biology" at the All-Russia Research Institute for Agricultural Microbiology (St. Petersburg, Russia). Libraries were sequenced on the Illumina MiSeq platform with MiSeq® Reagent Kit v3 (600 cycle) according to the manufacturer's instruction (Illumina Inc., USA).

Bioinformatics and fungal OTUs analysis. Two pipelines were used for sequencing data analysis. 1. Illumina reads processing were done by USEARCH software (Edgar, 2010). The key steps for USEARCH data treatment are described on https://www.drive5.com/usearch/. In further paragraphs we are briefly describing these steps. In the case of paired-end sequencing, Illumina sequencer makes sequences from both ends of a fragment and generates two files with forward and reverse reads. This data is written in FASTQ format, where each nucleotide corresponds with its quality score. For further treatment forward and reverse reads are merging using fastq_mergepairs command to give consensus sequences. This step includes resolving any mismatches found in the overlap alignments and calculation the posterior quality scores for the consensus sequences (Edgar, Flyvbjerg, 2014). To discard low-quality reads expected error filtering with fastq_filter command is used.

The next step is a dereplication, which means sorting unique sequences in order of decreasing abundance in the dataset. After this, singletons (sequences that are present exactly once) are discarded, since they are likely to have errors. However the remaining reads can still have errors. So the goal of the final step (denoising) is to identify a set of correct biological sequences. The denoising can be made by UPARSE algorithm which clusters sequences with 97 % or more (Edgar, 2018) similarity and then chooses the most abundance sequence in each cluster. Also chimeric sequences, which occurs by combining parts of two or more biological sequences are detected and deleted at this step.

Raw forward and reverse reads were merged with minimal length parameter ("-fastq_minmergelen") 130 bp and maximal difference parameter ("-fastq_maxdiffs") 30 bp. Then low-quality read ends including primer sequences were trimmed, reads were filtered based on expected error value (E_max = 1). Singletons were removed from the dataset. Then data were divided on operational taxonomic units (OTUs) with a 97 % similarity cut-off by UPARSE algorithm (Edgar, 2013). Chimeric sequences were removed. For further analysis the most represented sequence from each OTU was chosen. Data were checked for cross-talk errors. Sequences of AMF species were selected by BLAST+ (Altschul et al., 1990).

2. The primers we used are universal for a broad range of species; and after amplification in the ITS2 region the extracted plant DNA prevail over fungi DNA. A second pipeline was made for the selection of rare and unique reads with a high homology of AMF sequences. After quality control (FastQC) forward and reverse reads were trimmed and merged with minimal length parameter 230 bp by trimmomatic (Bolger et al., 2014) and fastq-join software (Aronesty, 2013). Then sequences were demultiplicated and sorted in the descending order of their frequency. AMF sequences were selected via character for AMF motifs, then aligned and checked via BLAST.

Obtained sequences were submitted to the GenBank database (https://www.ncbi.nlm.nih.gov/). Evolutionary analyses were conducted by using the Maximum Likelihood method in MEGA7 software (Kumar et al., 2016) with implementation of the Tamura-Nei model (Tamura, Nei, 1993) and 1,000 bootstrap analyses.

Results

Sequencing of 9 isolates yielded approximately 381,249 pair of reads, from 19,236 to 81,054 joined sequences for each isolate. The following OTUs of AMF isolates were identified via BLAST at the genera or species level and submitted to NCBI: MK948362-MK948371 (isolate number 01-053), MK948403-MK948404 (01-056a), MK968150 (01-056b), MK948427-MK948429 (02-060), MK948492-MK948496 (03-097), MK948434-MK948436 (04-067), MK948447 (04-068), MK948486-MK948491 (05-077), MK948503-MK948504 (05-104). The length of obtained sequences varied from 340 to 366 bp. Variability in GC content was distinct in different genera: a narrow range was shown in genera *Claroideoglomus* (38–39 %), *Rhizophagus* (36–39 %) and *Paraglomus* (42–46 %), whereas GC content in *Acaulospora* varied from 31 to 46 %. Owing to the significant variability of the ITS region in AM, alignment of sequences belonging to different genera and orders is ineffective, and in some cases impossible. As a result, four separate phylogenetic trees were constructed for the four genera mentioned above (Fig. 1–4).

Unusual deletion was determined by ITS2 sequence alignment. This deletion of 5–6 bp in alignment positions 97–102 was identified in various species of genus *Rhizophagus*, and in all OTUs of genus *Paraglomus*, and in *Racocetra weresubiae* (FR750135) (NCBI, 2019). This fact can serve to indicate the presence of a deletion-specific site related to secondary RNA structure. However, it can also indicate a relationship of sequences in this region, which is of greater interest because it is well known that AM mycelia contain a significant number of nuclei (Hosny et al., 1998), some of which carry this deletion in ITS.

Some samples (01-053, 01-056a, 02-060, 04-068, 05-077, and 03-097) demonstrated a high similarity with *Rhizophagus* (see Fig. 1). OTU isolates 01-053, 01-056a, 02-060, 04-068, 05-077 fell with high accuracy into the clade formed by *Rhizophagus irregularis*, while OTUs of 03-097 were placed in one clade with *R. invermaius* (bootstrap index = 97). This genus includes a small number of species, about 20 according to A. Schüßler (2019), but in the NCBI only 8 species are represented by ITS sequences. Thus one may consider that *Rhizophagus* is a genus that requires further sequencing

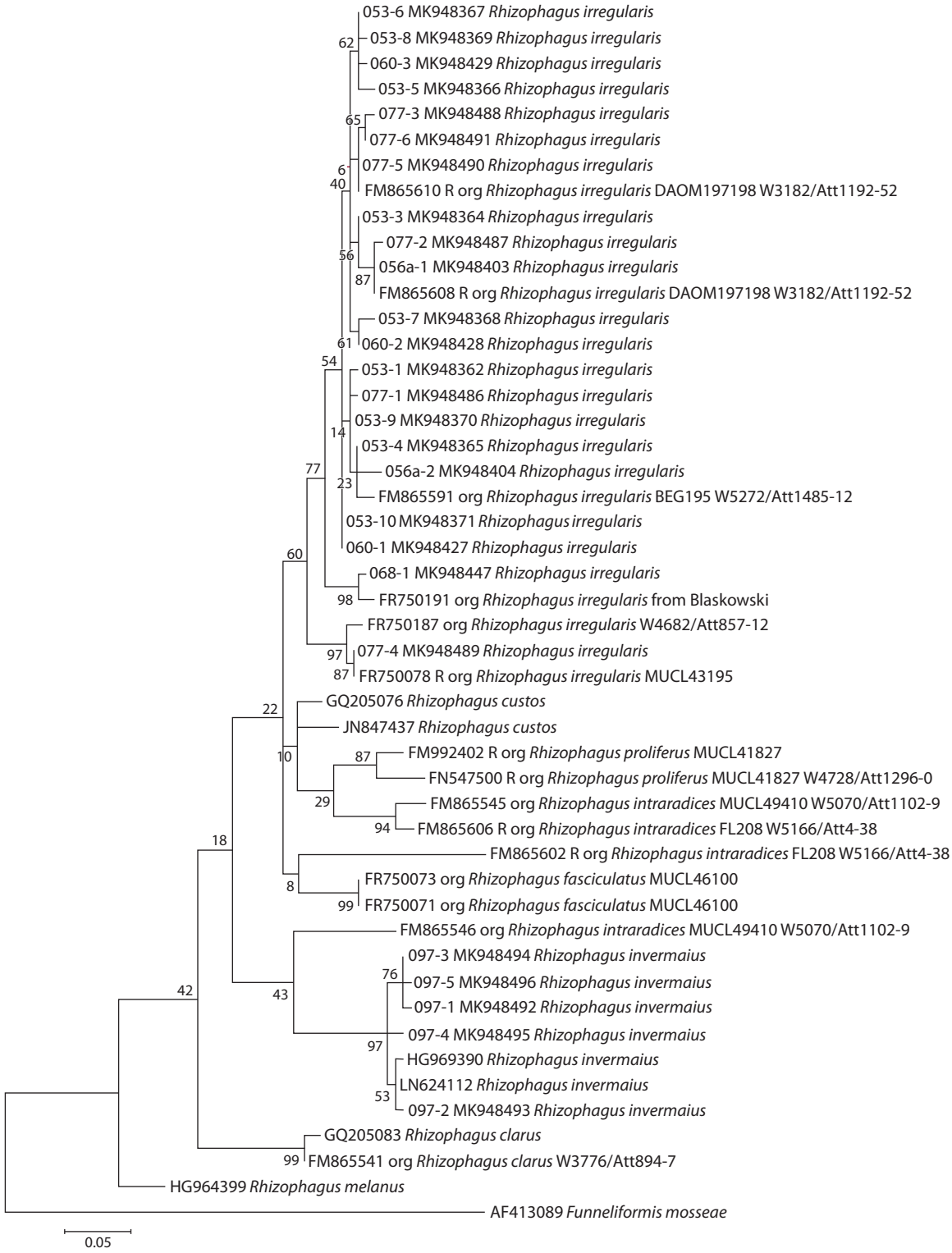


Fig. 1. Phylogenetic tree of ITS-sequences from the genus *Rhizophagus*.

research. At the same time, it is beyond doubt that this study provides a molecular genetic identification at the species level of isolates 01-053, 01-056a, 02-060, 04-068, 05-077, 03-097.

Due to the ambiguous position of OTUs of isolate 05-104 on a pre-built phylogenetic tree of the *Acaulospora* genus, as well as due to the possibility that OTUs of this isolate could incorporate other close genera from the Acaulosporaceae family, species from the *Archaeospora*, *Ambispora*, *Diversispora*, *Gigaspora*, *Racocetra*, *Redeckera* and *Scutellospora* genera were added to the tree. We found that both OTUs of 05-104 isolates were included in the *Acaulospora* clade with high bootstrap support (see Fig. 2).

However, owing to the significant differences from other species of *Acaulospora*, we define the MK948503 sequence

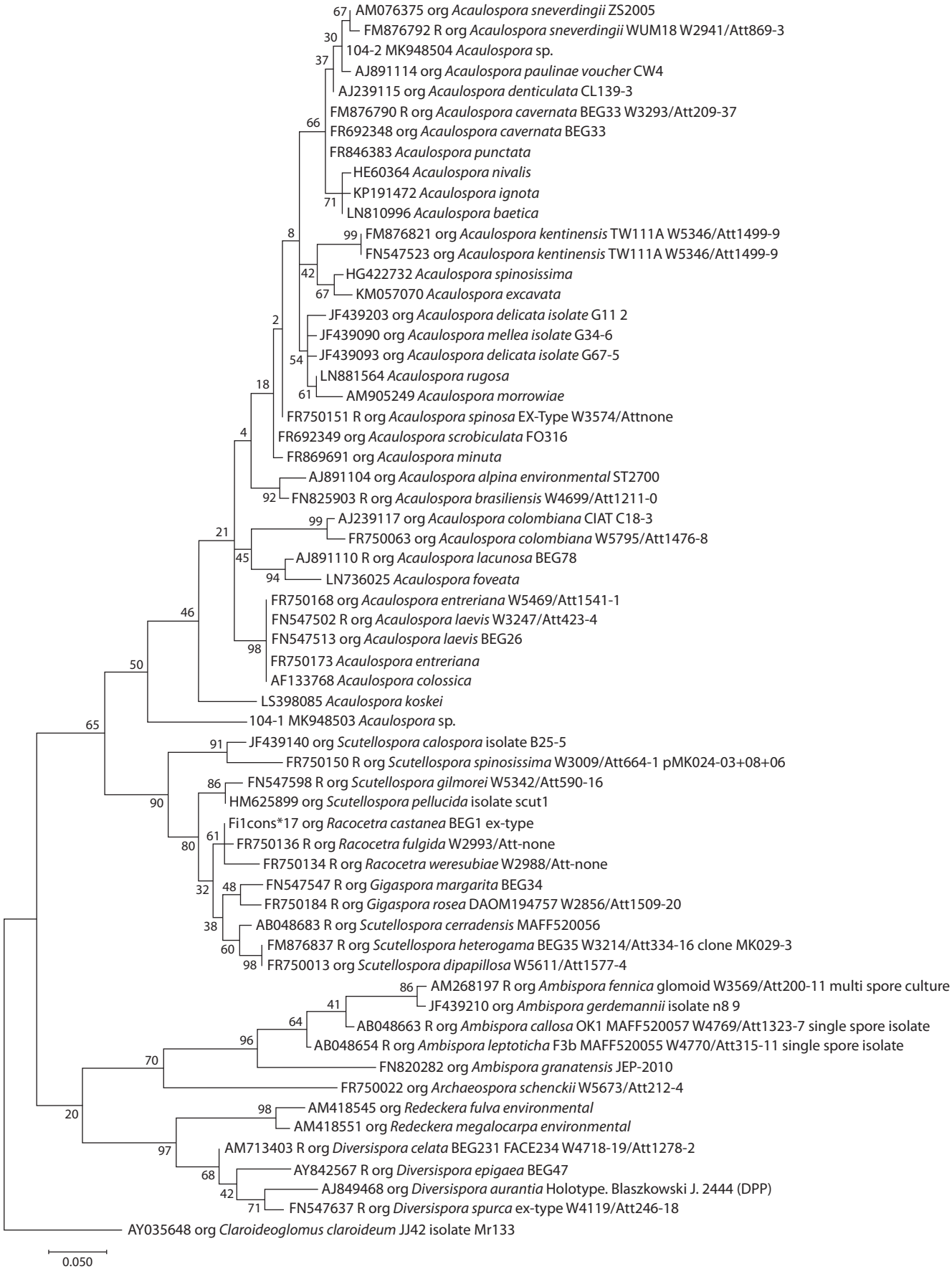


Fig. 2. Common phylogenetic tree of ITS-sequences from *Acaulospora*, *Archaeospora*, *Ambispora*, *Diversispora*, *Gigaspora*, *Racocetra*, *Redeckera* and *Scutellospora* genera.

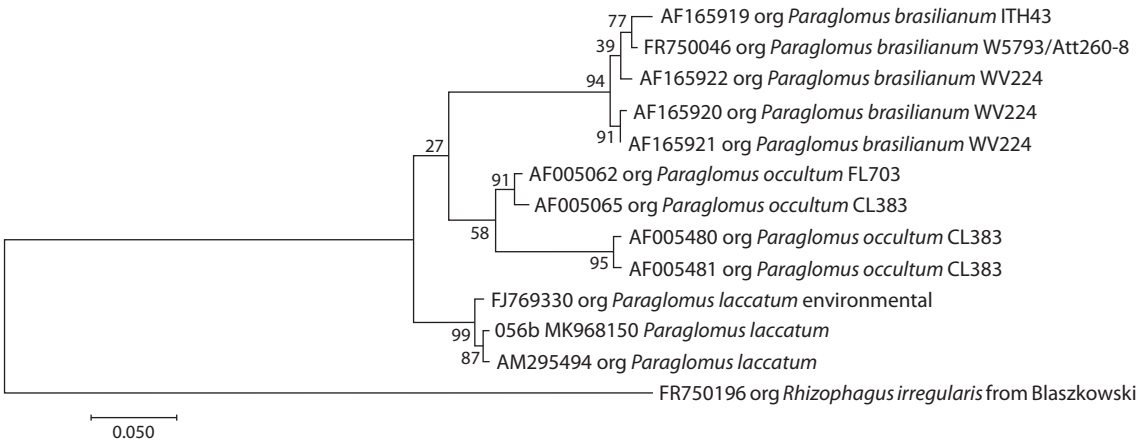


Fig. 3. Phylogenetic tree of ITS-sequences from *Paraglomus* genus.

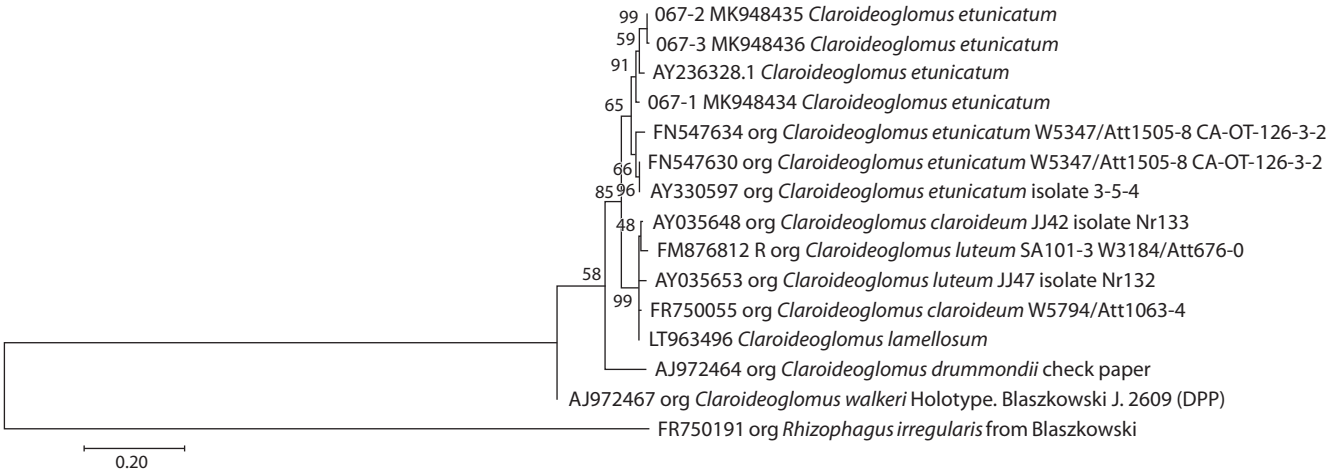


Fig. 4. Phylogenetic tree of ITS-sequences from *Claroideoglomus* genus.

(05-104 isolate) as a virtual taxon. The MK948504 sequence (05-104 isolate), though it has a high similarity with the *A. sieverdingii* and *A. paulinae* species, did not show similarities with the indicated species according to morphological data, therefore it was also identified as a virtual taxon. At the same time, it can be reliably stated that these two OTUs belong to different taxa, since they have significant differences in GC content (31 and 46 %). But differences did not appear in the case of morphological analysis, highlighting the obstacles of distinguishing among *Acaulospora* species.

The OTUs of 01-056b isolate formed a well-supported subclade in the *Paraglomus* genus (bootstrap index = 99) (see Fig. 3). Furthermore, the Paraglomerales order hosts two more genera (*Innospora* and *Pervetustus*) in addition to the *Paraglomus* genus, but the ITS sequences for them are unknown. Thus the outgroup for the phylogenetic tree was taken from another order. The 01-056b isolate was identified as *Paraglomus laccatum*. The sequences of the *P. laccatum* 01-056b isolate turned out to be the shortest in comparison with isolates of other genera and ITS. Only three of eight species of this genera have ITS sequences in NCBI (Schüßler, 2019).

The *Claroideoglomus* genus is the most studied of the above mentioned genera. The ITS data were reported for six out of eight species (Schüßler, 2019). The OTU of isolate 04-067 with high support belongs to the clade formed by the species *Claroideoglomus etunicatum* (see Fig. 4).

To verify whether the molecular genetic identification of AMF was effective, we conducted a morphological identification of nine isolates (Supplementary)[1]. The principal morphological characteristics for comparison were: size, shape, and color of the spores in air by the CMYK standard; shape, thickness and color of spore layers in Melzer's reagent by the CMYK standard; the shape and thickness of subtending hypha; the wall thickness of subtending hypha; the presence of a septum in subtending hypha, type of mycorrhiza in *P. australis*; mycorrhization parameter (*F* – frequency of mycorrhizal infection in *P. australis* roots).

On the basis of morphological analysis of spores, it was concluded that the AM isolates under examination – supported on *P. australis* culture in the ARRIAM collection – are of the following taxa: 1) 01-053 isolate – *Rhizophagus irregularis*; 2) 01-056a – *R. irregularis*; 3) 01-056b – *Paraglomus laccatum*; 4) 02-060 – *R. irregularis*; 5) 03-097 – *R. invermaius*; 6) 04-067 – *Claroideoglomus etunicatum*; 7) 04-068 – *R. irre-*

[1] Supplementary are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2020-24/appx0.pdf

gularis; 8) 05-077 – *R. irregularis*; 9) 05-104 – *Acaulospora* sp. defined only at the genus level, since it was not possible to clearly determine distinguishing morphological features.

Among the morphological features of the spores formed by various isolates of AMF the following should be highlighted. The spores of all AMF isolates classified as *R. irregularis* (isolates 01-053, 01-056a, 02-060, 04-068 and 05-077) have the cylindrical subtending hypha in oblong spores or slightly flared (expanded) in spherical spore at a distance of more than 10 μm from the spores. The size, shape and color of the spores in air in these isolates were similar. Only spores of isolates 01-053 and 01-056a with a greenish tint, had insignificant differences: color varied from 19-5-54-0 (pistachio) and 23-4-43-0 (very light yellow-green) to 26-17-66-0 (yellow-green) and 29-25-58-0 (dark khaki). It should be noted that the color of the L3 layer in spores of *R. irregularis* isolates had significant polymorphism: from 10-37-99-10 (corn-yellow), 1-45-91-25 (deep yellow), 10-53-98- 0 (yellow-orange), 13-54-99-1 (deep orange-yellow) and 26-52-97-2 (deep yellow) to 17-48-99-2 (almost pure orange), 15-62-98-0 (deep orange-yellow), 2-67-82-58 (brown-red), 30-80-99-30, 32-78-99-25 (saturated red-brown). The spore color of the 01-056b isolate, referred to the *Paraglomus laccatum* species, on the contrary varied in a narrow range: from colorless to 4-4-0-0 (ghostly white). The second difference in the *P. laccatum* 01-056b isolate was the absence of a septum in the subtending hypha in the presence of narrowing at the site of attachment of the spore as determined by the greater thickness of the L2 layer. The spores of isolate 04-067, referred to *Claroideoglomus etunicatum*, were more yellow: from 5-5-0-0 (ghostly white) to 2-10-62-5 (orange-yellow Craiol) in air and in Melzer's reagent L2 had a yellow color from 2-15-60-2 (orange-yellow Craiol) to 3-35-92-5 (saturated yellow). The darkest spores were observed in the isolate 03-097, referred to the *Rhizophagus invermaius* species: from 30-52-88-51 (deep yellow-brown) to 59-67-63-72 (brown-olive) in air. The spore morphology of isolate 05-104 was significantly different from the other AMF. The spore color of isolate 05-104 attributed to the *Acaulospora* genus, ranged from colorless (transparent) to 3-15-70-4 (yellow ivory), spores cycatrix (the scarring is a remnant of the connection between the spore wall and the wall of the spore of the saccule during spore synthesis is 5.5–7.1 μm. Since this isolate is poorly maintained in *P. australis* culture, the morphology of the spores could not be studied in detail. According to molecular genetic analysis data, it may be approximately the same probability as *A. paulinae*, *A. denticulata or A. sieverdingii*.

All isolates of the *Rhizophagus* genus (01-053, 01-056a, 02-060, 03-097, 04-068 and 05-077), collected both in the forest belt and arable land, had a higher activity in *P. australis* roots ($F > 80$ %) than isolates of other AMF genera (see Supplementary). All the AMF isolates under examination formed mycorrhiza of the Arum-type, but only seven of nine isolates were able to be isolated into culture: 01-053, 01-056a, 02-060, 03-097, 04-067, 04-068, 05-077 isolates should be considered as AMF strains.

Because their mycorrhization activity in *P. australis* is low (see Supplementary), sustained maintenance of the *P. laccatum* 01-056b isolate and the *Acaulospora* sp. 05-104 isolate will depend on the selection of optimal conditions for their growth (substrate composition and type of host plant).

Discussion

The Illumina MiSeq approach allowed us to determine eight out of nine AMF isolates from the All Russian Research Institute for Agricultural Microbiology collection. It is very powerful method, which enables the identification of a large number of AMF taxa in fungal communities, especially if the proportion of targeted marker sequences in a sample is small. There are other effective methods of NGS, but several comparative studies indicate that MiSeq in some cases is more efficient as it can provide longer reads and fewer errors in comparison to HiSeq and IonTorrent techniques, respectively (Salipante et al., 2014; Razzauti et al., 2015).

These comparisons show the clear advantage of the sequence-targeted NGS approach as contrasted with the alternative cloning-sequencing method for AMF species identification. However, the main disadvantage of the Illumina MiSeq are relatively short reads (250 bp ×2), which do not allow the use of long markers, such as the entire cloned SSU-ITS1-5.8S-ITS2-LSU region that was most often used for AMF barcoding from 2009 to 2012 (Krüger et al., 2009; De Castro et al., 2018). A prerequisite for the correct AMF identification by the MiSeq method is the employment of a short marker region for sequencing (400–500 bp). The most commonly used length-appropriate marker is ITS2. The advantage of this region is that it provides sufficient variability for identification at the species level. Other less variable regions such as D1–D2 of the LSU allow identification only at the genus level (Krüger et al., 2009).

The use of universal primers for ITS2 region is the optimal choice for the identification of poorly studied taxa, which in recent years has helped to identify a significant number of virtual taxa of AMF (Öpik et al., 2014). These taxa can subsequently receive species names in the presence of individual morphological features and stable maintenance of isolates/strains in culture. For example, we studied the 05-104 sample in which there are two OTUs that belong to two different virtual *Acaulospora* sp. taxa. Another advantage of using the ITS2 region for identification is that a substantially representative sequence database has been stored in NCBI in comparison to other marker regions, for example, for the SSU region (NCBI, 2018).

It is important to note that OTUs related to one *R. irregularis* species and collected in different ecotopes (isolates 01-053, 01-056a, 02-060, 04-068 and 05-077) did not cluster separately on the phylogenetic tree (see Fig. 1). This suggests that all studied *R. irregularis* isolates shared one ribotype. Thus, it was not possible to identify the ecotype-related features of *R. irregularis* isolates, perhaps they are objectively missing.

Also, it should be noted that NGS methods produce huge data arrays. We paid special attention to choose the correct tools for data treatment. Several pipelines have been developed to process rDNA sequences that are generated using Illumina's MiSeq platform. Among them mothur (Schloss et al., 2016), QIIME (Caporaso et al., 2010), USEARCH and VSEARCH (Rognes et al., 2016) are the most popular. Generally, all of them have similar steps in sequences treatment. We used USEARCH since it has a large set of tools and detailed documentation. However, two approaches were used to search for AMF OTUs among a great pool of sequences. Both approaches were equally effective since they allowed us to identify the

same set of taxa. At the same time the use of the USEARCH software has significantly reduced the time spent on data processing, which makes it possible to recommend it as the main tool to identify AMF from NGS data.

Combining the results of molecular genetics and morphological identification, we assert that the effectiveness of the Illumina MiSeq method as applied to AMF identification is not inferior to morphological methods that are significantly more labor intensive. However, the NCBI database is still insufficient for identification of some AMF species. The reason for this is that more than half of the known AMF taxa are still absent in the database. For example, according to A. Schüßler (2019), there are up to 56 species in the genus *Acaulospora.* The NCBI contains 37 species just for *Acaulospora* sp., and ITS sequences are presented for only 28 species. Thus, the obtained OTUs of the 05-104 isolate may belong to one of *Acaulospora* species that has yet to be studied.

Conclusion

We have determined that the most effective method for AMF identification is Illumina MiSeq supplemented by application of universal primers for the ITS2 region. Considerable efforts of morphologists in collaboration with molecular geneticists are required to establish a reliable taxonomy of the Glomeromycotina subdivision and to improve the efficiency of the molecular genetic AMF identification as a key method in the future.

References

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. J. Mol. Biol. 1990;215:403-410. DOI 10.1016/S0022-2836(05)80360-2.

Aronesty E. Comparison of sequencing utility programs. TOBioiJ. 2013;7:1-8. DOI 10.2174/1875036201307010001.

Blaszkowski J. Definitions of morphological characters of spores of arbuscular fungi. 2003. Available at http://www.zor.zut.edu.pl/Glomeromycota/Definitions%20of%20spore%20characters.htm

Blaszkowski J. Species descriptions and illustrations. 2019. Available at http://www.zor.zut.edu.pl/

Bolger A.M., Lohse M., Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics. 2014;30(15):2114-2120. DOI 10.1093/bioinformatics/btu170.

Bruns T.D., Corradi N., Redecker D., Taylor J.W., Öpik M. Glomeromycotina: what is a species and why should we care? New Phytol. 2017. DOI 10.1111/nph.14913. Available at http://onlinelibrary.wiley.com/doi/10.1111/nph.14913/full

Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Peña A.G., Goodrich J.K., Gordon J.I., Huttley G.A., Kelley S.T., Knights D., Koenig J.E., Ley R.E., Lozupone C.A., Mcdonald D., Muegge B.D., Pirrung M., Reeder J., Sevinsky J.R., Turnbaugh P.J., Walters W.A., Widmann J., Yatsunenko T., Zaneveld J., Knight R. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods. 2010;7(5):335-336. DOI 10.1038/nmeth.f.303.

Daubois L., Beaudet D., Hijri M., de la Providencia I. Independent mitochondrial and nuclear exchanges arising in *Rhizophagus irregularis* crossed-isolates support the presence of a mitochondrial segregation mechanism. BMC Microbiol. 2016;16(1):1-12. DOI 10.1186/s12866-016-0627-5.

De Castro O., Avino M., Di Maio A., Menale B., Guida M. Sanger and next generation sequencing in the characterisation of arbuscular mycorrhizal fungi (AMF) in *Pancratium maritimum* L. (Amaryllidaceae), a representative plant species of Mediterranean sand dunes. Planta. 2018;248(6):1443-1453. DOI 10.1007/s00425-018-2981-z.

Doyle J.J., Doyle J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 1987;19:11-15.

Edgar R.C. Search and clustering orders of magnitude faster than BLAST. Bioinformatics (Oxford, England). 2010;26(19):2460-2461. DOI 10.1093/bioinformatics/btq461.

Edgar R.C. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. Nat. Methods. 2013;10:996-998. DOI 10.1038/nmeth.2604.

Edgar R.C. Updating the 97 % identity threshold for 16S ribosomal RNA OTUs. Bioinformatics. 2018;34(14):2371-2375. DOI 10.1093/bioinformatics/bty113.

Edgar R.C., Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics. 2015;31(21):3476-3482. DOI 10.1093/bioinformatics/btv401.

Helgason T., Daniell T.J., Husband R., Fitter A.H., Young J.P.W. Ploughing up the wood-wide web? Nature. 1998;394(6692):431-431. DOI 10.1038/28764.

Hosny M., Gianinazzi-Pearson V., Dulieu H. Nuclear DNA content of 11 fungal species in Glomales. Genome. 1998;41(3):422-428. DOI 10.1139/g98-038.

INVAM. Species descriptions from reference cultures. 2019. Available at http://fungi.invam.wvu.edu/the-fungi/species-descriptions.html

Krüger M., Stockinger H., Krüger C., Schüßler A. DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi. New Phytol. 2009;183(1):212-223. DOI 10.1111/j.1469-8137.2009.02835.x.

Kryukov A.A., Yurkov A.P. Optimization procedures for molecular-genetic identification of arbuscular mycorrhizal fungi in symbiotic phase on the example of two closely kindred strains. Mikologiya i Fitopatologiya = Mycology and Phytopathology. 2018;52(1):38-48. (in Russian)

Kryukov A.A., Yurkov A.P., Shishova M.F. Features of the molecular phylogeny of arbuscular mycorrhizal fungi based on the analysis of the SSU – ITS1–5.8SrRNA – ITS2 – LSU marker region. Estestvennye i Tekhnicheskie Nauki = Natural and Technical Sciences. 2017;2:15-21. (in Russian)

Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2016;33:1870-1874. DOI 10.1093/molbev/msw054.

Lee E.H., Eo J.K., Ka K.H., Eom A.H. Diversity of arbuscular mycorrhizal fungi and their roles in ecosystems. Mycobiology. 2013;41(3):121-125. DOI 10.5941/MYCO.2013.41.3.121.

Lin K., Limpens E., Zhang Z., Ivanov S., Saunders D.G.O., Mu D., Pang E., Cao H., Cha H., Tao Lin T., Qian Zhou Q., Yi Shang Y., Ying Li Y., Trupti Sharma T., van Velzen R., de Ruijter N., Aanen D.K., Win J., Sophien Kamoun S., Ton Bisseling T., Geurts R., Huang S. Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. PLoS Genet. 2014;10(1):e1004078. DOI 10.1371/journal.pgen.1004078.

NCBI. National Center for Biotechnology Information. 2018. Available at https://www.ncbi.nlm.nih.gov

Öpik M., Davison J., Moora M., Zobel M. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. Botany. 2014;92(2):135-147. DOI 10.1139/cjb-2013-0110.

Öpik M., Metsis M., Daniell T.J., Zobel M., Moora M. Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. New Phytol. 2009;184(2):424-437. DOI 10.1111/j.1469-8137.2009.02920.x.

Öpik M., Vanatoa A., Vanatoa E., Moora M., Davison J., Kalwij J.M., Reier Ü., Zobel M. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). New Phytol. 2010;188(1):223-241. DOI 10.1111/j.1469-8137.2010.03334.x.

Phillips J.M., Hayman D.S. Improved procedures for clearing roots and staining parasitic and vesicular-arbuscular mycorrhizal fungi for rapid assessment of infection. Transact. Brit. Mycol. Soc. 1970;55:158-161. DOI 10.1016/S0007-1536(70)80110-3.

Razzauti M., Galan M., Bernard M., Maman S., Klopp C., Charbonnel N., Vayssier-Taussat M., Eloit M., Cosson J.-F. A comparison between transcriptome sequencing and 16S metagenomics for detection of bacterial pathogens in wildlife. PLoS Negl. Trop. Dis. 2015; 9(8):e0003929. DOI 10.1371/journal.pntd.0003929.

Rodionov A.V., Gnutikov A.A., Kotsinyan A.R. ITS1–5.8S rDNA–ITS2 sequense in the 35S pRNA genes as a marker in the reconstruction of cereal (Poaceae) phylogeny. Uspekhi Sovremennoy Biologii = Advances in Current Biology. 2016;136(5):416. (in Russian)

Rognes T., Flouri T., Nichols B., Quince C., Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584. DOI 10.7717/peerj.2584.

Salipante S.J., Kawashima T., Rosenthal C., Hoogestraat D.R., Cummings L.A., Sengupta D.J., Harkins T.T., Cookson B.T., Hoffman N.G. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. Appl. Environ. Microbiol. 2014;80(24):7583-7591. DOI 10.1128/aem.02206-14.

Savary R., Masclaux F.G., Wyss T., Droh G., Corella J.C., Machado A.P., Morton J.B., Sanders I.R. A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis*. ISME J. 2017;12(1):17-30. DOI 10.1038/ismej.2017.153.

Schenck N.C., Pérez Y. Manual for the Identification of VA Mycorrhizal Fungi. 3rd ed. Gainesville, FL: Synergistic Publ., 1990.

Schloss P.D. Application of a database-independent approach to assess the quality of operational taxonomic unit picking methods. mSystems. 2016;1(2):e00027-16.

Schoch C.L., Seifert K.A., Huhndorf S., Robert V., Spouge J.L., Levesque C.A., Chen W., Bolchacova E., Voigt K., Crous P.W., Miller A.N. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. Proc. Natl. Acad. Sci. USA. 2012;109:6241-6246. DOI 10.1073/pnas.1117018109.

Schüßler A. Glomeromycota species list. 2019. Available at http://www.amf-phylogeny.com/amphylo_species.html

Senés-Guerrero C., Schüßler A. A conserved arbuscular mycorrhizal fungal core-species community colonizes potato roots in the Andes. Fungal Divers. 2015;77(1):317-333. DOI 10.1007/s13225-015-0328-7.

Senés-Guerrero C., Torres-Cortés G., Pfeiffer S., Rojas M., Schüßler A. Potato-associated arbuscular mycorrhizal fungal communities in the Peruvian Andes. Mycorrhiza. 2014;24(6):405-417. DOI 10.1007/s00572-013-0549-0.

Stockinger H., Krüger M., Schüssler A. DNA barcoding of arbuscular mycorrhizal fungi. New Phytol. 2010;187(2):461-474. DOI 10.1111/j.1469-8137.2010.03262.x.

Stockinger H., Peyret-Guzzon M., Koegel S., Bouffaud M.-L., Redecker D. The largest subunit of RNA polymerase II as a new marker gene to study assemblages of arbuscular mycorrhizal fungi in the field. PLoS One. 2014;9(10):e107783. DOI 10.1371/journal.pone.0107783.

Tamura K., Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 1993;10:512-526. DOI 10.1093/oxfordjournals.molbev.a040023.

Tedersoo L., Anslan S., Bahram M., Põlme S., Riit T., Liiv I., Kõljalg U., Kisand V., Nilsson H., Hildebrand F., Bork P., Abarenkov K. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. MycoKeys. 2015;10:1-43. DOI 10.3897/mycokeys.10.4852.

Trouvelot A., Kough J.L., Gianinazzi-Pearson V. Mesure du taux de mycorhization VA d'un système radiculaire. Recherche de méthodes ayant une signification fonctionnelle. In: Gianinazzi-Pearson V., Gianinazzi S. (Eds.). Physiological and Genetical Aspects of Mycorrhizae. Paris: INRA-Press, 1986. (in French)

UNITE. Unified system for the DNA based fungal species linked to the classification Ver. 7.2. 2017. Available at https://unite.ut.ee

Vasar M., Andreson R., Davison J., Jairus T., Moora M., Remm M., Young J.P.W., Zobel M., Öpik M. Increased sequencing depth does not increase captured diversity of arbuscular mycorrhizal fungi. Mycorrhiza. 2017;27(8):761-773. DOI 10.1007/s00572-017-0791-y.

Vorob'ev N.I., Yurkov A.P., Provorov N.A. Programma Vychisleniya Indeksov Mikorizatsii Kornei Rastenii = The Program for Calculation of Indexes of Plant Root Mycorrhization. Registration certificate of computer software Svid-vo No. 2016612112 ot 12.02.16 o gos. reg. programmy EVM. FGBNU VNIISKHM (RF). Zayavl. 21.12.2015. Opubl. 20.03.16 No. 2016612112, 12.02.16. FSBSI ARRIAM (RU). Appl. December 21, 2015. Publ. March 20, 2016. (in Russian)

White T.J., Bruns T., Lee S., Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Innis M.A., Gelfand D.H., Sninsky J.J., White T.J. (Eds.). PCR Protocols: a Guide to Methods and Applications. San Diego: Academic Press, 1990.

Yurkov A.P., Jacobi L.M., Gapeeva N.E., Stepanova G.V., Shishova M.F. Development of arbuscular mycorrhiza in highly responsive and mycotrophic host plant – black medick (*Medicago lupulina* L.). Russian Journal of Developmental Biology. 2015;46(5):263-275. DOI 10.1134/S1062360415050082.

Yurkov A.P., Kozhemyakov A.P., Stepanova G.V. The effectiveness of some microbial bioproducts based on bacteria and fungi arbuscular mycorrhiza. Multifunctional Adaptive Fodder Production: Collection of Proceedings (Mnogofunktsional'noe Adaptivnoe Kormoproizvodstvo: Sbornik Trudov). 2018a;19:20-28. (in Russian)

Yurkov A.P., Kryukov A.A., Gorbunova A.O., Kozhemyakov A.P., Stepanova G.V., Machs E.M., Rodionov A.V., Shishova M.F. Molecular genetic identification of arbuscular mycorrhiza fungi. Ekologicheskaya Genetika = Ecological Genetics. 2018b;16(2)11-23. DOI 10.17816/ecogen16211-23. (in Russian)

Yurkov A.P., Kryukov A.A., Jacobi L.M., Kozhemyakov A.P., Shishova M.F. The relationship between the activity indicators and the effectiveness of arbuscular mycorrhiza fungi strains of various origin. Tavricheskiy Vestnik Agrarnoy Nauki = Tavrichesky Bulletin of Agrarian Science. 2017a;4(12):31-41. (in Russian)

Yurkov A.P., Laktionov Yu.V., Kozhemyakov A.P., Stepanova G.V. Analysis of the symbiotic efficiency of bacterial and fungal preparations on feed crops according to seed yield. Kormoproizvodstvo = Fodder Production. 2017b;3:16-21. (in Russian)

ORCID ID

A.A. Kryukov orcid.org/0000-0002-8715-6723
E.M. Machs orcid.org/0000-0001-9347-5379
Y.V. Mikhaylova orcid.org/0000-0001-9278-0937
A.V. Rodionov orcid.org/0000-0003-1146-1622
P.M. Zhurbenko orcid.org/0000-0002-2102-4568
A.P. Yurkov orcid.org/0000-0002-1072-5166

Acknowledgements. The work was performed using the material and technical base of the Center for Collective Use of Scientific Equipment "Genomic Technologies, Proteomics and Cell Biology" of the All-Russian Research Institute for Agricultural Microbiology and was supported by Russian Foundation for Basic Research grants (18-016-00220, 19-29-05275, and a portion of 17-00-00340, 17-00-00337; molecular-genetic identification of arbuscular mycorrhizal fungi). Part of this work was performed in the framework of the state task No. 0664-2019-0026 (morphological identification of arbuscular mycorrhizal fungi).

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

Received June 17, 2019. Revised November 15, 2019. Accepted November 17, 2019. Published online December 27, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):168-175
DOI 10.18699/VJ20.610

Изменение кишечного микробиома пациентов с язвенным колитом после трансплантации кишечной микробиоты

А.Ю. Тикунов[1], В.В. Морозов[1], А.Н. Швалов[2], А.В. Бардашева[1], Е.В. Шрайнер[1], О.А. Максимова[3], И.О. Волошина[3], В.В. Морозова[1], В.В. Власов[1], Н.В. Тикунова[1]

[1] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора Российской Федерации, р. п. Кольцово, Новосибирская область, Россия
[3] ООО «Центр персонализированной медицины», Новосибирск, Россия
e-mail: tikunova@niboch.nsc.ru

Аннотация. Микробиота кишечника человека – это динамическая система, находящаяся под воздействием организма хозяина и внешних факторов. Возникающие нарушения кишечной микробиоты могут привести к патологическим состояниям, включая воспалительные и онкологические заболевания желудочно-кишечного тракта. Одним из возможных способов воздействия на микробиоту кишечника является фекотрансплантация (ФТ) – введение кишечной микробиоты от здорового донора в кишечный тракт пациента. В настоящее время в ряде стран этот метод используется для нормализации микробиоты кишечника, в основном при хронических воспалительных заболеваниях кишечника. В России (Новосибирск) уже несколько лет ведутся пилотные исследования эффективности ФТ при язвенном колите. Цель данной работы – оценить изменение микробиома кишечника 20 пациентов с язвенным колитом после однократного проведения ФТ. Основной метод – сравнительный анализ библиотек последовательностей 16S рибосомальной РНК, созданных на основе образцов, полученных от пациентов с язвенным колитом до и после ФТ и секвенированных на платформе Illumina MiSeq. Результаты исследования показали, что ФТ привела к увеличению среднего биоразнообразия последовательностей в образцах, полученных после ФТ, по сравнению с образцами, собранными до ФТ, хотя разница не была статистически достоверной. Доля последовательностей Firmicutes, являющихся доминирующей компонентой кишечной микробиоты здоровых людей, уменьшилась (~32 % vs. >70 %), а доля последовательностей Proteobacteria увеличилась (>9 % vs. <5 %). В некоторых образцах, собранных до ФТ, были обнаружены последовательности патогенных представителей Firmicutes и Proteobacteria, включая *Acinetobacter* spp., *Enterococcus* spp., *Klebsiella pneumoniae*, *Proteus mirabilis*, *Staphylococcus aureus*, *Stenotrophomonas maltophylia*, *Streptococcus* spp. В большинстве случаев после ФТ доля таких последовательностей резко сократилась. Исключение составили последовательности *Clostridium difficile*, содержание которых в образцах почти половины пациентов составляло менее 0.5 %; после ФТ доля последовательностей *C. difficile* значительно уменьшилась лишь у трех пациентов. Следует отметить, что после ФТ повысилось на порядок содержание *Lactobacillus* spp. и существенно расширился их видовой состав. По результатам исследования можно сделать предварительное заключение о том, что даже однократная процедура ФТ приводит к повышению биоразнообразия микробиоты пациентов и оптимизации ее таксономического состава.
Ключевые слова: микробиом; язвенный колит; 16S рРНК профилирование; трансплантация кишечной микробиоты.

Для цитирования: Тикунов А.Ю., Морозов В.В., Швалов А.Н., Бардашева А.В., Шрайнер Е.В., Максимова О.А., Волошина И.О., Морозова В.В., Власов В.В., Тикунова Н.В. Изменение кишечного микробиома пациентов с язвенным колитом после трансплантации кишечной микробиоты. Вавиловский журнал генетики и селекции. 2020; 24(2):168-175. DOI 10.18699/VJ20.610

Fecal microbiome change in patients with ulcerative colitis after fecal microbiota transplantation

A.Y. Tikunov[1], V.V. Morozov[1], A.N. Shvalov[2], A.V. Bardasheva[1], E.V. Shrayner[1], O.A. Maksimova[3], I.O. Voloshina[3], V.V. Morozova[1], V.V. Vlasov[1], N.V. Tikunova[1]

[1] Institute of Chemical Biology and Fundamental Medicine of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] State Research Center of Virology and Biotechnology Vector, Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
[3] Limited liability company "Personalized Medicine Center", Novosibirsk, Russia
e-mail: tikunova@niboch.nsc.ru

Abstract. Intestinal human microbiota is a dynamic system that is under the pressures of its host organism and external factors. Microbiota disruption caused by these factors can lead to severe diseases including inflammatory and oncological diseases of the gastrointestinal tract. One of the possible approaches in managing the intestinal microbiota is fecal microbiota transplantation (FT) – transfer of the microbiota from the stool of a healthy donor to the intestinal tract of a recipient patient. Currently, this procedure is recognized as an efficacious method to normalize the

intestinal microbiota mainly in inflammatory diseases of the gastrointestinal tract. In Russia, pilot studies of the effectiveness of FT in patients with ulcerative colitis have been conducted for several years, and these studies were started in Novosibirsk. The aim of this study was to assess the change of intestinal microbiome in 20 patients with ulcerative colitis after a single FT procedure. The main method is a comparative analysis of 16S ribosomal RNA sequence libraries constructed using fecal samples obtained from patients with ulcerative colitis before and after FT and sequenced on the Illumina MiSeq platform. The obtained results showed that FT led to an increase in average biodiversity in samples after FT compared to samples before FT; however, the difference was not significant. In the samples studied, the proportion of Firmicutes sequences, the major gastrointestinal microbiota of healthy people, was decreased (~32 % vs. >70 %), while the proportion of Proteobacteria sequences was increased (>9 % vs. <5 %). In some samples collected before FT, sequences of pathogenic Firmicutes and Proteobacteria were detected, including *Acinetobacter* spp., *Enterococcus* spp., *Klebsiella pneumoniae*, *Proteus mirabilis*, *Staphylococcus aureus*, *Stenotrophomonas maltophylia*, *Streptococcus* spp. In most cases, the proportion of such sequences after FT substantially decreased in appropriate samples. The exception was the *Clostridium difficile* sequences, which accounted for <0.5 % of the sequences in samples from almost half of the patients and after FT, the share of such *C. difficile* sequences was significantly reduced only in samples from three patients. It should be noted that the proportion of *Lactobacillus* spp. increased ten-fold and their species composition significantly expanded. According to the obtained results, a preliminary conclusion can be made that even a single FT procedure can lead to an increase in the biodiversity of the gastrointestinal microbiota in patients and to the optimization of the taxonomic composition of the microbiota.
Key words: microbiome; ulcerative colitis; 16S rRNA profiling; fecal microbiota transplantation.

For citation: Tikunov A.Y., Morozov V.V., Shvalov A.N., Bardasheva A.V., Shrayner E.V., Maksimova O.A., Voloshina I.O., Morozova V.V., Vlasov V.V., Tikunova N.V. Fecal microbiome change in patients with ulcerative colitis after fecal microbiota transplantation. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):168-175. DOI 10.18699/VJ20.610 (in Russian)

Введение

Широкое применение технологий NGS (next generation sequencing) обеспечило детальную характеризацию микробных сообществ, преимущественно бактериальных, ассоциированных с организмом человека. В настоящее время микробиота кишечника человека рассматривается как динамическая система, находящаяся под воздействием организма хозяина и внешних факторов (Fujimura et al., 2010; Qin et al., 2010). Известно, что основными компонентами нормальной кишечной микробиоты человека являются представители Firmicutes и Bacteroidetes, хотя коровая группа конкретных видов бактерий не совпадает у разных здоровых индивидуумов (Donaldson et al., 2016), и различные по составу варианты нормальной микробиоты способны обеспечить стабильное функционирование этого сложного микробного сообщества (Lozupone et al., 2012).

Дисбаланс кишечной микробиоты, возникающий под действием внешних или внутренних факторов, может привести к патологическим состояниям, включая не только воспалительные и онкологические заболевания желудочно-кишечного тракта, но и нарушения иммунной системы, диабет второго типа, сосудистые заболевания и даже нарушения функций головного мозга (O'Hara, Shanahan, 2006). Одним из возможных способов воздействия на микробиоту кишечника, наряду с использованием антибиотиков, пробиотиков и пребиотиков, является фекотрансплантация (ФТ) – введение кишечной микробиоты от здорового донора в кишечный тракт пациента. Считается, что до нашей эры применение ФТ практиковали в Китае, а в наше время эту процедуру впервые провели в 1958 г. при лечении пациента с энтероколитом (Eiseman et al., 1958). Однако лишь недавно метод стал успешно использоваться для нормализации микробиоты кишечника при различных заболеваниях, включая хронические воспалительные заболевания кишечника (Aas et al., 2003; Khoruts et al., 2010; Angelberger et al., 2013; Pigneur, Sokol, 2016; Vaughn et al., 2016; Kang et al., 2017; Paramsothy et al., 2017; Staley et al., 2017).

Наибольшая результативность применения ФТ показана при псевдомембранозном колите, ассоциированном с *Clostridium* (*Clostridioides*) *difficile* (Drekonja et al., 2015; Khoruts, Sadowsky, 2016; Cheng, Fisher, 2017; Staley et al., 2017; Goldenberg et al., 2018). Быстрое улучшение состояния пациентов и высокая эффективность ФТ (более 80 %) при этих инфекциях была подтверждена в многоцентровых рандомизированных клинических испытаниях (van Nood et al., 2013; Cammarota et al., 2015), и сейчас в США и в европейских странах рекомендуется применять ФТ уже после второго или третьего эпизода *C. difficile*-ассоциированного колита (Debast et al., 2014).

Имеются свидетельства того, что ФТ может быть полезной и при язвенном колите (ЯК). Эффективность применения ФТ для лечения пациентов с ЯК в различных исследованиях существенно варьировала – от 20 до 92 % (Angelberger et al., 2013; Kellermayer et al., 2015). В рандомизированных клинических испытаниях ремиссия зарегистрирована на уровне 25–27 %, что, впрочем, статистически значимо превышало эффект плацебо (Moayyedi et al., 2015; Rossen et al., 2015; Paramosthy et al., 2017). Патогенез ЯК не вполне ясен; полагают, что он имеет сложную природу и опосредован нарушениями кишечной микробиоты, генетической предрасположенностью и экологическими факторами (Shen et al., 2018). Как и при *C. difficile*-ассоциированном колите, биоразнообразие микробиоты кишечника при ЯК существенно снижено, уменьшено количество представителей Bacteroidetes и Firmicutes. При этом в микробиоте пациентов с ЯК увеличивается количество представителей Proteobacteria и Actinomycetes, выявляются *C. difficile*, *Helicobacter pylori*, *Salmonella* spp., *Yersinia* spp. и энтероинвазивные *Escherichia coli* (Okhusa et al., 2003; Saebo et al., 2005; Gradel et al., 2009; Sonnenberg, Genta, 2012; Deshpande et al., 2013; Reddy, Brandt, 2013; Shen et al., 2018). До сих пор не ясно,

является ли ЯК результатом нарушенного иммунного ответа на нормальную микробиоту, или это нормальный иммунный ответ на дисбаланс в микрофлоре кишечника (Cheng, Fisher, 2017).

Несмотря на кажущуюся очевидность, механизмы положительного действия ФТ при ЯК не вполне ясны. Полагают, что эффективность ФТ связана с увеличением разнообразия микробиоты кишечника, что приводит к повышению обилия «полезных» бактерий и препятствует колонизации кишечника патогенными бактериями (Broecker et al., 2016; Chehoud et al., 2016; Khoruts, Sadowsky, 2016). Однако неизвестно, вовлечены ли в этот процесс другие механизмы, включая возможное влияние виробиоты, действие иммунной системы пациента, привнесение регуляторных высоко- и низкомолекулярных соединений при ФТ; также до сих пор не определен список видов бактерий, обусловливающих нормализацию микробиоты. Цель данного исследования – оценить изменение микробиома кишечника пациентов с ЯК после проведения ФТ на основе профилирования 16S рибосомальной РНК в образцах, полученных до и после лечения.

Материалы и методы

В работе использовали образцы фекалий, полученные от 20 пациентов (27–57 лет) с диагнозом ЯК. Диагноз подтверждали на основании результатов изучения уровня фекального кальпротектина, данных фиброколоноскопии и гистологического исследования биоптатов, взятых из разных отделов толстой и подвздошной кишок. Образцы пациентов, в которых рутинными методами были обнаружены *C. difficile*, в исследование не вовлекались. Все пациенты предоставили информированное согласие с проводимым исследованием и анонимной обработкой данных. Образцы собирали за один-два дня до ФТ и через 7–12 дней после ФТ. Донорами были молодые здоровые добровольцы (20–39 лет) без хронических заболеваний, не перенесшие инфекции и не подвергавшиеся госпитализации по крайней мере последние два месяца. Все доноры прошли обследование, включающее в себя общий и биохимический анализы крови, а также ИФА крови на наличие лямблий, токсокар, описторхов, аскарид, трихинелл. Кроме того, с использованием стандартных тест-систем подтверждали отсутствие у доноров возбудителя сифилиса, ВИЧ-1 и ВИЧ-2, вирусов гепатита B и C. Также рутинными методами проводили анализ фекалий на дисбиоз и на отсутствие патогенной микрофлоры (*C. difficile*, *Campilobacter jejuni*, *Salmonella* spp., *Shigella* spp., энтероинвазивная *Escherichia coli*, *Cryptosporidium* spp., *Cyclospora* spp., *Giardia* spp., *Isospora* spp.), ротавирусов A, норовирусов I и II и аденовирусов F, а также гельминтов и их яиц. Пилотное исследование было одобрено Локальным этическим комитетом Автономной некоммерческой организации «Центр новых медицинских технологий в Академгородке».

По 50 мг каждого образца от пациентов суспендировали в 300 мкл 0.9 % NaCl и центрифугировали при 2 тыс. об. в течение 10 мин. Суммарную ДНК очищали из 100 мкл осветленной клеточной суспензии с помощью набора для выделения ДНК из клеток тканей и крови («БиоЛабМикс», Россия) с добавлением лизоцима для повышения эффективности извлечения ДНК из грамположительных бактерий. С использованием полученной ДНК в качестве матрицы, фьюжн-праймеров (NEB-FF 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTACGGGAGGCAGCAG-3′, NEB-FR 5′-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGACTACCGGGGTATCT-3′) и высокоточной полимеразы Q5 (New England Biolabs, США) проводили амплификации фрагмента гена 16S рРНК, содержащего вариабельные участки V3–V4. Продукты амплификации очищали электрофоретически в геле из легкоплавкой SeaKem GTG-агарозы (Lonza, США). Обогащение полученных ампликонов, введение баркодов и служебных последовательностей для дальнейшего секвенирования на платформе MiSeq выполняли, используя полимеразу Q5 и набор олигонуклеотидов Dual index set (New England Biolabs), согласно инструкции производителя. Полученные библиотеки очищали на магнитных частицах AMPure XP (Beckman Coulter, США); концентрацию ДНК измеряли с помощью набора Qubit dsDNA HS (Life Technologies, США). По результатам измерений библиотеки объединяли в пул таким образом, чтобы соотношение ДНК библиотек в пуле было приблизительно равным. Секвенировали на высокопроизводительном секвенаторе MiSeq с набором реагентов MiSeq reagent kit v2 2×250-cycles (Illumina, США).

Результаты секвенирования анализировали с использованием пакета программного обеспечения UGENE v.1.32 (Unipro UGENE, Россия). Полученные риды картировали на базу данных 16S рРНК, размещенную на Национальном сервере NSBI (США), с использованием пакета Clark. Предварительно из последовательностей ридов удаляли последовательности адаптеров и проводили фильтрацию ридов по качеству. Риды анализировали двумя методами: с помощью генерации операционных таксономических единиц (OTU) с последующим картированием последовательностей на полученные OTU в пакете программ Usearch-9.2 и путем классификации ридов алгоритмом Kraken по базе данных известных последовательностей 16S рРНК Silva v.132 (full). В первом случае OTU генерировали алгоритмом unoise2 с отбраковкой химерных последовательностей и учетом ошибок чтения. Таблицы полученных частот встречаемости OTU были обработаны в среде R3.3.3. Во втором случае риды картировали на базу данных 16S рРНК Silva с помощью алгоритма seed-kraken с использованием разреженного *k*-мера со специальной решеткой, позволяющей увеличить специфичность классификации. Индекс Шеннона рассчитывали в пакете программ R; достоверность различий между индексами Шеннона определяли с помощью t-теста Хатчесона. Визуализацию результатов анализа библиотек последовательностей методом главных координат PCoA проводили на основе матриц дистанций с использованием пакета программ vegan.

Результаты и обсуждение

На основе ДНК, выделенной из образцов фекалий от 20 пациентов с ЯК до и после ФТ, было сконструировано 40 библиотек фрагментов гена 16S рРНК. Фрагменты гена включали вариабельные участки V3 и V4, используемые обычно для таксономической классификации бактерий

(Chakravorty et al., 2007; Wang, Qian, 2009). В результате секвенирования библиотек, которые были созданы на основе образцов, собранных до ФТ, получено от 106411 до 1751663 ридов (в среднем 141010). Библиотеки из образцов, собранных после ФТ, содержали от 107042 до 173855 ридов (в среднем 142060). По результатам классификации в библиотеках из первой и второй групп к определенному типу прокариот были таксономически отнесены в среднем 99.7 и 99.6 % ридов соответственно. Лишь незначительная часть последовательностей осталась неклассифицированной. Всего выявлены последовательности 13 типов бактерий, основными из которых были Firmicutes, Bacteroidetes и Proteobacteria. В трех образцах выявлены последовательности архей, принадлежащих к роду *Methanobrevibacter* (тип Euryarchaeota), однако доля таких последовательностей в соответствующих библиотеках не превышала 0.1 %.

Анализ полученных данных показал, что биоразнообразие бактериальных сообществ в образцах от пациентов до и после ФТ различается. Так, индекс Шеннона для выборки образцов от пациентов после лечения (3.43 ± 0.71) был выше, чем для образцов, взятых до лечения (3.05 ± 0.67), хотя разница статистически недостоверна (рис. 1). Однако при попарном сравнении индексов Шеннона для образцов, полученных от одного пациента до и после ФТ, различия были во всех случаях статистически достоверными (SD, 0.011–0.019); при этом у 15 пациентов биоразнообразие микробных сообществ значимо увеличилось и лишь у 5 пациентов – уменьшилось.

Большее биоразнообразие сообществ в образцах, полученных после ФТ, подтвердилось при анализе данных методом главных координат (рис. 2). Видно, что 60 % образцов, собранных до лечения, находятся в области отрицательных значений первой главной координаты, тогда как 65 % образцов, взятых после лечения, находятся в области положительных значений. Аналогично в области положительных значений второй главной координаты находится

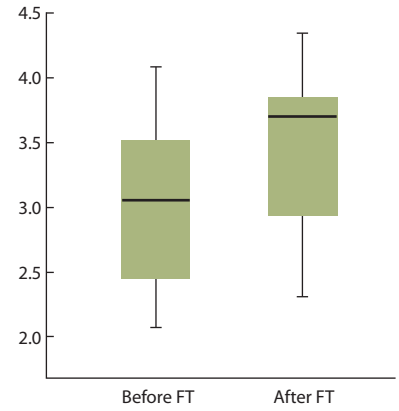


Fig. 1. Shannon index, reflecting the biodiversity of sequences in samples obtained from patients before and after FT.

лишь 45 % образцов, полученных до лечения, и 70 % образцов после ФТ. Отметим, что в образцах от 15 пациентов (75 %) после проведения ФТ значение увеличилось в области и первой, и второй главных координат; еще у четырех пациентов значение увеличилось в области хотя бы одной

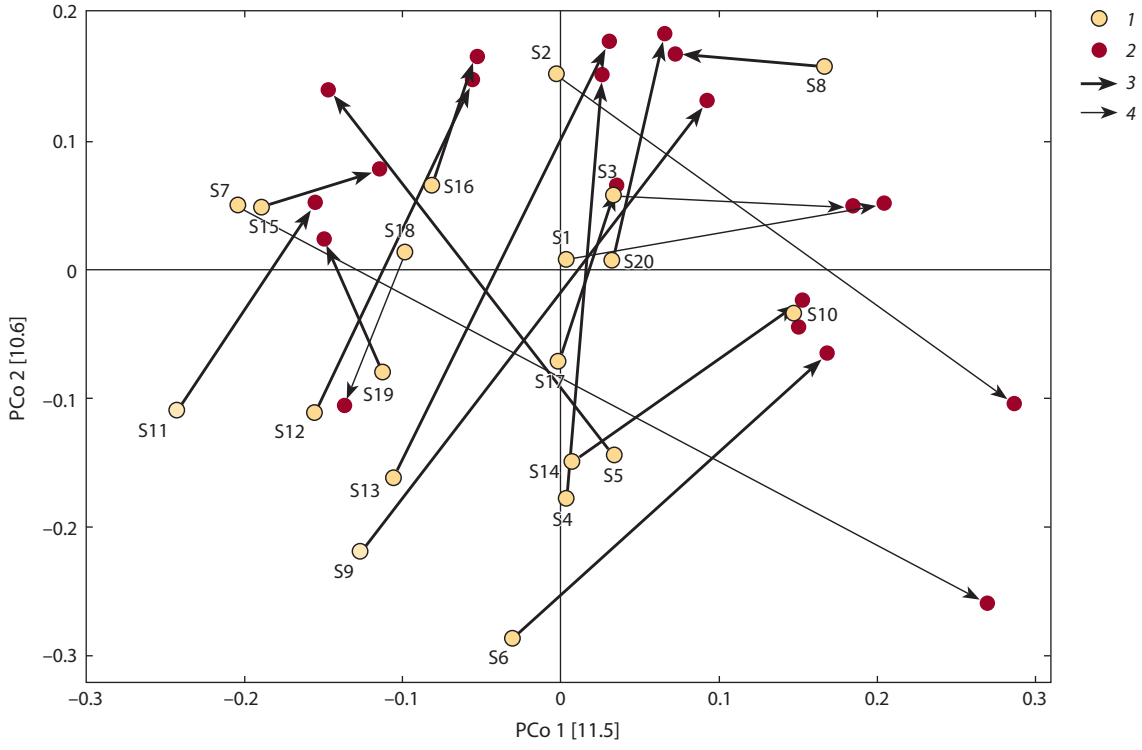


Fig. 2. Visualization of sequence library analysis by the method of principal coordinates PCoA made on matrices of distances:

1, data for libraries from samples obtained before FT; *2*, from samples obtained after FT. *3*, *4*, lines connecting two samples from one patient for which biodiversity (*3*) increases or (*4*) decreases after FT. Designations S1–S20 correspond to sample numbers before FT. The values of the first and second principal coordinates are presented on the *OX* and *OY* axes, respectively.

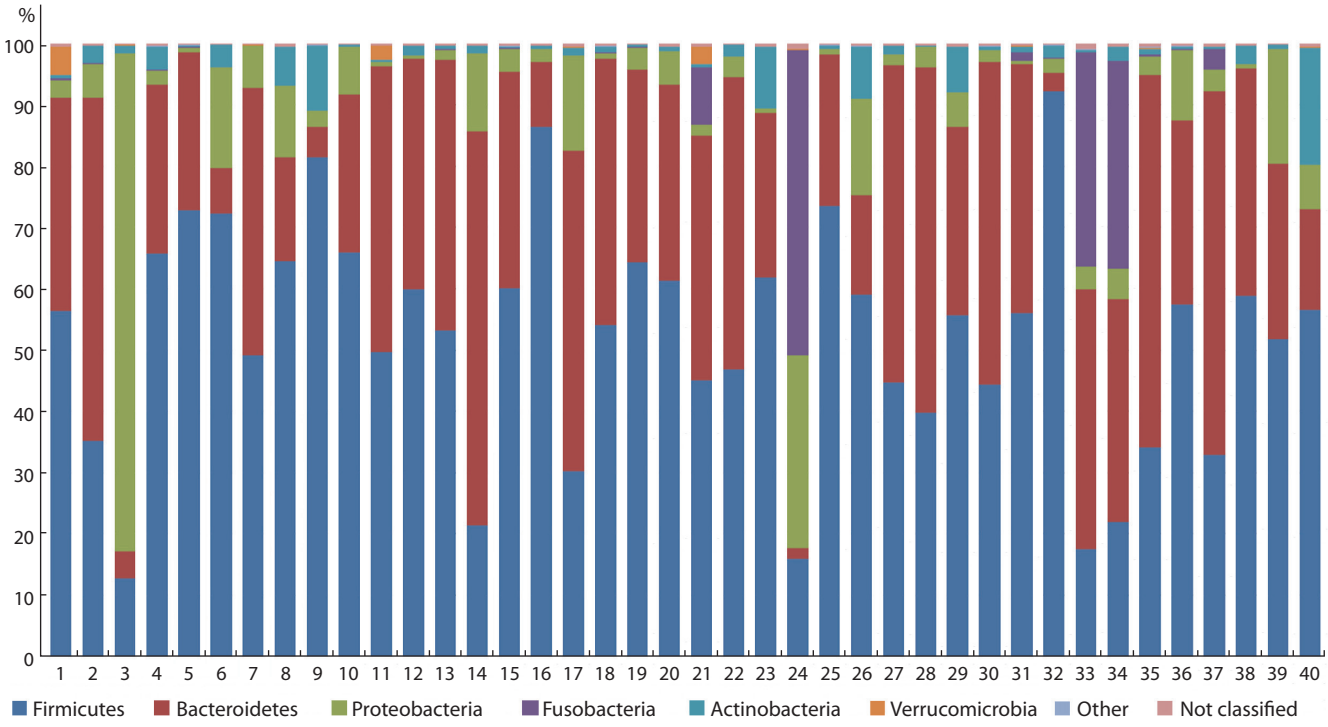


Fig. 3. Taxonomic classification of OTUs at the phylum level based on the Silva v.132 (full) database.

Here and in Fig. 4 samples 1–20 were obtained from patients before FT (enumeration follows Fig. 2); samples 21–40 were collected after FT. Samples 1 and 21, 2 and 22, etc. were obtained from the first, second, etc. patients, respectively. Phyla with representation exceeding 0.1 % are shown.

из главных координат; только для пары образцов от одного пациента (№ 18) биоразнообразие уменьшилось после ФТ в области значений обеих главных координат (см. рис. 2). Тем не менее при уменьшении значений в области хотя бы одной главной координаты после ФТ индекс Шеннона для этой же пары образцов статистически значимо уменьшался.

Полученные данные подтверждают возможность повышения биоразнообразия бактериального сообщества кишечника после проведения ФТ. Следует подчеркнуть, что в настоящем исследовании мы сравнивали исходный микробиом пациентов и микробиом после первой процедуры ФТ. Согласно опубликованным сведениям зарубежных авторов, даже после нескольких процедур ФТ статистически значимое повышение биоразнообразия микробиоты регистрируется не у всех пациентов (Angelberger et al., 2013).

Анализ таксономического состава показал, что в образцах до и после ФТ в среднем 55.7 % (от 12.7 до 72.8 %) и 48.3 % (от 15.9 до 92.3 %) выявленных последовательностей соответственно принадлежали к типу Firmicutes (рис. 3), 32.3 % (от 4.4 до 64.4 %) и 33.7 % (от 1.7 до 60.7 %) – к типу Bacteroidetes, а 9.2 % (от 0.4 до 81.3 %) и 6.2 % (от 0.7 до 31.5 %) – к типу Proteobacteria. Четвертыми по встречаемости были последовательности Fusobacteria, в среднем 3.4 % по всем библиотекам, однако их доля значительно увеличилась после ФТ: от <0.1 (0–0.3 %) до 6.7 % (0.1–9.4 %). После ФТ увеличилась также доля последовательностей Actinobacteria, хотя и не так значительно – от 1.9 (0.1–10.7 %) до 3.1 % (0.2–19.4 %). Последовательности Verrucomicrobia присутствовали в каждом образце в небольших количествах, а их встречаемость после ФТ уменьшилась в среднем с 0.4 до 0.2 %. Остальные последовательности встречались лишь в отдельных образцах, не превышая 0.1 % от всех последовательностей в этом образце.

Известно, что микробиота здоровых людей состоит из постоянных и транзиторных видов, относящихся более чем к 17 типам, включая Firmicutes (>70 %), Bacteroidetes (>30 %), протеобактерии (<5 %), актинобактерии (<2 %), Fusobacteria и Verrucomicrobia (<1 %) (Belizário et al., 2018). Полученные нами результаты коррелируют с данными других исследователей, свидетельствующими о пониженном биоразнообразии микробиоты кишечника при ЯК (Manichanh et al., 2012; Machiels et al., 2014; Bajer et al., 2017). Так, в исследуемых образцах от пациентов в среднем присутствовало существенно меньше последовательностей Firmicutes, что хорошо согласуется с наблюдениями других авторов (Machiels et al., 2014). В основном Firmicutes были представлены последовательностями классов Clostridia (рис. 4), в среднем 47.4 и 40.6 % в библиотеках из образцов, полученных до и после ФТ соответственно. Из них доминировали *Faecalibacterium prausnitzii* (15.7 и 11.3 % соответственно) и *Roseburia hominis* (2.3 и 0.5 %), ответственные за расщепление широкого спектра углеводов, включая крахмал и инулин, с образованием бутиратов (Duncan et al., 2007; Machiels et al., 2014). Также были представлены последовательности классов Negativicutes (6.8 и 4.0 %) и Bacilli (1.3 и 2.8 %). Следует отметить, что, несмотря на относительно невы-

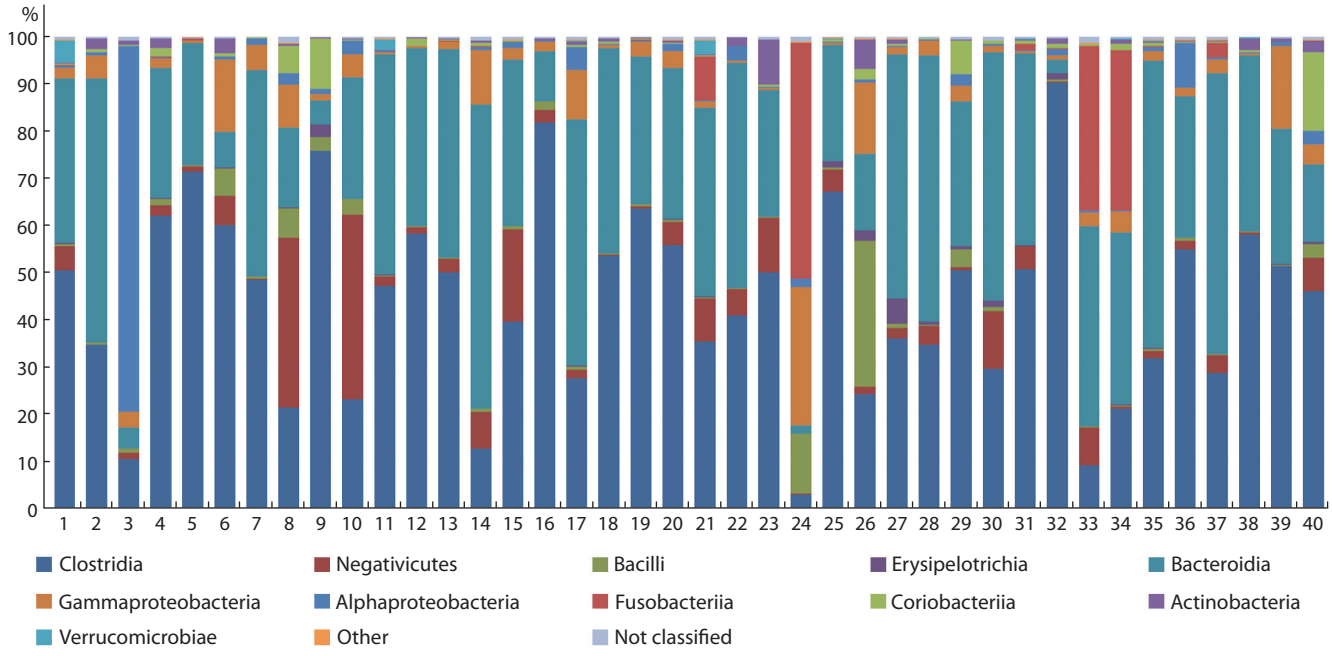


Fig. 4. Taxonomic classification of OTUs at the class level based on the Silva v.132 (full) database.
Classes with representation exceeding 0.1 % are shown.

сокую встречаемость *Bacillus* spp. и *Lactobacillus* spp., их доля после ФТ повысилась в 3.5 и 11 раз соответственно. При этом существенно расширился видовой состав лактобацилл, которые, как известно, не только участвуют в расщеплении лактозы и других углеводов, но и являются антагонистами по отношению к патогенным микроорганизмам, вытесняя их из микробного сообщества кишечника человека.

Доля последовательностей, принадлежащих к типу Bacteroidetes, в исследуемых образцах от пациентов с ЯК не была снижена (см. рис. 3 и 4) и составила около трети от всех последовательностей, что отличается от наблюдений других исследователей (Machiels et al., 2014). В основном этот тип был представлен последовательностями *Bacteroides* spp. и *Prevotella* spp., причем после ФТ доля последовательностей бактероидов уменьшилась в среднем с 19.7 до 9.6 %, а последовательностей, принадлежащих к роду *Prevotella*, существенно увеличилась, с 3.7 до 14.6 %. Известно, что большинство бактероидов, обитающих в кишечнике человека, способны разлагать разнообразные растительные полисахариды (Flint et al., 2012), причем в кишечной микробиоте жителей западных стран преобладают *Bacteroides* spp., а в микробиоте населения из стран с преимущественно растительной диетой – *Prevotella* spp. (Ley, 2016).

Отметим, что доля последовательностей Proteobacteria (см. рис. 3) в образцах, собранных до ФТ (в среднем 9.2 %), превышала таковую, обычно регистрируемую у здоровых людей (<5 %) (Belizário et al., 2018). Это объясняется повышенным содержанием *Salmonella* spp. в микробных сообществах пациентов с ЯК и наличием в отдельных образцах в большом количестве последовательностей патогенных бактерий, способных вызвать желудочно-кишечные заболевания (более 9 % *Acinetobacter* spp., 0.5–1 % *Klebsiella pneumoniae*), что установлено и в других исследованиях (Gradel et al., 2009; Shen et al., 2018). В единичных образцах были обнаружены также *Proteus mirabilis* и *Stenotrophomonas maltophilia*, доля которых не превышала 0.1 %. Надо сказать, что после проведения ФТ содержание перечисленных последовательностей патогенных Proteobacteria в образцах существенно уменьшилось, порой до 0.02 % и менее.

Кроме последовательностей патогенных Proteobacteria, в некоторых образцах, полученных от пациентов до ФТ, были обнаружены последовательности патогенных представителей Firmicutes. Так, в девяти образцах найдены последовательности *C. difficile*, доля которых превышала 0.5 %. Повышенная встречаемость *C. difficile* в кишечной микробиоте пациентов с ЯК отмечалась и ранее (Deshpande et al., 2013; Reddy, Brandt, 2013). Помимо последовательностей *C. difficile*, в некоторых образцах обнаружены последовательности *Staphylococcus aureus* (0.1–0.9 %), *Streptococcus* spp. и *Enterococcus* spp. (~0.1 %). Как и в случае с патогенными Proteobacteria, доля этих последовательностей после ФТ в соответствующих образцах резко уменьшилась. Исключение составили образцы с *C. difficile*: доля их последовательностей уменьшилась после ФТ только в трех образцах из девяти, составив <0.2 %.

Заключение

Таким образом, исследованы биоразнообразие и таксономический состав последовательностей фрагмента гена 16S рРНК, ассоциированных с кишечной микробиотой у 20 пациентов с ЯК до и после ФТ. Результаты показали, что однократное проведение процедуры ФТ привело к увеличению среднего биоразнообразия последовательностей в образцах, полученных после ФТ, по сравнению с образцами, собранными до ФТ, хотя разница не была

статистически достоверной. Доля последовательностей Firmicutes, являющихся доминирующей компонентой кишечной микробиоты здоровых людей, была снижена (~32 % vs. >70 %), а доля последовательностей Proteobacteria увеличена (>9 % vs. <5 %). В некоторых образцах, собранных до ФТ, обнаружены значимые содержания последовательностей патогенных представителей Firmicutes и Proteobacteria, включая *Acinetobacter* spp., *Enterococcus* spp., *K. pneumoniae*, *P. mirabilis*, *S. aureus*, *St. maltophilia*, *Streptococcus* spp. В большинстве случаев после однократной процедуры ФТ доля таких последовательностей резко сократилась. Исключение составили последовательности *C. difficile*, которые были обнаружены (около 0.5 % и выше) в образцах почти половины пациентов с ЯК; после ФТ доля последовательностей *C. difficile* значительно уменьшилась лишь у трех пациентов. Следует отметить, что после ФТ содержание *Lactobacillus* spp. повысилось на порядок и существенно расширился видовой состав лактобацилл.

Результаты исследования позволяют сделать предварительный вывод, что даже однократная процедура ФТ может привести к повышению биоразнообразия микробиоты и оптимизации ее таксономического состава. Однако для того чтобы сделать заключение об эффективности такого лечения, длительности ремиссии и стабильности изменений микробиоты кишечника у пациентов с ЯК, требуются дальнейшие наблюдения за этими пациентами и анализ кишечного микробиома после последующих процедур ФТ.

Список литературы / References

Aas J., Gessert C.E., Bakken J.S. Recurrent *Clostridium difficile* colitis: case series involving 18 patients treated with donor stool administered via a nasogastric tube. Clin. Infect. Dis. 2003;36(5):580-585.

Angelberger S., Reinisch W., Makristathis A., Lichtenberger C., Dejaco C., Papay P., Novacek G., Trauner M., Loy A., Berry D. Temporal bacterial community dynamics vary among ulcerative colitis patients after fecal microbiota transplantation. Am. J. Gastroenterol. 2013;108(10):1620-1630. DOI 10.1038/ajg.2013.257.

Bajer L., Kverka M., Kostovcik M., Macinga P., Dvorak J., Stehlikova Z., Brezina J., Wohl P., Spicak J., Drastich P. Distinct gut microbiota profiles in patients with primary sclerosing cholangitis and ulcerative colitis. World J. Gastroenterol. 2017;23(25):4548-4558. DOI 10.3748/wjg.v23.i25.4548.

Belizário J.E., Faintuch J., Garay-Malpartida M. Gut microbiome dysbiosis and immunometabolism: new frontiers for treatment of metabolic diseases. Mediators Inflamm. 2018;2018:1-12. DOI 10.1155/2018/2037838.

Broecker F., Klumpp J., Moelling K. Long-term microbiota and virome in a Zürich patient after fecal transplantation against *Clostridium difficile* infection. Ann. N.Y. Acad. Sci. 2016;1372(1):29-41. DOI 10.1111/nyas.13100.

Cammarota G., Masucci L., Ianiro G., Bibbò S., Dinoi G., Costamagna G., Sanguinetti M., Gasbarrini A. Randomised clinical trial: faecal microbiota transplantation by colonoscopy vs. vancomycin for the treatment of recurrent *Clostridium difficile* infection. Aliment. Pharmacol. Ther. 2015;41(9):835-843. DOI 10.1111/apt.13144.

Chakravorty S., Helb D., Burday M., Connell N., Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J. Microbiol. Methods. 2007;69(2):330-339.

Chehoud C., Dryga A., Hwang Y., Nagy-Szakal D., Hollister E.B., Luna R.A., Versalovic J., Kellermayer R., Bushman F.D. Transfer of viral communities between human individuals during fecal microbiota transplantation. MBio. 2016;7(2):e00322. DOI 10.1128/mBio.00322-16.

Cheng Y.W., Fischer M. The present status of fecal microbiota transplantation and its value in the elderly. Curr. Treat. Options Gastroenterol. 2017;15(3):349-362. DOI 10.1007/s11938-017-0143-1.

Debast S.B., Bauer M.P., Kuijper E.J. European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for *Clostridium difficile* infection. Clin. Microbiol. Infect. 2014;20(Suppl.2):1-26. DOI 10.1111/1469-0691.12418.

Deshpande A., Pasupuleti V., Pant C., Rolston D.D., Sferra T.J. Diagnostic testing for *Clostridium difficile* infection in patients with inflammatory bowel disease. J. Clin. Gastroenterol. 2013;47(8):737-738. DOI 10.1097/MCG.0b013e318295d4ec.

Donaldson G.P., Lee S.M., Mazmanian S.K. Gut biogeography of the bacterial microbiota. Nat. Rev. Microbiol. 2016;14(1):20-32. DOI 10.1038/nrmicro3552.

Drekonja D., Reich J., Gezahegn S., Greer N., Shaukat A., MacDonald R., Rutks I., Wilt T.J. Fecal microbiota transplantation for *Clostridium difficile* infection: a systematic review. Ann. Intern. Med. 2015;162(9):630-638. DOI 10.7326/M14-2693.

Duncan S.H., Louis P., Flint H.J. Cultivable bacterial diversity from the human colon. Lett. Appl. Microbiol. 2007;44(4):343-350.

Eiseman B., Silen W., Bascom G.S., Kauvar A.J. Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. Surgery. 1958;44(5):854-859.

Flint H.J., Scott K.P., Duncan S.H., Louis P., Forano E. Microbial degradation of complex carbohydrates in the gut. Gut Microbes. 2012;3(4):289-306.

Fujimura K.E., Slusher N.A., Cabana M.D., Lynch S.V. Role of the gut microbiota in defining human health. Expert Rev. Anti. Infect. Ther. 2010;8(4):435-454. DOI 10.1586/eri.10.14.

Goldenberg S.D., Batra R., Beales I., Digby-Bell J.L., Irving P.M., Kellingray L., Narbad A., Franslem-Elumogo N. Comparison of different strategies for providing fecal microbiota transplantation to treat patients with recurrent *Clostridium difficile* infection in two English hospitals: a review. Infect. Dis. Ther. 2018;7(1):71-86. DOI 10.1007/s40121-018-0189-y.

Gradel K.O., Nielsen H.L., Schønheyder H.C., Ejlertsen T., Kristensen B., Nielsen H. Increased short- and long-term risk of inflammatory bowel disease after salmonella or campylobacter gastroenteritis. Gastroenterology. 2009;137(2):495-501. DOI 10.1053/j.gastro.2009.04.001.

Kang D.W., Adams J.B., Gregory A.C., Borody T., Chittick L., Fasano A., Khoruts A., Geis E., Maldonado J., McDonough-Means S., Pollard E.L., Roux S., Sadowsky M.J., Lipson K.S., Sullivan M.B., Caporaso J.G., Krajmalnik-Brown R. Microbiota transfer therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. Microbiome. 2017;5(1):10. DOI 10.1186/s40168-016-0225-7.

Kellermayer R., Nagy-Szakal D., Harris R.A., Luna R.A., Pitashny M., Schady D., Mir S.A., Lopez M.E., Gilger M.A., Belmont J., Hollister E.B., Versalovic J. Serial fecal microbiota transplantation alters mucosal gene expression in pediatric ulcerative colitis. Am. J. Gastroenterol. 2015;110(4):604-606. DOI 10.1038/ajg.2015.19.

Khoruts A., Dicksved J., Jansson J.K., Sadowsky M.J. Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent *Clostridium difficile*-associated diarrhea. J. Clin. Gastroenterol. 2010;44(5):354-360. DOI 10.1097/MCG.0b013e3181c87e02.

Khoruts A., Sadowsky M.J. Understanding the mechanisms of faecal microbiota transplantation. Nat. Rev. Gastroenterol. Hepatol. 2016;13(9):508-516. DOI 10.1038/nrgastro.2016.98.

Ley R.E. *Prevotella* in the gut: choose carefully. Nat. Rev. Gastroenterol. Hepatol. 2016;13(2):69-70. DOI 10.1038/nrgastro.2016.4.

Lozupone C.A., Stombaugh J.I., Gordon J.I., Jansson J.K., Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220-230. DOI 10.1038/nature11550.

Machiels K., Joossens M., Sabino J., De Preter V., Arijs I., Eeckhaut V., Ballet V., Claes K., Van Immerseel F., Verbeke K., Ferrante M., Ver-

haegen J., Rutgeerts P., Vermeire S. A decrease of the butyrate-producing species *Roseburia hominis* and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. Gut. 2014;63(8):1275-1283. DOI 10.1136/gutjnl-2013-304833.

Manichanh C., Borruel N., Casellas F., Guarner F. The gut microbiota in IBD. Nat. Rev. Gastroenterol. Hepatol. 2012;9(10):599-608. DOI 10.1038/nrgastro.2012.152.

Moayyedi P., Surette M.G., Kim P.T., Libertucci J., Wolfe M., Onischi C., Armstrong D., Marshall J.K., Kassam Z., Reinisch W., Lee C.H. Fecal microbiota transplantation induces remission in patients with active ulcerative colitis in a randomized controlled trial. Gastroenterology. 2015;149(1):102-109. DOI 10.1053/j.gastro.2015.04.001.

O'Hara A.M., Shanahan F. The gut flora as a forgotten organ. EMBO Rep. 2006;7(7):688-693.

Ohkusa T., Okayasu I., Ogihara T., Morita K., Ogawa M., Sato N. Induction of experimental ulcerative colitis by *Fusobacterium varium* isolated from colonic mucosa of patients with ulcerative colitis. Gut. 2003;52(1):79-83.

Paramsothy S., Kamm M.A., Kaakoush N.O., Walsh A.J., van den Bogaerde J., Samuel D., Leong R.W.L., Connor S., Ng W., Paramsothy R., Xuan W., Lin E., Mitchell H.M., Borody T.J. Multidonor intensive faecal microbiota transplantation for active ulcerative colitis: a randomised placebo-controlled trial. Lancet. 2017;389(10075):1218-1228. DOI 10.1016/S0140-6736(17)30182-30184.

Pigneur B., Sokol H. Fecal microbiota transplantation in inflammatory bowel disease: the quest for the holy grail. Mucosal Immunol. 2016;9(6):1360-1365. DOI 10.1038/mi.2016.67.

Qin J., Li R., Raes J., Arumugam M., Burgdorf K.S., Manichanh C., Nielsen T., Pons N., Levenez F., Yamada T., Mende D.R., Li J., Xu J., Li S., Li D., Cao J., Wang B., Liang H., Zheng H., Xie Y., Tap J., Lepage P., Bertalan M., Batto J.M., Hansen T., Le Paslier D., Linneberg A., Nielsen H.B., Pelletier E., Renault P., Sicheritz-Ponten T., Turner K., Zhu H., Yu C., Li S., Jian M., Zhou Y., Li Y., Zhang X., Li S., Qin N., Yang H., Wang J., Brunak S., Doré J., Guarner F., Kristiansen K., Pedersen O., Parkhill J., Weissenbach J., MetaHIT Consortium, Bork P., Ehrlich S.D., Wang J. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464(7285):59-65. DOI 10.1038/nature08821.

Reddy S.S., Brandt L.J. *Clostridium difficile* infection and inflammatory bowel disease. J. Clin. Gastroenterol. 2013;47(8):666-671. DOI 10.1097/MCG.0b013e31828b288a.

Rossen N.G., Fuentes S., van der Spek M.J., Tijssen J.G., Hartman J.H., Duflou A., Löwenberg M., van den Brink G.R., Mathus-Vliegen E.M., de Vos W.M., Zoetendal E.G., D'Haens G.R., Ponsioen C.Y. Findings from a randomized controlled trial of fecal transplantation for patients with ulcerative colitis. Gastroenterology. 2015;149(1):110-118. DOI 10.1053/j.gastro.2015.03.045.

Saebo A., Vik E., Lange O.J., Matuszkiewicz L. Inflammatory bowel disease associated with *Yersinia enterocolitica* O:3 infection. Eur. J. Intern. Med. 2005;16(3):176-182.

Shen Z.H., Zhu C.X., Quan Y.S., Yang Z.Y., Wu S., Luo W.W., Tan B., Wang X.Y. Relationship between intestinal microbiota and ulcerative colitis: mechanisms and clinical application of probiotics and fecal microbiota transplantation. World J. Gastroenterol. 2018;24(1):5-14. DOI 10.3748/wjg.v24.i1.5.

Sonnenberg A., Genta R.M. Low prevalence of *Helicobacter pylori* infection among patients with inflammatory bowel disease. Aliment. Pharmacol. Ther. 2012;35(4):469-476. DOI 10.1111/j.1365-2036.2011.04969.x.

Staley C., Hamilton M.J., Vaughn B.P., Graiziger C.T., Newman K.M., Kabage A.J., Sadowsky M.J., Khoruts A. Successful resolution of recurrent *Clostridium difficile* infection using freeze-dried, encapsulated fecal microbiota; pragmatic cohort study. Am. J. Gastroenterol. 2017;112(6):940-947. DOI 10.1038/ajg.2017.6.

van Nood E., Vrieze A., Nieuwdorp M., Fuentes S., Zoetendal E.G., de Vos W.M., Visser C.E., Kuijper E.J., Bartelsman J.F., Tijssen J.G., Speelman P., Dijkgraaf M.G., Keller J.J. Duodenal infusion of donor feces for recurrent *Clostridium difficile*. N. Engl. J. Med. 2013;368(5):407-415. DOI 10.1056/NEJMoa1205037.

Vaughn B.P., Vatanen T., Allegretti J.R., Bai A., Xavier R.J., Korzenik J., Gevers D., Ting A., Robson S.C., Moss A.C. Increased intestinal microbial diversity following fecal microbiota transplant for active Crohn's disease. Inflamm. Bowel Dis. 2016;22(9):2182-2190. DOI 10.1097/MIB.0000000000000893.

Wang Y., Qian P.Y. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One. 2009;4:e7401. DOI 10.1371/journal.pone.0007401.

ORCID ID

A.Y. Tikunov orcid.org/0000-0001-5613-5447
N.V. Tikunova orcid.org/0000-0002-1687-8278

Acknowledgements. This work was supported by the Russian Science Foundation, project 18-74-00082; and the Program of Basic Research for Governmental Academies of Sciences 2013-2020, project AAAA-A17-117020210026-2.

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

Received April 22, 2019. Revised June 10, 2019. Accepted June 23, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):176-184
DOI 10.18699/VJ20.611

The current state of the problem of *in vitro* gene pool preservation in poultry

Y.L. Silyukova, O.I. Stanishevskaya ✉, N.V. Dementieva

Russian Research Institute of Farm Animal Genetics and Breeding – Branch of the L.K. Ernst Federal Science Center for Animal Husbandry, Pushkin, St. Petersburg, Russia
✉ e-mail: olgastan@list.ru

Abstract. This review presents the current progress in and approaches to *in vitro* conservation of reproductive cells of animals, including birds, such as cryopreservation and freeze-drying, as well as epigenetic conditions for restoring viable spermatozoa and female gametes after conservation. Cryopreservation is an effective way to preserve reproductive cells of various species of animals and birds. *In vitro* gene pool conservation is aimed primarily to the restoration of extinct breeds and populations and to the support of genetic diversity in populations prone to genetic drift. It is the combination of *ex situ in vivo* and *ex situ in vitro* methods that can form the basic principles of the strategy of animal genetic diversity preservation. Also, use of cryopreserved semen allows faster breeding in industrial poultry farming. Despite numerous advances in semen cryobiology, new methods that can more efficiently restore semen fertility after cryopreservation are being sought. The mechanisms underlying the effect of cryopreservation on the semen parameters of cocks are insufficiently understood. The review reflects the results of recent research in the field of cryopreservation of female and male germ cells, embryonic cells, the search for new ways in the field of genetic diversity *in vitro* (the development of new cryoprotective media and new conservation technologies: freeze-drying). Molecular aspects of cryopreservation and the mechanisms of cryopreservation influence on the epigenetic state of cells are highlighted. Data on the results of studies in the field of male reproductive cell lyophilization are presented. The freeze-drying of reproductive cells, as a technology for cheaper access to the genetic material of wild and domestic animals, compared to cryopreservation, attracts the attention of scientists in Japan, Israel, Egypt, Spain, and France. There is growing interest in the use of lyophilized semen in genetic engineering technologies. Methods of freeze-drying are developed taking into account the species of birds. Organizational and legal ways of solving the problems of *in vitro* conservation of genetic resources of farm animals, including birds, are proposed.
Key words: semen; spermatozoa; marker proteins; cryopreservation; freeze-drying; poultry; gene pool preservation; cryobank; cryoresistance.

For citation: Silyukova Y.L., Stanishevskaya O.I., Dementieva N.V. The current state of the problem of *in vitro* gene pool preservation in poultry. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):176-184. DOI 10.18699/VJ20.611

Современное состояние проблемы сохранения генетических ресурсов сельскохозяйственных птиц *in vitro*

Ю.Л. Силюкова, О.И. Станишевская ✉, Н.В. Дементьева

Всероссийский научно-исследовательский институт генетики и разведения сельскохозяйственных животных – филиал Федерального научного центра животноводства – ВИЖ им. академика Л.К. Эрнста, Пушкин, Санкт-Петербург, Россия
✉ e-mail: olgastan@list.ru

Аннотация. Настоящий обзор представляет современные достижения и подходы по сохранению репродуктивных клеток животных *in vitro*, такие как криоконсервация и лиофилизация, а также эпигенетические предпосылки для получения жизнеспособных сперматозоидов и женских гамет после реконсервации. Криоконсервация – эффективный путь сохранения репродуктивных клеток различных видов сельскохозяйственных животных, включая птиц. Метод сохранения генофонда *in vitro* через поддержание в криогенных условиях клеток или тканей в основном направлен на восстановление исчезнувших пород/популяций, на поддержание генетического разнообразия в популяциях, подверженных генетическому дрейфу. Именно сочетание методов *ex situ in vivo* и *ex situ in vitro* может сформировать основу эффективной стратегии сохранения генетического разнообразия животных. Кроме того, использование криконсервированного семени лучших представителей линии или породы позволяет ускорить прогресс селекции в промышленном птицеводстве. Несмотря на многочисленные достижения в области криобиологии половых клеток, продолжается поиск методов, обеспечивающих более эффективное восстановление жизнеспособности спермиев после криоконсервации. Механизмы, лежащие в основе влияния процедуры криоконсервации на параметры семени сельскохозяйственных птиц, полностью не изучены. В обзоре отражены результаты со-

временных исследований в области проблематики криоконсервации женских и мужских половых клеток, эмбриональных клеток, поиска новых путей решения в области сохранения генетического разнообразия *in vitro* (разработка новых криопротекторных сред и новых технологий сохранения). Освещены молекулярно-генетические аспекты криоконсервации и механизмы влияния криоконсервации на эпигенетическое состояние клеток. Представлены данные по результатам исследований в области лиофильной сушки репродуктивных клеток самцов. Интерес к технологии лиофилизации семени как возможности более дешевого способа сохранения и транспортировки генетического материала диких и домашних животных, по сравнению с криоконсервацией, в мире стремительно растет; исследования ведутся в Японии, Израиле, Египте, Испании, Франции. Растет и интерес к использованию лиофилизированного семени в технологиях генной инженерии. Методы лиофильной сушки разрабатываются с учетом видовой принадлежности. В обзоре предложены также организационно-правовые пути решения проблемы сохранения генетических ресурсов сельскохозяйственных животных, включая птиц, *in vitro*.
Ключевые слова: семя; сперматозоид; маркерные белки; криоконсервация; лиофилизация; птицеводство; сохранение генетических ресурсов; криобанк; криорезистентность.

Introduction

The conservation of genetic resources of farm animals is a global challenge, and it attracts efforts of the world community. The Food and Agriculture Organization (FAO) of the United Nations and its specialized units coordinate these efforts (FAO, 2015).

Conservation programs for genetic resources include the following tasks: economyc (livestock husbandry support, response to changes in the environment, market needs, regulatory requirements, and the availability of imports and exports); social and cultural issues; conservation of biodiversity; and the maintenance of resources for academic or educational purposes, genetics, genomics, and adaptation to climate and other environmental changes.

The method of *ex situ in vitro* gene pool preservation through cryogenic maintenance (cryobank) of cells or tissues that can be used for breed/population restoration is recognized a necessary supplement to the *in vivo* method (FAO, 2015). It is the combination of *ex situ in vivo* and *ex situ in vitro* methods that can form grounds for an effective strategy for preserving animal genetic diversity.

Methods developed to freeze reproductive cells of male farm birds can be successfully applied to wild species to preserve their genetic diversity: red jungle chicken (Rakhaa et al., 2016), capercaillie (Kowalczyk et al., 2012), and pheasant (Saint Jaime et al., 2003). Due to the significant decrease in genetic diversity in pure lines of industrial crosses (Muir et al., 2008), the use of cryopreserved semen of the best representatives of a line or breed in artificial insemination under conditions of poultry industry expands the range of variation and accelerates breeding.

Semen preservation

At present, cryopreservation of reproductive cells of males is the most important, practically the only method of preserving the gene pool of farm birds *in vitro*. Various protocols have been developed for the conservation of poultry semen, and the effectiveness of their use depends on many factors (Tselyutin, Tour, 2013; Thélie et al., 2019). The problem of reduced functional ability of the semen after a freeze-thaw cycle has not yet been resolved; the level of the fertilizing ability of thawed semen is not satisfactory. According to different authors, depending on the freezing methods, individual and breed characteristics of chickens, egg fertilization varies from 2 to 85 % (Blesbois et al., 2007; Long et al., 2010; Seigneurin, Blesbois, 2010; Çiftci, Aygün, 2018). The average level of fertilization with cryopreserved semen is low, usually less than 30 % (Fulton, 2006); however, some recent publications show 65 % average fertility of frozen-thawed semen (Silyukova et al., 2019). The reduced viability of embryos derived from cryopreserved semen caused by DNA fragmentation (Watson, 2000; Liptói, Hidas, 2006; Morris et al., 2012) also compromises the economic feasibility of its use for practical breeding purposes. Therefore, works aimed at the improvement of the composition of diluents for cryopreservation, the selection of cryoprotectant and freezing methods (in straws or pellets), the freezing protocols (low/fast), and so on are still under way (Thieu Ngoc Lan Phuong et al., 2014; Svoradová et al., 2017).

Most studies on cryopreservation of avian semen are conducted with mixed ejaculates from several males, although it is known that the genetic contributions of males differ due to the effect of selective fertilization (Sakharova, Popov, 2001), and males have different quality indicators of sperm after cryopreservation (Pleshanov et al., 2018, 2019). Therefore, there are concerns that the use of cryopreserved sperm may lead to an increase in inbreeding when using mixed sperm. In order to avoid this problem, when preserving rare and endangered breeds of chickens, it is necessary that the cryobank store individual ejaculates.

In Russia, studies in this direction are being conducted at the L.K. Ernst Federal Science Center for Animal Husbandry and its branch, the Russian Research Institute of Farm Animal Genetics and Breeding (Iolchiev et al., 2018; Mavrodina et al., 2018a, b; Pleshanov, Stanishevskaya, 2018; Stanishevskaya, Pleshanov, 2018a–c).

Technologies. Many scientific publications describe different protocols for cryopreservation of semen of domesticated and wild bird species. Technologies differ in the type of cryoprotectant, the method of packaging (straw, pellets, etc.), the speed of freezing and thawing (fast/slow) and temperature regimes. The effectiveness of the protocols can be assessed by analyzing the semen functional state in the laboratory (determination of sperm concentration, motility, morphology, and live/dead ratio) and assessing the fertilizing capacity of sperm in artificial insemination *in vivo* (Varadi et al., 2013; Thieu Ngoc Lan Phuong et al., 2014). It was found that high rates of avian semen freezing–thawing were preferable in

terms of improving its survival in contrast to the protocol for mammalian semen (Shahverdi et al., 2015; Madeddu et al., 2016). There is significant individual, intrabreed, and intraspecific variation in bird semen cryostability, which demands that different cryopreservation strategies for different species and breeds be developed (Blesbois et al., 2007).

Methods of semen quality assessment. Semen cryopreservation is very important for *ex situ* management of avian genetic diversity, but the use of this method is limited due to the high variability of success rates. To calculate the number of sperm doses in the formation of the cryobank, it is necessary to predict the fertilizing capacity of cryopreserved semen. Unfortunately, in determining the effectiveness of their development, many researchers are limited only to the assessment of sperm motility. This test is not informative enough in terms of predicting the fertilizing capacity of semen.

A more effective prognosis for the fertilizing capacity of the semen is provided by assessment of morphological disorders including fluorescent staining of living and dead cells, flow cytometry, and evaluation of spermatozoon motility parameters using computer-assisted sperm analysis (CASA). The CASA system permits one to estimate the percentage of viable and morphologically normal cells (PVN), mass mobility (MMOT) and various parameters of movement, including the percentage of motile sperm (PMOT) and biophysical tests (resistance to osmotic stress (OSM), membrane permeability (FLUID)) (Blesbois et al., 2008; Svoradová et al., 2018). However, the set of these tests does not fully reflect the functionality of semen.

The functional capacity of thawed semen can be reliably determined *in vitro* by analyzing the interaction of the sperm with the inner layer of the perivitelline layer of the egg yolk (Robertson et al., 1997; Long et al., 2010). In evaluation of the functional state of spermatozoa *in vitro* in the laboratory, it is advisable to use the perivitelline membrane of egg yolk. The assessment is based on the number of hydrolysis points (spermatozoon penetrations) per unit area of the inner perivitelline layer (Robertson et al., 1997). This method, compared to traditional quality assessment, is more informative for predicting the fertilizing capacity of semen.

Preservation of female gametes. By now, no method has been developed to preserve germ cells of female birds. The presence of large amounts of yolk in the eggs of birds hampers the use of existing cryopreservation methods (Fulton, 2006). This is a serious issue in the preservation of a breed/population, as it cannot be fully preserved without the genetic contribution of individuals of both sexes: there is a loss of maternal hereditary material, including the mitochondrial genome. Currently existing methods of preservation of reproductive cells of birds (semen) allow restoration of endangered breeds/populations only by grading.

A relatively new technology is the transplantation of cryopreserved gonad cells from neonatal chickens to adult recipients for reproduction of donor offspring. This method of transplantation can contribute to the conservation of endangered bird species and maintain their genetic variability (Benesova, Trefil, 2016). Cryopreservation of ovarian tissue is actually the only effective way to preserve *in vitro* female germ plasma in birds. A method of vitrification of donor ovarian tissues in straw is proposed. Ovarian tissue fragments from 1-week old females are transferred to a metal rod, then vitrified in liquid nitrogen using special media. By this method, the ovarian tissue of birds can be stored and transported. Successful transplant survival was demonstrated by Liu et al. (2012). In addition, ovarian transplantation can be used for research in genetics and developmental biology (Song, Silversides, 2007; Liptoi et al., 2013). Since this technique involves significant surgery and requires the use of immunosuppressants, today it appears to be costly and technologically challenging in routine implementation.

Preservation of embryonic cells. Chicken primary germ cells (PGCs) can be isolated and cultured *in vitro*. PGCs act as a valuable source material for cellular genetic engineering, germ plasm production, and genetic conservation of species and populations (Kino et al., 1997). Indeed, bird PGCs can be reproduced in culture and conserved without irreversible alteration of their biological properties (van der Lavoir et al., 2006; Nandi et al., 2016; Tonus et al., 2016). Two main methods are commonly used for cryopreservation of PGCs in birds: slow freezing (SLF) and ultrafast freezing (Vitrif) (Tonus et al., 2017). These cells can be used to repair gonadal tissues with germ cells of the donor line. This method is not currently applicable for the preservation of entire embryos (Fulton, 2006). Both methods require further research, but we can now definitely state that in the future this approach to *in vitro* preservation of cells will provide grounds for the development of a practical gene bank and systematic genomic bank for birds.

Genetics of cryostability of reproductive cells

It has been established that semen cryostability is a genetically determined trait (Pleshanov et al., 2019), but the mechanisms of the effect of cryopreservation on the epigenetic state of cells have not yet been fully investigated. Semen freezing–thawing can injure genes, including *SNORD116/PWSAS* and *UBE3A*, associated with fertility (Valcarce et al., 2013).

The study of changes in boar semen after cryopreservation revealed differences in 41 proteins (Chen et al., 2014). Proteins SOD1, TPI1, ODF2, and AKAP3 have been proposed as markers affecting semen resistance to freezing. In *Gallus gallus domesticus*, ontology genes were found for SOD1, TPI1, and ODF2. Consider these genes in more detail. The protein encoded by the SOD1 gene (Superoxide dismutase 1) binds copper and zinc ions and promotes the breakdown of superoxide radicals into molecular oxygen and hydrogen peroxide (Bogle et al., 2017; Wu, 2019). Another isozyme of this protein is found in mitochondria, and its functions have not yet been studied. The enzyme TPI1 (Triosephosphate isomerase 1), which consists of two identical proteins, catalyzes the isomerization of glyceraldehyde 3-phosphate (G3P) and dihydroxyacetone phosphate (DHAP) in glycolysis and gluconeogenesis (Chen et al., 2014).

It has been found that HSP90 (Heat shock protein 90) proteins are associated with sperm motility, and their pool decreases significantly after freezing–thawing (Huang et al., 2009).

Significant protein changes in human sperm before and after cryopreservation were detected by Wang et al. (2014): mito-

chondrial matrix proteins ACO2 (Aconitase 2) and OXCT1 (3-oxoacid CoA-transferase 1); filamentous protein TEKT1 (Tektin 1), which is necessary for the formation of ciliary and flagellar microtubules; glycolytic enzyme ENO1 (Enolase 1), the intermediate filament protein vimentin; and the amino acid tyrosine. These molecules are associated with sperm motility, viability, and acrosome integrity (Wang et al., 2014).

As a result of semen freeze–thawing, the amounts of antioxidant proteins such as SOD1, PRDX6 (Peroxiredoxin 6), TXNDC2 (Thioredoxin containing domain 2), GSTM3 (Glutathione-S-transferase mu 3), membrane proteins CYB5R2 (Cytochrome b5 reductase 2), pellucid zone proteins ZPBP1 and ZPBP2 (Zona pellucida binding protein), acrosomal proteins ACRBP (Acrosin binding protein), and SPACA3 (Sperm acrosome associated 3) were found to decrease. Simultaneously, the amounts of other proteins whose accumulation is observed in cells under stress – ANX1, ANX3, and ANX4 (Annexin A); clusterin (CLU Clusterin); importin-1b (KPNB1); Karyopherin subunit beta 1, HIST1H4A (Histone cluster 1 H4 family member a); TUBA1A (Tubulin alpha 1a); and SPAG17 (Sperm associated antigen 17) increased (Bogle et al., 2017).

A study of the effect of cryopreservation on *Gallus gallus domesticus* spermatozoa has shown an increase in the amounts of 36 proteins and a decrease in 19 proteins after thawing. These proteins are linked to spermatozoa metabolism (Cheng et al., 2015). Proteins such as ACRBP, FN1 (Fibronectin 1), HSP90AA1 (Heat shock protein 90), and VDAC2 (Voltage dependent anion channel 2) are biomarkers that predict tolerance of cryopreservation in boar semen (Vilagran et al., 2015, 2016).

During fertilization, spermatozoa deliver the paternal mRNA to the egg and thus play an important role in the early development of the embryo. During freezing, transcripts and mRNA–protein interactions in spermatozoa may be lost, which may affect embryo development (Valcarce et al., 2013). Correlations between sperm mRNA and early embryo development in humans and some animals were identified in (Hezavehei et al., 2018). Studies by Valcarce et al. (2013) showed a decrease in the expression of the PRM1, PRM2, PEG1/MEST, and ADD1 genes associated with human sperm fertility after cryopreservation. Some studies confirmed changes in transcripts of some proteins and micro-RNA. Attempts are being made to explain some epigenetic modifications that may occur in spermatozoa during freezing (Hezavehei et al., 2018).

Cryopreservation of semen is a very important method of assisted reproduction, but the freezing–thawing procedure is harmful, because it leads to a decrease in the motility and viability of spermatozoa, premature capacitation, and, as a consequence, poorer effectiveness of artificial fertilization. Therefore, the addition of some proteins normalizes the process of condensation and accelerates fertilization *in vitro*. For example, use of TrxA-FNIIx4-His6 is a promising biotechnological approach for cryopreservation of ram semen and maintenance of spermatozoon viability (Ledesma et al., 2019).

In addition to preserving genetic material by creating a sperm bank, it is possible to create a cryobank of embryos. In cattle, the effect of resveratrol on embryos after cryopreservation was evaluated. Its effects on mitochondrial function preservation, DNA integrity, SIRT1 (Sirtuin 1) expression and embryo development ability have been studied. Embryo survival was significantly improved when embryos were incubated in a medium containing 0.5 μM of resveratrol after thawing. Besides, SIRT1 expression and cell-free mtDNA content in the medium were higher in the case of embryos treated with resveratrol. It should be noted that slow freezing affects mitochondrial integrity and function in blastocysts (Hayashi et al., 2019). It is important to improve *in vitro* maturation (IVM) conditions for immature oocytes after cryopreservation, especially if a limited number of oocytes are collected from specific donors. Culture systems with fresh oocytes significantly accelerate the meiotic development of vitrified oocytes and significantly increase the rate of blastocyst formation after parthenogenetic activation and transfer of somatic cell nuclei (Jia et al., 2019).

The understanding of the molecular mechanisms that determine epigenetic processes occurring in reproductive cells during freezing–thawing will improve the effectiveness of the technologies used to preserve species, breeds, and populations of rare and endangered animals and birds.

Lyophilization

The preservation of semen by freeze-drying is an innovative method. The advantages of lyophilized semen are that it can be (1) stored at 4 °C for a long time and (2) stored and transported at room temperature without the use of liquid nitrogen or dry ice as cooling agents.

It is expected that sperm lyophilization, rather than cryopreservation, can become a new simple method of preserving genetic resources and be used, among other things, to produce transgenic animals (Kaneko, 2012). The state of research in the field of freeze-drying of wild and domestic animal semen indicates an increasing interest in this method of preserving genetic resources. Methods of lyophilization in relation to microorganisms and plant cells have been developed and successfully applied. Interest in the lyophilization of reproductive cells, as a possibility of a cheaper way to preserve and transport (including space) genetic material of wild and domestic animals, compared with cryopreservation, is growing rapidly in the world; research is underway in Japan, Israel, Egypt, Spain, and France. Methods of lyophilic drying are developed with regard to species features. Promising results have been achieved in mice, rats, hamsters, cattle, sheep, rabbits, chimpanzees, giraffes, jaguars, etc., but it is too early to talk about the problem as solved, since the functional characteristics of sperm are not fully preserved (Hopshi et al., 1994; Foote, 2002; Liu et al., 2004; Kawase et al., 2005; Li et al., 2009; Gil et al., 2014; Kaneko et al., 2014; Shahba et al., 2016; Wakayama et al., 2017; Arav et al., 2018). The main issues are associated with damage of the motility apparatus of spermatozoa, membranes, and DNA. As for birds, including poultry, research on the freeze-drying of their semen has not been carried out, at least not published.

Problems of cryopreservation

Cryopreservation triggers damage processes not only at the mechanical level of membrane damage, but also chemical and physical processes of denaturation of proteins and lipids of membrane bilayers. These processes result in sublethal freez-

ing and the launch of cryocapacitation, generation of reactive oxygen species, and aberrations in sperm proteins, lipids, and sugars (Pini et al., 2018).

It is well known that the tolerance of cold shock and cryostability in spermatozoa of different livestock species, including farm birds, varies greatly. Cryopreserved semen of any animal species has reduced fertility compared to fresh sperm. The causes of fertility loss include the susceptibility to cold shock, cooling rate, diluent composition, and osmotic stress. There are also factors that affect the functional state of frozen/thawed spermatozoa: membrane stability, oxidative damage, membrane receptor integrity, and nucleus structure (Watson, 2000; Iolchiev et al., 2018). In the course of cryopreservation and thawing, sperm can experience both irreversible damage, expressed as the absence of motility and various morphological disorders, and reversible, associated mainly with a temporary injury of the structure and membrane permeability disturbance.

It is believed that the high content of intracellular protein together with osmotic "shrinkage" of the sperm membrane associated with the formation of extracellular ice leads to intracellular vitrification of sperm during cooling. At high cooling rates, sperm damage is the result of osmotic imbalance occurring during thawing, rather than intracellular ice formation during freezing. Osmotic imbalance occurs at high cooling rates due to limited diffusion of ice crystallization in the extracellular fluid; that is, the amount of ice formed during cooling is less than expected from the phase equilibrium diagram (Morris et al., 2012).

A significant interbreeding variability in the cryostability of cock semen, estimated by the activity of thawed semen, is known: the coefficient of variation (C_v) can be up to 23–25 % (Pleshanov et al., 2018; Stanishevskaya, Pleshanov, 2018a). A greater individual variability in cock spermatozoon activity in the freeze/thaw cycle has been shown in (Pleshanov et al., 2017; Pleshanov, Stanishevskaya, 2018; Stanishevskaya, Pleshanov, 2018a, b). The coefficient of variation (C_v) of spermatozoon activity was 6.1 % in native sperm and 19.5 % in frozen/thawed, which points to a broad norm of sperm response to low temperatures.

The generally accepted parameters of ejaculate selection for cryopreservation are volume, concentration, and spermatozoon motility. These criteria do not provide a complete prediction of the degree of reproductive cell cryotolerance, which is largely due to the state of the membranes, as the membranes are first to be damaged in the freezing–thawing process.

One of the ways to assess the degree of cryopreservation of spermatozoon membranes is staining with Sperm VitalStain dye (Nidacon International AB, Sweden), which allows assessment of the degree of cryopreservation by changing the color of damaged cells (Pleshnov, Stanishevskaya, 2018). Lipid fractions of membranes, such as glycolipids, phospholipids, sterols, cholesterol, the cholesterol/phospholipids ratio, etc., affect the state of cell membranes; their permeability, microviscosity, and fluidity; molecular mobility of lipids in the membrane; the process of capacitation, the interaction of egg and spermatozoon membranes; and the result of fertilization (Blesbois et al., 2005; Ahmed et al., 2014; Eubaid et al., 2015; Partyka et al., 2016; Pleshanov et al., 2017).

Recent studies of sperm cryostability have established the effect of the amino acid profiles of seminal plasma in different breeds of chickens on DNA fragmentation (Santiago-Moreno et al., 2019). The composition of intracellular spermatozoon protein has been found to be associated with indicators of osmotic imbalance after thawing (Morris et al., 2012). The results of these studies open up new aspects of sperm cryobiology, which is a prerequisite for the development of new technologies for semen preservation, including vitrification and lyophilization.

Problems of early embryonic mortality. It is well known that the use of frozen/thawed semen reduces not only the percentage of their fertility but also the viability of embryos. The mortality rate of embryos in the early stages of development can reach 8–17 % (Stanishevskaya, Pleshanov, 2018c). This research area is insufficiently studied, as it is technically difficult to investigate the causes of arrested development, since signs of early embryonic death are not determined. DNA damage is probably a major cause of early embryonic mortality caused by functional damage to sperm nuclear structures (Watson, 2000; Liptói, Hidas, 2006). In addition, the influence of toxic endo/exocellular cryoprotectants used in sperm freezing and their concentrations, which may also cause embryo death at an early stage of development, should not be ruled out (Mosca et al., 2019).

Thus, the genetic diversity of the preserved material is reduced at different stages of postsingamy due to the elimination of individuals with reduced cryoresistance of reproductive cells.

Cryoprotectants. A necessary condition for successful cryopreservation of reproductive cells is the use of cryoprotectants. Cryoprotectants acting inside cells penetrate into cells and prevent the formation of intracellular ice, but at high concentrations they exert a damaging effect. Exocellular cryoprotectants act outside the sperm in the extracellular space and protect cells by dehydrating the intracellular space and limiting the action of osmotic shock during thawing.

Glycerol, one of the best known cryoprotectants, is the most effective and less toxic to cock sperm, but unfortunately has a contraceptive effect after insemination of hens and requires removal before insemination. The most widely used penetrating cryoprotectants are dimethylsulfoxide, dimethylacetamide, dimethylformamide and ethyleneglycol. Semen samples can be thawed without further processing, and high fertility levels are obtained with these substances depending on the cooling rate and the type of semen packaging (Santiago-Moreno et al., 2011). Nonpenetrating cryoprotectants, also known as osmoprotectants, are low-molecular-weight hydrophilic nontoxic molecules that stabilize internal solutes under osmotic stress in cells. These cryoprotectants are often used in combination with penetrators (Benesona, Trefil, 2016; Mosca et al., 2016; Svoradová et al., 2017).

Recent approaches to the development of cryoprotectants of fundamentally different actions are based on antifreeze glycoproteins (AFGP) and antifreeze proteins (AFP), found in the blood and tissues of poikilothermic organisms living in freezing environments (insects and marine fish). The obtained substances inhibit the growth of ice crystals in a noncolligative manner. The use of AFP opens up a promising direction

for cryopreservation of living tissues and cells. The efficacy of some fish AFP or fish AFGP against hypothermic damage has been reported while preserving swine and cattle oocytes, whole rat liver, and model membranes. To preserve sperm, attempts have been made to develop cryopreservation methods with the addition of fish AFP in different species with different efficiencies. AFP and AFGP of marine fish have recently been found to improve buffalo sperm cryopreservation results (Qadeer et al., 2016).

There are studies on the use of recombinant AFP-based *Dendroides canadensis* (DAFP) larvae for cryopreservation. Addition of DAFP to the diluent protects buffalo (*Bubalus bubalis*) semen during freezing–thawing and increases the fertility of cryopreserved semen (Qadeer et al., 2016).

Cryobanks and their contribution to the conservation of genetic resources

Collections of genetic banks are of paramount importance in preventing the extinction of breeds due to extreme genetic conditions, such as small breed/population size and high incidence of genetic defects as a result of intensive breeding and genetic drift. Stored material from animals that do not carry undesired or lethal mutations can be used to reduce the frequency of defects to an acceptable level.

Biobanks are a ready-to-use source of genetically diverse and specialized DNA. Conserved materials are used in studying genetic diversity, genomic associations, gene functions, and other issues. Importantly, over time, genetic banks can provide samples from different generations, which contributes to the accuracy of genomic selection. These advantages will be easier to implement if the information is catalogued taking into account the phenotype and genotype and the stored samples have genomic certificates (Wildt, 2000; Comizzoli, 2015).

The problem of *in vitro* preservation of genetic diversity, including farm birds, is being solved in many countries of the world. One of the advantages of preserving genetic diversity *in vitro* in cryobanks is the economic component (Woelders, 2006; Santiago-Moreno et al., 2011; Silversides et al., 2013; FAO, 2015).

Recently, a new approach to interaction between organizations preserving the gene pool *ex situ in vivo* and *ex situ in vitro* has been developed. The goal of the genetic bank is not only to obtain and preserve reserve biological material but also to closely cooperate with collections in live breeding to expand genetic diversity while preserving *ex situ in vivo*.

Genetic banks for the conservation of farm birds can take into account the experience of the European Union, the European Federation of Animal Science (EAAP), and FAO, which have established European and international consultative forums to discuss and take practical measures for the conservation of genetic resources worldwide. However, the implementation of this idea is a complex process, which requires interdisciplinary cooperation and clear definition of goals (Mara et al., 2013).

The legislation of the Russian Federation provides for a regulatory framework (Strategy) for the conservation of rare and endangered species of animals, plants and fungi (Order No. 212-r of the Government of the Russian Federation of 17.02.2014), including preservation *in vitro*. As for the problem of preserving the genetic diversity of farm animals and birds, the Federal law No. 123-FZ "On livestock breeding" of 03.08.1995 does not provide for such regulation. It is necessary to develop and adopt a law and by-laws that would determine the legal status of genetic cryobanks in the overall system of genetic resource conservation.

References

Ahmed M., Sattar A., Iqbal S., Shahzad Q., Zahid Tahir M., Hammad Fayyaz M., Ahmad M. Cholesterol loaded cyclodextrin improves sperm survival in tris-based extender with less egg yolk used in buffalo bulls semen. Thai J. Vet. Med. Suppl. 2014;44:129-130.

Arav A., Idda A., Nieddu S., Natan Y., Ledda S. High post-thaw survival of ram sperm after partial freeze-drying. J. Assist. Reprod. Genet. 2018;35(7):1149-1155. DOI 10.1007/s10815-018-1145-1.

Benesova B., Trefil P. Possibilities for preserving genetic resources in birds. World's Poult. Sci. J. 2016;72(3):628-641. DOI 10.1017/S0043933916000489.

Blesbois E., Grasseau I., Segineurin F. Membrane fluidity and the ability of domestic bird spermatozoa to survive cryopreservation. Reproduction. 2005;129(3):371-378. DOI 10.1530/rep.1.00454.

Blesbois E., Grasseau I., Seigneurin F., Mignon-Grasteau S., Saint Jalme M., Mialon-Richard M.M. Predictors of success of semen cryopreservation in chickens. Theriogenology. 2008;69(2):252-261. DOI 10.1016/j.theriogenology.2007.09.019.

Blesbois E., Seigneurin F., Grasseau I., Limouzin C., Besnard J., Gourichon D., Coquerelle G., Rault P., Tixier-Boichard M. Semen cryopreservation for ex situ management of genetic diversity in chicken: creation of the French avian cryobank. Poult. Sci. 2007;86(3):555-564. DOI 10.1093/ps/86.3.555-564.

Bogle O.A., Kumar K., Attardo-Parrinello C., Lewis S.E., Estanyol J.M., Ballescà J.L., Oliva R. Identification of protein changes in human spermatozoa throughout the cryopreservation process. Andrology. 2017;5:10-22. DOI 10.1111/andr.12279.

Chen X., Zhu H., Hu C., Hao H., Zhang J., Li K., Zhao X., Qin T., Zhao K., Zhu H., Wang D. Identification of differentially expressed proteins in fresh and frozen-thawed boar spermatozoa by iTRAQ-coupled 2D LC-MS/MS. Reproduction. 2014;147(3):321-330. DOI 10.1530/REP-13-0313.

Cheng C.Y., Chen P.R., Chen C.J., Wang S.H., Chen C.F., Lee Y.P., Huang S.Y. Differential protein expression in chicken spermatozoa before and after freezing-thawing treatment. Anim. Reprod. Sci. 2015;152:99-107. DOI 10.1016.anireprosci.2014.11.011.

Çiftci Y., Aygün A. Poultry semen cryopreservation technologies. World's Poult. Sci. J. 2018;74(4):699-710. DOI 10.1017/s0043933918000673.

Comizzoli P. Biobanking efforts and new advances in male fertility preservation for rare and endangered species. Asian J. Androl. 2015;17:640-645. DOI 10.4103/1008-682X.153849.

Eubaid H.J., Al-Haidary B.A., Tawfiq L.J. Possible role for cholesterol in human seminal fluidin relations to other semen parameters & fertility. J. Babylon Univ./Pure Appl. Sci. 2015;23(2):654-659.

FAO. 2015. The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture. Ed. by B.D. Scherf, D. Pilling. FAO Commission on Genetic Resources for Food and Agriculture Assessments. Rome, 2015. Available at: http://www.fao.org/3/a-i4787e/index.html

Foote R. The history of artificial insemination; selected notes and notables. J. Anim. Sci. 2002;80:1-10.

Fulton J.E. Avian genetic stock preservation: an industry perspective. Poult. Sci. 2006;85:227-231.

Gil L., Olaciregui M., Luno V., Malo C., González N., Martínez F. Current status of freeze-drying technology to preserve domestic animals sperm. Reprod. Domest. Anim. 2014;49(4):72-81. DOI 10.1111/rda.12396.

Hayashi T., Kansaku K., Abe T., Ueda S., Iwata H. Effects of resveratrol treatment on mitochondria and subsequent embryonic development of bovine blastocysts cryopreserved by slow freezing. Anim. Sci. J. 2019;90(7):849-856. DOI 10.1111/asj.13219.

Hezavehei M., Sharafi M., Kouchesfahani H.M., Henkel R., Agarwal A., Esmaeili V., Shahverdi A. Sperm cryopreservation: a review on current molecular cryobiology and advanced approaches. Reprod. BioMed. Online. 2018;37(3):327-339. DOI 10.1016/j.rbmo.2018.05.012.

Hopshi K., Yanagida K., Katayose H., Yazawa H. Pronuclear formation and cleavage of mammalian eggs after microsurgical injection of freeze-dried sperm nuclei. Zigote. 1994;2:237-242. DOI 10.1017/S0967199400002033.

Huang S.Y., Pribenszky C., Kuo Y.H., Teng S.H., Chen Y.H., Chung M.T., Chiu Y.F. Hydrostatic pressure pre-treatment affects the protein profile of boar sperm before and after freezing-thawing. Anim. Reprod. Sci. 2009;112:136-149. DOI 10.1016/j.anireprosci.2008.04.016.

Iolchiev B.S., Bagirov V.A., Zhilinsky M.A., Volkova N.A., Zinovieva N.A. Change of biological parameters of poultry semen at cryopreservation. Sel'skohozyaystvennaya Biologiya = Agricultural Biology. 2018;6(53):1230-1237. DOI 10.15389/agrobiology.2018.6.1230eng.

Jia B.Y., Xiang D.C., Zhang B., Quan G.B., Shao Q.Y., Hong Q.H., Wu G.Q. Quality of vitrified porcine immature oocytes is improved by coculture with fresh oocytes during in vitro maturation. Mol. Reprod. Dev. 2019;86:1615-1627. DOI 10.1002/mrd.23249.

Kaneko T. New possibilities of sperm freeze-drying. J. Fert. in Vitro. 2012;2:5. DOI 10.4172/2165-7491.1000e119.

Kaneko T., Ito H., Sakamoto H., Onuma M., Inoue-Murayama M. Sperm preservation by freeze-drying for the conservation of wild animals. PLoS One. 2014;9(11):e113381. DOI 10.1371/journal.pone.0113381.g001.

Kawase Y., Araya H., Kamada N., Jishage K., Suzuki H. Possibility of long-term preservation of freeze-dried mouse spermatozoa. Biol. Reprod. 2005;72(3):568-573. DOI 10.1095/biolreprod.104.035279.

Kino K., Pain B., Leibo S.P., Cochran M., Clark M.E., Etches R.J. Production of chicken chimeras from injection of frozen-thawed blastodermal cells. Poult. Sci. 1997;76(5):753-760. DOI 10.1093/ps/76.5.753.

Kowalczyk A., Lukaszewicz E.T., Rzońca Z. Successful preservation of capercaillie (*Tetrao urogallus* L.) semen in liquid and frozen states. Theriogenology. 2012;77:899-907. DOI 10.1016/j.theriogenology.2011.09.015.

Ledesma A., Zalazar L., Buchelly Imbachi F., Pastore J.I., Brown P., Eddy E.M., Hozbor F., Cesari A. Recombinant peptide reverses cryo-capacitation in ram sperm and improves in vitro fertilization. Anim. Reprod. Sci. 2019;207:61-72. DOI 10.1016/j.anireprosci.2019.05.016.

Li M.-W., Willis B.J., Griffey S.M., Spearow J.L., Lloyd K.C. Assessment of three generations of mice derived by ICSI using freeze-dried sperm. Zygote. 2009;17(03):239-251. DOI 10.1017/s0967199409005292.

Liptói K., Hidas A. Investigation of possible genetic background of early embryonic mortality in poultry. World's Poult. Sci. J. 2006;62(02):326-337. DOI 10.1079/wps2005101.

Liptoi K., Horvath G., Gal J., Varadi E., Barna J. Preliminary results of the application of gonadal tissue transfer in various chicken breeds in the poultry gene conservation. Anim. Reprod. Sci. 2013;141(1-2):86-89. DOI 10.1016/j.anireprosci.2013.06.016.

Liu J., Cheng K.M., Silversides F.G. Novel needle-in-straw vitrification can effectively preserve the follicle morphology, viability, and vascularization of ovarian tissue in Japanese quail (*Coturnix japonica*). Anim. Reprod. Sci. 2012;134(3-4):197-202. DOI 10.1016/j.anireprosci.2012.08.002.

Liu J.-L., Kusakabe H., Chang C.-C., Suzuki H., Schmidt D.W., Julian M., Pfeffer R., Bormann C.L., Tian X.C., Yanagimachi R., Yang X. Freeze-dried sperm fertilization leads to full-term development in rabbits. Biol. Reprod. 2004;70(6):1776-1781. DOI 10.1095/biolreprod.103.025957.

Long J.A., Bongalhardo D.C., Pelaez J., Saxena S., Settar P., O'Sullivan N.P., Fulton J.E. Rooster semen cryopreservation: effect of pedigree line and male age on post-thaw sperm function. Poult. Sci. 2010;89(5):966-973. DOI 10.3382/ps.2009-00227.

Madeddu M., Mosca F., Abdel Sayed A., Zaniboni L., Mangiagalli M.G., Colombo E., Cerolini S. Effect of cooling rate on the survival of cryopreserved rooster sperm: comparison of different distances in the vapor above the surface of the liquid nitrogen. Anim. Reprod. Sci. 2016;171:58-64. DOI 10.1016/j.anireprosci.2016.05.014.

Mara L., Casu S., Carta A., Dattena M. Cryobanking of farm animal gametes and embryos as a means of conserving livestock genetics. Anim. Reprod. Sci. 2013;38(1-2):25-38. DOI 10.1016/j.anireprosci.2013.02.006.

Mavrodina T., Stanishevskaya O., Cherepanov S., Silyukova Y. Influence of sperm quality (cryopreserved and native) on the duration of spermatozoa storage in reproductive tracts of turkeys. Anim. Reprod. Sci. 2018a;194:e13. DOI 10.1016/j.anireprosci.2018.04.034.

Mavrodina T., Stanishevskaya O., Cherepanov S., Silyukova Y. Influence of osmolality of the media for dilution and cryopreservation of turkey toms' sperm on fertilization ability of thawed sperm. Reprod. Domest. Anim. 2018b;53(2):164.

Morris G.J., Acton E., Murray B.J., Fonseca F. Freezing injury: the special case of the sperm cell. Cryobiology. 2012;64:71-80. DOI 10.1016/j.cryobiol.2011.12.002.

Mosca F., Madeddu M., Abdel Sayed A., Zaniboni L., Iaffaldano N., Cerolini S. Combined effect of permeant and non-permeant cryoprotectants on the quality of frozen/thawed chicken sperm. Cryobiology. 2016;73(3):343-347. DOI 10.1016/j.cryobiol.2016.10.001.

Mosca F., Zaniboni L., Abdel Sayed A., Madeddu M., Iaffaldano N., Cerolini S. Effect of dimethylacetamide and N-methylacetamide on the quality and fertility of frozen/thawed chicken semen. Poult. Sci. 2019;98(11):6071-6077. DOI 10.3382/ps/pez303.

Muir W., Wong G.-K., Zhang Y., Wang J., Groenen M., Crooijmans R., Megens H.-J., Zhang H., Okimoto R., Vereijken A., Jungerius A., Albers G., Lawley C., Delany M., MacEachern S., Cheng H. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. Proc. Natl. Acad. Sci. USA. 2008;105(45):17312-17317. DOI 10.1073/pnas.0806569105.

Nandi S., Whyte J., Taylor L., Sherman A., Nair V., Kaiser P., McGrew M.J. Cryopreservation of specialized chicken lines using cultured primordial germ cells. Poult. Sci. 2016;95:1905-1911.

Partyka A., Bonarska-Kujawa D., Sporniak M., Strojecki M., Niżański W. Modification of membrane cholesterol and its impact on frozen-thawed chicken sperm characteristics. Zygote. 2016;24(5):714-723.

Pini T., Leahy T., Graaf S. Sublethal sperm freezing damage: manifestations and solutions. Theriogenology. 2018;118:172-181. DOI 10.1016/j.theriogenology.2018.06.006.

Pleshanov N., Cherepanov S., Stanishevskaya O. Chicken sperm cryopreservation as a tool of maintenance genetic diversity in small scale populations (Proc. of the XVth Eur. Poultry Conf. Dubrovnik). World's Poult. Sci. J. 2018:445.

Pleshanov N., Stanishevskaya O. Evaluation of the cocks spermatozoa membranes' damaging during cryopreservation with use of Sperm VitalStain colorant. Reprod. Domest. Anim. 2018;53(S2):183.

Pleshanov N., Stanishevskaya O., Silyukova Y. Fertilizing ability of the cryopreserved cock sperm depending on cholesterol levels and cholesterol impact on the degree of damaging of the membrane structures in spermatozoa. Genetika i Razvedenie Zhivotnyh = Genetics and Breeding of Animals. 2017;3:34-40. (in Russian)

Pleshanov N., Stanishevskaya O., Silyukova Y. Inheritance of cock's sperm cryostability. Reprod. Domest. Anim. 2019;54(3):133.

Qadeer S., Khan M., Shahzad Q., Azam A., Ansari M., Rakha B., Ejaz R., Husna A., Duman J., Akhter S. Efficiency of beetle (*Dendroides canadensis*) recombinant antifreeze protein for buffalo semen freezability and fertility. Theriogenology. 2016;86(7):1662-1669. DOI 10.1016/j.theriogenology.2016.05.028.

Rakhaa B., Ansarib M., Akhterc S., Hussaina I., Blesboisd E. Cryopreservation of Indian red jungle fowl (*Gallus gallusmurghi*) semen. Anim. Reprod. Sci. 2016;174:45-55. DOI 10.1016/j.anireprosci.2016.09.004.

Robertson L., Brown H., Staines H., Wishart G. Characterization and application of an avian *in vitro* spermatozoa-egg interaction assay using the inner perivitelline layer from laid chicken eggs. J. Reprod. Fertility. 1997;110(2):205-211.

Saint Jalme M., Lecoq R., Seigneurin F., Blesbois E., Plouzeau E. Cryopreservation of semen from endangered pheasants: the first step towards a cryobank for endangered avian species. Theriogenology. 2003;59(3-4):875-888. DOI 10.1016/s0093-691x(02)01153-6.

Sakharova S.A., Popov I.I. Improvement of chicken gene pool conservation methods. Ptahivnictvo = Poultry Farming. 2001;51:135-141. (in Ukrainian)

Santiago-Moreno J., Bernal B., Pérez-Cerezales S., Castaño C., Toledano-Díaz A., Esteso M.C., Gutiérrez-Adán A., López-Sebastián A., Gil M.G., Woelders H., Blesbois E. Seminal plasma amino acid profile in different breeds of chicken: role of seminal plasma on sperm cryoresistance. PLoS One. 2019;14(1):e0209910. DOI 10.1371/journal.pone.0209910.

Santiago-Moreno J., Castaño C., Toledano-Díaz A., Coloma M.A., López-Sebastián A., Prieto M.T., Campo J.L. Semen cryopreservation for the creation of a Spanish poultry breeds cryobank: optimization of freezing rate and equilibration time. Poult. Sci. 2011;90(9):2047-2053.

Seigneurin F., Blesbois E. Update on semen cryopreservation methods in poultry species. In: Proc. of the XIIIth Eur. Poultry Conf. 2010;172. Available at: https://www.visitvalencia.com/sites/default/files/pdfs/valencia_convention_bureau/candidaturas/epc2022valenciabid.pdf

Shahba M., El-Sheshtawy R., El-Azab A., Abdel-Ghaffar A., Ziada M., Zaky A. The effect of freeze-drying media and storage temperature on ultrastructure and DNA of freeze-dried buffalo bull spermatozoa. Asian Pac. J. Reprod. 2016;5(6):524-535. DOI 10.1016/j.apjr.2016.11.002.

Shahverdi A., Sharafi V., Gourabi H., Yekta A.A., Esmaeili V., Sharbatoghli M., Janzamin E., Hajnasrollahi M., Mostafayi F. Fertility and flow cytometric evaluations of frozen-thawed rooster semen in cryopreservation medium containing low-density lipoprotein. Theriogenology. 2015;83:78-85. DOI 10.1016/j.theriogenology.2014.07.044.

Silversides F.G., Robertson M., Liu J. Cryoconservation of avian gonads in Canada. Poult. Sci. 2013;92(10):2613-2617.

Silyukova Y., Pleshanov N., Stanishevskaya O. The influence membranes damage andactivity of roosters' sperm on the fertilization of eggs when using cured cryopreserved sperm. Reprod. Domest. Anim. 2019;54(3):101.

Song X.Y., Silversides F. Offspring produced from orthotopic transplantation of chicken ovaries. Poult. Sci. 2007;86:107-111. DOI 10.1093/ps/86.1.107.

Stanishevskaya O., Pleshanov N. Cryotolerance of cocks' sperm depending on their breed and individual properties. Anim. Reprod. Sci. 2018a;194:e1-e27. DOI 10.1016/j.anireprosci.2018.04.031.

Stanishevskaya O., Pleshanov N. Evaluation of the cocks spermatozoa membranes' damaging during cryopreservation with use of Sperm VitalStain colorant. Reprod. Domest. Anim. 2018b;53(2):183.

Stanishevskaya O., Pleshanov N. Livability of chicken embryos, obtained after insemination by frozen/thawed sperm, depending on storage duration of incubation eggs. Reprod. Domest. Anim. 2018c;53(2):199.

Svoradová A., Kuželová L., Vašíček J., Baláži A., Hanusová E., Chrenek P. *In vitro* effect of various cryoprotectants on the semen quality of endangered Oravka chicken. Zygote. 2017;26(01):33-39. DOI 10.1017/s0967199417000685.

Svoradová A., Kuželová L., Vašíček J., Olexíková L., Baláži A., Kulíková B., Hrnčár C., Ostro A., Bednarczyk M., Chrenek P. The assessment of cryopreservation on the quality of endangered Oravka rooster spermatozoa using casa and cytometry. CryoLetters. 2018;39(6):359-365.

Thélie A., Bailliard A., Seigneurin F., Zerjal T., Tixier-Boichard M., Blesbois E. Chicken semen cryopreservation and use for the restoration of rare genetic resources. Poult. Sci. 2019;98(1):447-455. DOI 10.3382/ps/pey360.

Thieu Ngoc Lan Phuong, Váradi E., Vegi B., Liptói K., Barna J. Comparison between low/programmable freezing and fast freezing protocols of Hungarian guinea fowl semen. Athens J. Nat. Formal Sci. 2014;1(3):175-183. DOI 10.13140/2.1.2727.2322.

Tonus C., Cloquette K., Ectors F., Piret J., Gillet L., Antoine N., Desmecht D., Vanderplasschen A., Waroux O., Grobet L. Long term-cultured and cryopreserved primordial germ cells from various chicken breeds retain high proliferative potential and gonadal colonisation competency. Reprod. Fertil. Dev. 2016;28:628-639.

Tonus C., Connan D., Waroux O., Vandenhove B., Wayet J., Gillet L., Desmecht D., Antoine N., Ectors F.J., Grobet L. Cryopreservation of chicken primordial germ cells by vitrification and slow freezing: a comparative study. Theriogenology. 2017;88:197-206. DOI 10.1016/j.theriogenology.2016.09.022.

Tselyutin K.V., Tour B.K. Artificial Insemination and Cryopreservation of Sperm (Roosters, Turkeys, Gander, Drakes). St. Peterburg, 2013. (in Russian)

Valcarce D.G., Carton-Garcia F., Riesco M.F., Herraez M.P., Robles V. Analysis of DNA damage after human sperm cryopreservation in genes crucial for fertilization and early embryo development. Andrology. 2013;1:723-730. DOI 10.1111/j.2047-2927.2013.00116.x.

van der Lavoir M., Diamond J., Leighton P., Mather-Love C., Heyer B., Bradshaw R., Kerchner A., Hooi L., Gessaro T., Swanberg S., Delany M., Etches R. Germline transmission of genetically modified primordial germ cells. Nature. 2006;441(7094):766-769.

Varadi E., Vegi B., Liptoi K., Barna J. Methods for cryopreservation of Guinea fowl sperm. PLoS One. 2013;8(4):e62759. DOI 10.1371/journal.pone.0062759.

Vilagran I., Castillo-Martín M., Prieto-Martínez N., Bonet S., Yeste M. Triosephosphate isomerase (TPI) and epididymal secretory glutathione peroxidase (GPX5) are markers for boar sperm quality. Anim. Reprod. Sci. 2016;165:22-30. DOI 10.1016/j.anireprosci.2015.12.001.

Vilagran I., Yeste M., Sancho S., Castillo J., Oliva R., Bonet S. Comparative analysis of boar seminal plasma proteome from different freezability ejaculates and identification of Fibronectin 1 as sperm freezability marker. Andrology. 2015;3(2):345-356. DOI 10.1111/andr.12009.

Wakayama S., Kamada Y., Yamanaka K., Kohda T., Suzuki H., Shimazu T., Tada M.N., Osada I., Nagamatsu A., Kamimura S., Nagatomo H., Mizutani E., Ishino F., Yano S., Wakayama T. Healthy offspring from freeze-dried mouse spermatozoa held on the International Space Station for 9 months. Proc. Natl. Acad. Sci. USA. 2017;14(23):5988-5993. DOI 10.1073/pnas.1701425114.

Wang S., Wang W., Xu Y., Tang M., Fang J., Sun H., Sun Y., Gu M., Liu Z., Zhang Z., Lin F., Wu T., Song N., Wang Z., Zhang W., Yin C. Proteomic characteristics of human sperm cryopreservation. Proteomics. 2014;14:298-310. DOI 10.1002/pmic.201300225.

Watson P.F. The causes of reduced fertility with cryopreserved semen. Anim. Reprod. Sci. 2000;60-61:481-492.

Wildt D. Genome resource banking for wildlife research, management, and conservation. ILAR J. 2000;41(4):228-234. DOI 10.1093/ilar.41.4.228.

Woelders H. Animal genetic resources conservation in the Netherlands and Europe: poultry perspective. Poult. Sci. 2006;85:216-222.

Wu Y., Guo L., Liu Z., Wei H., Zhou Y., Tan J., Peng J. Microelements in seminal and serum plasma are associated with fresh semen quality in Yorkshire boars. Theriogenology. 2019;132:88-94. DOI 10.1016/j.theriogenology.2019.04.002.

ORCID ID

Y.L. Silyukova orcid.org/0000-0003-1905-6373
O.I. Stanishevskaya orcid.org/0000-0001-9504-3916
N.V. Dementieva orcid.org/0000-0003-0210-9344

Acknowledgements. The study was supported by the Russian Science Foundation, project No. 19-16-00009.

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

Received October 10, 2019. Revised November 13, 2019. Accepted December 4, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):185-190
DOI 10.18699/VJ20.612

Полногеномные ассоциативные исследования распространения пороков развития и других селекционно значимых качественных признаков у потомства хряков крупной белой породы российской селекции

А.А. Траспов ✉, О.В. Костюнина, А.А. Белоус, Т.В. Карпушкина, Н.А. Свеженцева, Н.А. Зиновьева

Федеральный научный центр животноводства – ВИЖ им. академика Л.К. Эрнста, пос. Дубровицы, Подольский городской округ, Московская область, Россия
✉ e-mail: traspovalex@gmail.com

Аннотация. Выявление областей генома, прямо или опосредованно связанных с признаками пороков развития у домашних свиней, может способствовать идентификации генетических мишеней, используемых в качестве биомаркеров индивидуальных особенностей формирования экстерьера, их метаболического статуса, а также подверженности генетическим заболеваниям. Такие исследования напрямую связаны с повышением экономической эффективности, поскольку позволяют выявлять и исключать из селекционного процесса животных-носителей нежелательных генов, фенотип которых может не проявляться. В данной работе проведен поиск подобных целевых генов и геномных регионов с помощью полногеномных ассоциативных исследований (GWAS) с использованием ДНК-чипов PorcineSNP60K BeadChips (Illumina, San Diego, USA). Проанализировано 48 хряков свиней крупной белой породы селекционно-гибридного центра «Знаменский» Орловской области по 21 недостатку экстерьера и дефектам развития у 39 153 их потомков. Расчеты производили по линейной модели смешанного типа в программном пакете GEMMA. Из изначального сета в 61 000 SNP были отобраны 36 704 полиморфных SNP, в которых найдены 24 полиморфизма, входящих в 11 генов ($P < 0.1$), статистически значимо коррелирующих с признаками аномалий развития в геноме свиней, такими как атрезия анального отверстия (*ARMC7*, *FANCC*, *RND3*, *ENSSSCG00000017216*), проблемы с конечностями (*PAWR*, *NTM*, *OPCML*, *ENSSSCG00000040250*, *ENSSSCG00000017018*) и тремор поросят (*RIC3*, *ENSSSCG00000032665*). Также была выявлена коэкспрессия генов *NTM*, *OPCML* и *RND3*, участвующих в регуляции клеточной адгезии. Проведенная работа подтвердила актуальность применения подобного подхода в полногеномно-ассоциативных исследованиях для детектирования единичных SNP, связанных с отдельными признаками, даже для небольших выборок.
Ключевые слова: маркер-зависимая селекция; количественные признаки; SNP-чипы; полногеномные ассоциативные исследования; пороки развития свиней.

Для цитирования: Траспов А.А., Костюнина О.В., Белоус А.А., Карпушкина Т.В., Свеженцева Н.А., Зиновьева Н.А. Полногеномные ассоциативные исследования распространения пороков развития и других селекционно значимых качественных признаков у потомства хряков крупной белой породы российской селекции. Вавиловский журнал генетики и селекции. 2020;24(2):185-190. DOI 10.18699/VJ20.612

Whole-genome association studies of distribution of developmental abnormalities and other breeding-valuable qualitative traits in offspring of the Russian large-white boars

A.A. Traspov ✉, O.V. Kostyunina, A.A. Belous, T.V. Karpushkina, N.A. Svejenceva, N.A. Zinovieva

L.K. Ernst Federal Science Center for Animal Husbandry, Dubrovitzy, Podolsk municipal district, Moscow region, Russia
✉ e-mail: traspovalex@gmail.com

Abstract. Identifying genome regions that are directly or indirectly associated with developmental defects and malformations in domesticated pigs can help identify genomic traits used as biomarkers of the structural and functional composition of the body, their metabolic status and genetic diseases as well. Such studies are directly related to the improvement of the economic efficiency, as they allow identification and exclusion of defect animals, who may carry target genes not appearing phenotypically, from the breeding process. In the current work, we have searched for these kind of target genes and genome regions with conducting the genome-wide association studies using PorcineSNP60K BeadChips (Illumina, San Diego, USA). A total of 48 boars of a large white breed of the nucleus farm "Znamenskoe" were analyzed for 21 traits of indicated shortcomings of the exterior and defects of development in 39,153 their offspring. Calculations were made using a mixed type linear model in package GEMMA. In this study, we selected only 36,704 polymorphic SNPs from an initial 61,000-strong SNP set. After GWAS, we obtained 24 alleles in 11 corresponding genes ($P < 0.1$) in the genome of pigs, which are significantly correlated with traits of developmental abnormalities such as

anal atresia (*ARMC7*, *FANCC*, *RND3*, *ENSSSCG00000017216*), limb problems (*PAWR*, *NTM*, *OPCML*, *ENSSSCG00000040250*, *ENSSSCG00000017018*) and tremor of piglets (*RIC3*, *ENSSSCG00000032665*). Also, co-expression of the *NTM*, *OPCML* and *RND3* genes was revealed. This study confirms the relevance of using the single SNP detection according to the single trait approach in associative studies, even for small sample numbers.
Key words: marker assign selection; quantitative trait loci; SNP-chips; genome-wide association studies; malformations.

For citation: Traspov A.A., Kostyunina O.V., Belous A.A., Karpushkina T.V., Svejenceva N.A., Zinovieva N.A. Whole-genome association studies of distribution of developmental abnormalities and other breeding-valuable qualitative traits in offspring of the Russian large-white boars. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):185-190. DOI 10.18699/VJ20.612 (in Russian)

Введение

В свиноводстве болезни приводят к большим экономическим потерям не только из-за затрат на медикаментозное лечение, но и вследствие снижения продуктивных показателей больных животных. На данный момент у домашних свиней насчитывается более 130 болезней как наследственной, так и инфекционной этимологии (https:// thepigsite.com/disease-guide). Расширение знаний о причинах болезней позволит нивелировать их влияние на организм благодаря более совершенным программам разведения (Boddicker et al., 2012). Отклонения от нормального развития могут затрагивать различные органы и системы, ухудшая физическое состояние животного или даже приводя к смерти. Анатомические аномалии или дефекты, вызванные генетическими или экологическими факторами, встречаются по крайней мере у 1 % новорожденных поросят. В отдельных стадах такие аномалии могут встречаться с достаточно высокой частотой, приводя к значительным экономическим потерям (See et al., 2006).

Одна из стратегий снижения экономических потерь, обусловленных наследственными болезнями, – это выявление и исключение из разведения животных, генетически чувствительных к таким заболеваниям. Например, уже с начала 1990-х гг. селекционеры используют технологию маркерной селекции для выявления нежелательных аллелей гена *HAL*, вызывающего синдром стресса свиней, и гена *RN*, обусловливающего дефект «кислого мяса» (Salas, Mingala, 2017). Дополнение индексов племенной ценности (EBV) информацией, полученной на основании анализа непосредственно генотипа животного, делает возможным создание нового типа индекса – GEBV (genetic evaluation breeding value), характеризующегося более высокой точностью. Таким образом, дополнение традиционных методов оценки молекулярно-генетическими данными стало шагом вперед в направлении повышения интенсивности искусственного отбора (Племяшов, 2014). Выявление молекулярных маркеров, ответственных за желательные фенотипические эффекты, облегчает селекционный процесс и ускоряет получение прибыли в производстве (Ernst, Steibel, 2013). Исследования ассоциаций ДНК-маркеров в свиноводстве привлекают внимание ученых как в нашей стране (Долматова, Сковородин, 2010), так и за рубежом (Bruun et al., 2006; Ciobanu et al., 2011). Понимание генетических механизмов, ответственных за конкретные генетические аномалии, поможет производителям племенной продукции в разработке методов отбора, поскольку разные типы маркеров в геноме отвечают за разные фенотипические признаки. Так, с помощью MAS (маркер-зависимая селекция) можно проводить не только выбраковку, но и целевой отбор животных, устойчивых к заболеваниям. К примеру, отбор индивидуумов с отсутствием рецепторов адгезии *E. coli* на поверхности кишечника (*K88*) позволяет получить от них потомство, устойчивое к колибактериозу (Nyachoti et al., 2012). Раскрытие механизмов появления генетических аномалий поможет производителям в разработке методик отбора животных с «желательными» генотипами.

Современные методы полногеномных исследований (детекция SNP, полногеномное секвенирование) находят применение в выявлении генетических факторов и, как следствие, в понимании биологических процессов, лежащих в основе развития экстерьера у свиней. С учетом возможного сцепленного наследования и коэкспрессии соседних генов, детекции отдельных SNP может быть недостаточно для детального изучения комплексных признаков заболеваний или резистентности к ним. Включение в селекционные программы ДНК-маркеров QTL в качестве дополнительного критерия позволяет повысить точность оценки племенной ценности животных в аспекте их продуктивности, с учетом потенциального носительства генетических дефектов или наличия резистентности к ряду заболеваний (Sermyagin et al., 2016, 2018).

Материалы и методы

Исследования проводили на хряках крупной белой породы и их потомках, разводимых в ООО «Знаменский селекционно-гибридный центр» Орловской области. Было произведено полногеномное генотипирование хряков ($n = 48$) с использованием ДНК-чипа средней плотности Porcine SNP60BeadChip (Illumina Inc., США). Контроль качества геномных данных выполняли в программном пакете Plink 1.9. Использовались параметры качества генотипирования 90 % для одного SNP (geno 0.1), для одного образца (mind 0.1), а также для частот минорных аллелей не более 0.5 % (maf 0.05) (Purcell et al., 2007). Всего фильтрацию прошли 36 704 полиморфных SNP.

База фенотипов потомков хряков была получена из ООО «Знаменский селекционно-гибридный центр». База данных содержала 31 нежелательный показатель для каждого животного. У потомков хряков ($n = 39\,153$) рассчитывали частоту встречаемости изучаемых признаков путем деления числа носителей фенотипического показателя на общее число потомков. Для проверки гипотезы о нормальном распределении использовали критерий согласия Пирсона (χ^2 для уровня значимости 0.05) с последующей нормализацией данных в пакете bestNormalize для языка R

(Peterson, 2017). В итоге мы получили 21 нормализированный показатель, характеризующий пороки экстерьера и другие нежелательные качественные показатели: крипторхизм (CR), недоношенность (AF), атрезия ануса (AA), черные и серые пятна на шкуре (BD, GD), недовес при рождении (LW), несоответствие породе (WB), общий вес при рождении (CW), проблемы с пищеварением (DP), гермафродитизм (HM), наличие пупочной и паховой грыж (CH, UH), низкий материнский индекс (LSI), пониженное либидо у хряков (LL), несовершенный эпителиогенез (SL), низкая интенсивность роста (SG), некачественное семя у хряков-производителей (LSQ), синдром спастического тремора поросят (TP), уродства (UP), искривление конечностей (CL) и аномалии копыт (HA).

Полногеномный анализ ассоциаций (GWAS) выполняли в программном пакете GEMMA, используя линейную модель смешанного типа для частот встречаемости:

$$y = \mathrm{W}\alpha + x\beta + u + \epsilon;$$
$$u \sim \mathrm{MVN}_n(0, \lambda\tau^{-1}\,\mathrm{K}),\ \epsilon \sim \mathrm{MVN}_n(0, \tau^{-1}\,\mathrm{I}_n),$$

где y – ковариантный признак (наличие/отсутствие заболевания или другого изучаемого качественного признака в виде бинарных значений 0 или 1); $\mathrm{W} = (\mathrm{w}_1, \ldots, \mathrm{w}_c)$ – матрица коварианс (фиксированные эффекты), α – перехватывающий коэффициент, x – маркерные генотипы; β – эффект маркера, u – случайные эффекты; ϵ – ошибки; τ^{-1} – дисперсия остаточных ошибок; λ – отношение между двумя компонентами дисперсии; K – матрица родства, соотнесенная с идентификационной матрицей I_n; MVN_n – многомерное нормальное распределение.

Матрица родства рассчитывалась по формуле (в данном случае X – матрица $n \times$ P генотипов)

$$G_c = \frac{1}{p}\sum_{i=1}^{p}(x_i - 1_n\bar{x}_i)(x_i - 1_n\bar{x}_i)^T,$$

$$G_c = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{v_{x_i}}(x_i - 1_n\bar{x}_i)(x_i - 1_n\bar{x}_i)^T,$$

где x_i – каждая i-я колонка с генотипами каждого i-го SNP; $\bar{x}_i$ – среднее для образца; v_{x_i} – варианса для каждого образца i-го SNP; $1n$ – вектор для каждого ($n \times 1$) первого образца.

В частности, SNP с меньшей малой частотой аллелей имеют тенденцию оказывать больший эффект (который обратно пропорционален его вариации в генотипе), и в подобном случае выбирается первая модель матрицы (Zhou, Stephens, 2012). Проверка альтернативной гипотезы H_1: $\beta \neq 0$ и H_0: $\beta = 0$ для каждого SNP, в свою очередь, проводилась по трем наиболее распространенным статистическим тестам (Wald, likelihood ratio test или score). В данной работе пакетом GEMMA была автоматически получена оценка максимального правдоподобия параметров λ и β (Maximum Likelihood Estimate) для дальнейшего вычисления соответствующего значения P (Zhou, Stephens, 2012). Фильтрация дисперсионных компонент λ была проведена с пороговым значением P<1e–10. Для подтверждения достоверного влияния SNP и определения значимых регионов в геноме животных были применены тесты Бонферрони (BFR) с пороговым значением $\mathrm{P} \leq 0.1$ ($\mathrm{P} < 2.86 \times 10^{-6}$) и ожидаемой долей ложных отклонений B. Efron по количеству SNP отдельно для каждого признака (Benjamini, Hochberg, 1995). При вычислении скорректированных индексов Q использовался список P-значений, полученных в результате одновременного тестирования многих гипотез (Wald, likelihood ratio или Score) (Benjamini, Hochberg, 1995). Q-значения измеряли долей ложно-позитивных индексов P в случае прохождения порогового интервала (Storey et al., 2017). В данном исследовании основным критерием был установлен пороговый интервал $\mathrm{P} < 0.1$.

Для поиска генов, ассоциированных с изучаемыми признаками, использовали данные VEP (variant effect predictor) (McLaren et al., 2016). Для визуализации значений P и геномного контроля λ были построены Manhatten и QQ графики в пакете qqman с помощью языка программирования R (Storey et al., 2017; Turner, 2017). Идентификацию генов и их функциональную аннотацию осуществляли по базе взаимосвязей STRING (https://string-db.org/cgi/input.pl). Матрицы гаплотипов были построены посредством программного пакета Haploview (Barrett et al., 2005).

Результаты

По результатам исследования были установлены значимые (согласно критериям BFR, с пороговым значением $\mathrm{P} < 0.1$) ассоциативные связи для трех из 21 проанализированного качественного показателя хряков-производителей: атрезия ануса, AA; синдром спастического тремора поросят, TP; аномалии копыт, HA. Пять SNP с высокими значениями достоверности обнаружены для признака AA (P = 1.16e–06…3.68e–09), пять – для TP (P = 1.721e–06…1.24e–08) и четырнадцать – для HA (P = 1.766e–06…1.737e–09) (см. таблицу).

Для AA и TP ($\lambda \sim 1$) уровень инфляции статистики был на номинальном уровне, а для всех признаков коэффициент геномного контроля был близок к единице, как показано на графиках квантиль–квантиль (QQ-график, рис. 1). Однако у HA выявлено наличие близкорасположенных SNP, входящих в одинаковые группы генов со значительными превышениями P порога (ASGA0104521, P = 1.737e–09). Значимые нуклеотидные полиморфизмы были локализованы внутри отдельных генов. В ходе расчета LD между SNP с самыми высокими значениями P были отобраны полиморфизмы со значениями r^2, наиболее близкими к 1 по отношению к ALGA0053356 (P = 3.115e–07, Pos 9:64845247). В результате в один блок с ним вошли ALGA0053410, MARC0024097, MARC0051180, DRGA0009397, что подтверждает ожидаемое функциональное родство отобранных SNP и генов, их включающих (рис. 2).

Анализ SNP, статистически значимо связанных с пороками развития поросят крупной белой породы, выявил несколько генов, имеющих отношение к различным биологическим процессам. Так, гены *ARMC7*, *FANCC* участвуют в репарации ДНК и клеточном цикле. *FANCC* принимает участие в передаче анемии Фанкони, *RND3* выступает как регулятор цитоскелетных структур клетки, препятствующих адгезии. *UBAP2* функционирует в процессе убиквитинирования и может проявлять повышенную экспрессию в надпочечниках и лимфатических узлах. Ген *PAWR* является опухолевым супрессором, который

Significant associations of single nucleotide polymorphisms (SNP) associated with the assessed traits in piglets and their positions in the *Sus scrofa* genome

Character	SNP	RS	Ch	Pos	P	A	a	β	AF	Gene
AA	ALGA0051997	rs81407818	9	27603177	2.666e–06	A	C	2.001394e–03	0.056	–
	ASGA0106167	rs81306460	10	31302650	1.158e–06	A	G	2.081251e–03	0.056	*FANCC*
	MARC0082230	rs81265837	12	6136945	1.419e–07	G	A	2.193750e–03	0.056	*ARMC7*
	ASGA0052617	rs81435284	12	6229056	1.419e–07	G	A	2.193750e–03	0.056	*ENSSSCG00000017216*
	ASGA0068580	rs80936660	15	1015164	3.681e–09	G	A	2.383802e–03	0.056	*RND3*
TP	H3GA0025901	rs81415828	9	1626616	1.24e–08	A	G	6.081343e–03	0.056	*ENSSSCG00000032665*
	ASGA0099429	–	9	4656663	1.24e–08	C	A	6.081343e–03	0.056	–
	ASGA0040658	rs81413027	9	1417890	3.107e–07	G	A	4.690571e–03	0.083	–
	MARC0013008	rs81275805	9	1438136	1.721e–06	G	A	3.828089e–03	0.097	*RIC3*
	ASGA0054790	rs81435622	12	45034446	9.638e–07	G	A	3.975254e–03	0.083	–
HA	ASGA0104521	rs81304512	2	141925578	1.737e–09	C	A	2.346000e–03	0.056	*ENSSSCG00000040250*
	ALGA0014069	rs81360100	2	85105474	1.755e–06	G	A	1.872074e–03	0.083	*PAWR*
	ASGA0027165	rs81386880	5	106682107	1.266e–07	A	G	2.036564e–03	0.069	–
	ASGA0090791	rs81309195	6	34358134	1.953e–06	A	G	1.764154e–03	0.1	–
	ALGA0053356	rs81412069	9	64845247	3.115e–07	A	C	2.010894e–03	0.083	*NTM*
	ALGA0053410	rs81412202	9	65286501	3.115e–07	A	C	2.010894e–03	0.083	*NTM*
	MARC0024097	rs81292427	9	65379314	3.115e–07	A	G	2.010894e–03	0.083	*OPCML*
	MARC0051180	rs81242341	9	65591835	3.115e–07	A	C	2.010894e–03	0.083	*OPCML*
	DRGA0009397	rs81294295	9	65650293	3.115e–07	A	G	2.010894e–03	0.083	–
	M1GA0012732	rs81415754	9	12012638	1.766e–06	A	C	1.807143e–03	0.083	–
	ALGA0053344	rs81412038	9	64410416	2.514e–06	G	A	1.799440e–03	0.1	*ENSSSCG00000017018*
	MARC0041414	rs81235185	16	60218765	1.766e–06	G	A	1.807143e–03	0.083	*ENSSSCG00000017018*
	MARC0112574	rs81283873	16	60220801	1.766e–06	G	A	1.807143e–03	0.083	–
	H3GA0055670	–	17	52247420	1.766e–06	G	A	1.807143e–03	0.083	–

Designations: SNP, single nucleotide polymorphism; RS, SNP identifier in the NIH dbSNP database; Ch, chromosome; Pos, position; P, validity; A, effector allele; a, reference allele; β, allele effect; AF, effector allele frequency; Gene, name of the gene housing the studied SNP.

селективно индуцирует апоптоз в раковых клетках через внутриклеточные и внеклеточные механизмы. *RIC3* влияет на свертывание и сборку рецепторных субъединиц в эндоплазматическом ретикулуме и адгезию на поверхности клетки.

Гены *NTM* и *OPCML* экспрессируются совместно и находятся на соседних участках 9-й хромосомы: 58700168–58967505 и 59037716–59271936 п. н. соответственно (www.ncbi.nlm.nih.gov/gene/100519556, www.ncbi.nlm.nih.gov/gene/100738337) (рис. 3). *NTM* способствует росту и адгезии на поверхности нейронов и тесно связан с родственным членом семейства, опиоидным связывающим белком-активатором клеточной адгезии OPCML.

Заключение

Полученные нами данные могут быть использованы при разработке селекционных программ, направленных на элиминацию пороков развития и других нежелательных количественных и качественных признаков свиней, в том числе являющихся сложными признаками. Увеличение степени разрешения сканирования от 100000 SNP и выше, а также увеличение размера выборки от нескольких сотен животных и более сделает возможным выявление значительно большего количества SNP-кандидатов с высоким уровнем достоверности ($P < 0.01$), а также уменьшение «генетического шума» (false positive components). В итоге такой метод детекции позволит не только выявлять животных-носителей генов-кандидатов нежелательных признаков, но и создать дешевые тест-системы для их идентификации. Хряков-производителей, имеющих подобные генетические особенности, необходимо оценивать с помощью комплексных моделей расчета племенной ценности с учетом выявленных маркеров (GEBV) и выбраковывать в случае крайне низких показателей продуктивности, а их потомство исключать из воспроизводства.

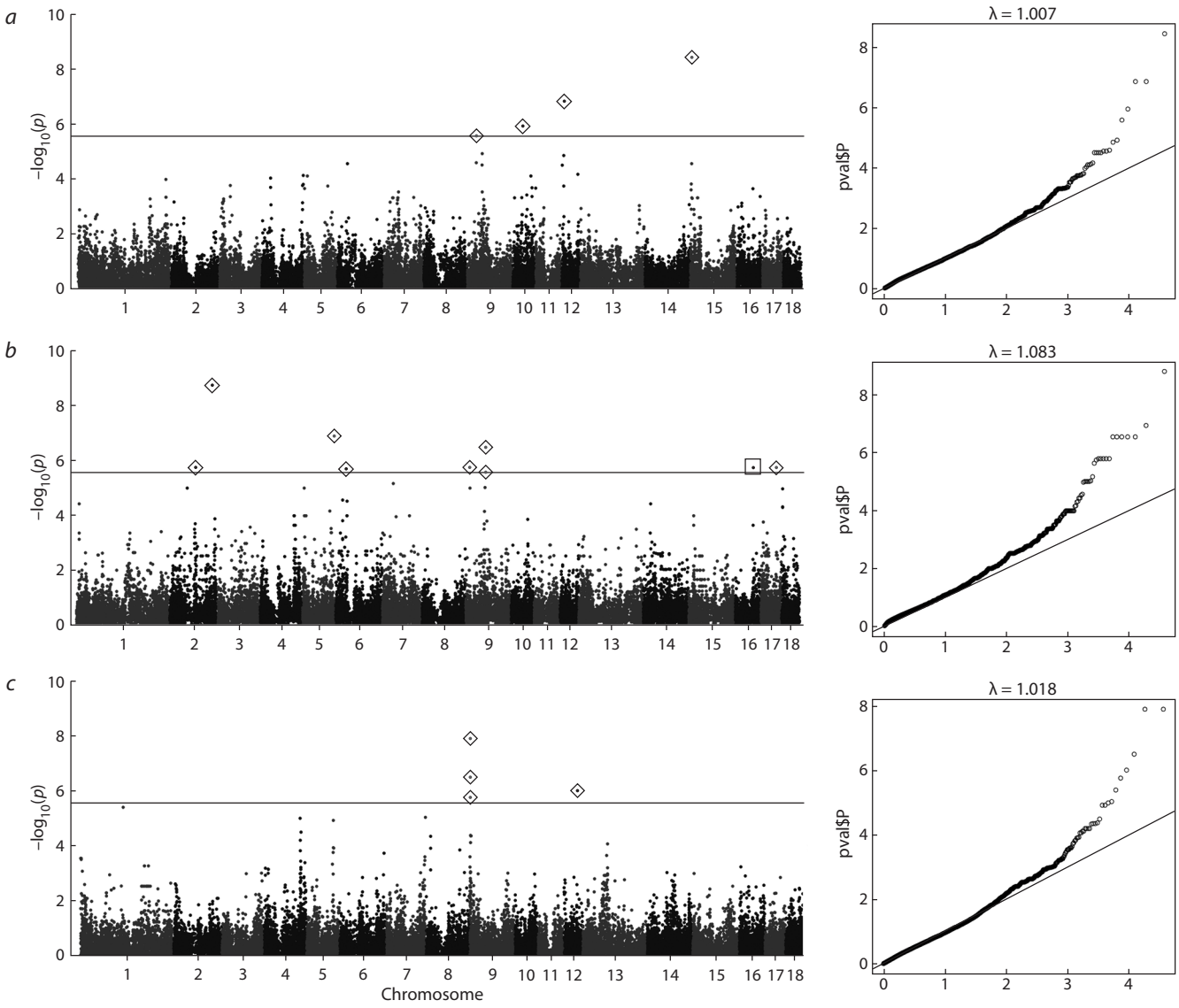


Fig. 1. Manhattan plot illustrating GWAS (genome-wide association study, top-left) and corresponding probability distributions of confidence P (Q-Q graph) in piglets of the studied populations of large white pigs (Znamenskiy Breeding Center):

a, anal atresia (AA); *b*, hoof problems (HA); and *c*, spastic tremor (TP). The solid line indicates the Bonferroni level (0.05). Reliable values satisfying the null Bonferroni hypothesis (BFR) are marked with diamonds.

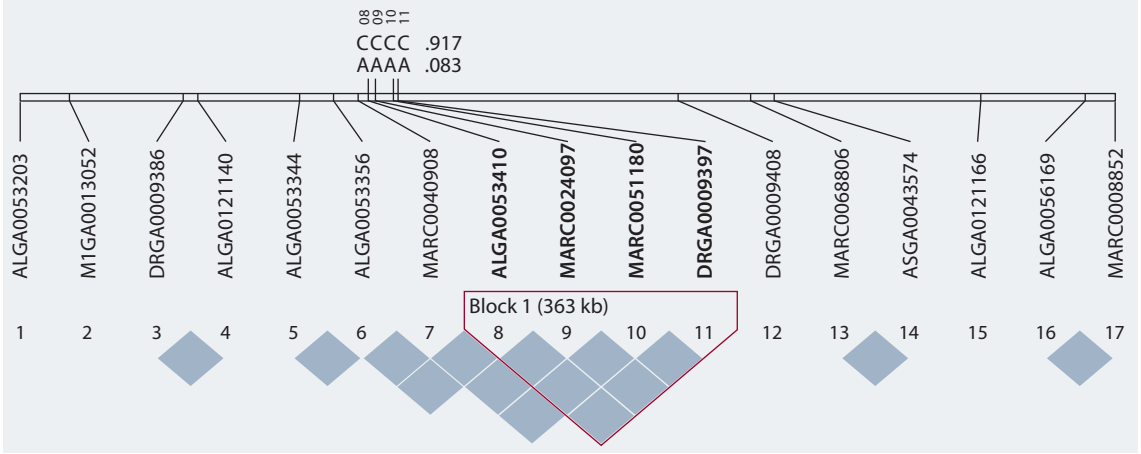


Fig. 2. Haploview LD chart illustrating 17 SNPs with the highest r^2 values of coupling between ALGA0053356 and four polymorphic variants – ALGA0053410, MARC0024097, MARC0051180, and DRGA0009397 to form a 363-kb haploblock in group HA.

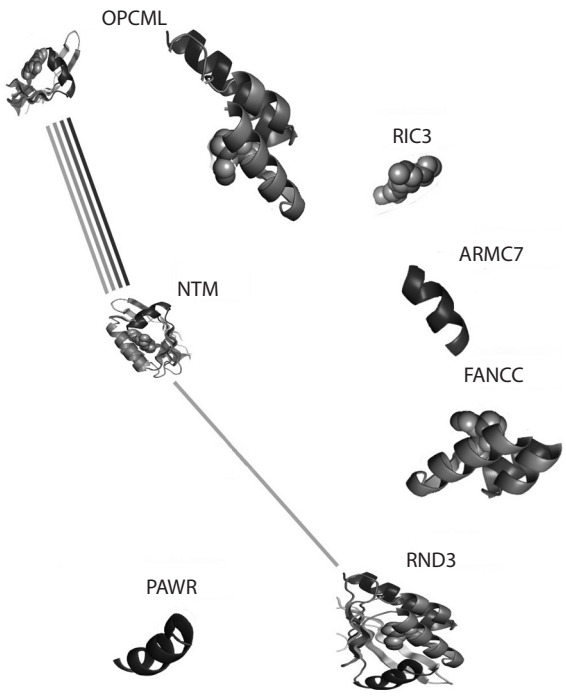


Fig. 3. String diagram (https://string-db.org) illustrating the strong protein-mediated interaction of the coexpressed *OPMCL* and *NTM* genes, as well as their indirect effect on *RND3*.

Проделанная работа иллюстрирует необходимость проведения дополнительных исследований с использованием методов GWAS в аспекте характеристики популяций сельскохозяйственных животных по ДНК-маркерам и идентификации комплексных генотипов, ассоциированных с селекционно значимыми признаками как пороков развития, так и продуктивных качеств. Данное направление крайне необходимо в современных условиях высокоэффективного воспроизводства сельскохозяйственных животных.

Список литературы / References

Долматова А.В., Сковородин Е.Н. Использование ДНК-полиморфизма в селекции свиней. В: Материалы междунар. науч.-практ. конф. «Современные проблемы интенсификации производства свинины в странах СНГ», посвященной 75-летнему юбилею заслуженного деятеля науки РФ, профессора В.Е. Уитько, 7–10 июля 2010 г. Ульяновск, 2010;138-143.
[Dolmatova A., Skovorodin E. The use of DNA polymorphism in the selection of pigs. In: Proc. of the Int. Sci. and Pract. Conf. “Modern Problems of Intensification of Pork Production in the CIS Countries”, dedicated to the 75th anniversary of the Honored scientist of the Russian Federation, professor V.E. Ulitko, July 7–10, 2010. Ulyanovsk, 2010;138-143. (in Russian)]

Племяшов К.В. Геномная селекция – будущее животноводства. Животноводство России. 2014;5:2-4.
[Plemyashov K. Genomic selection – the future of animal husbandry. Zhivotnovodstvo Rossii = Animal Husbandry of Russia. 2014;5:2-4. (in Russian)]

Barrett J., Fry B., Maller J., Daly M. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263-265. DOI 10.1093/bioinformatics/bth457.

Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. B. 1995;57(1):289-300. DOI 10.2307/2346101.

Boddicker N., Waide E.H., Rowland R.R.R., Lunney J.K., Garrick D.J., Reecy J.M., Dekkers J.C.M. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. Sci. 2012;90(6):1733-1746.

Bruun C.S., Jorgensen C.B., Nielsen V.H., Andersson L., Fredholm M. Evaluation of the porcine *melanocortin 4 receptor* (*MC4R*) gene as a positional candidate for a fatness QTL in a cross between Landrace and Hampshire. Anim. Genet. 2006;37(4):359-362. DOI 10.1111/j.1365-2052.2006.01488.x.

Ciobanu D.C., Lonergan S.M., Huff-Lonergan E.J. Genetics of meat quality and carcass traits. In: Rothschild M.F., Ruvinsky A. (Eds.). The Genetics of the Pig. 2nd ed. Wallingford: CAB International, 2011;355-389.

Ernst C.W., Steibel J.P. Molecular advances in QTL discovery and application in pig breeding. Trends Genet. 2013;29(4):215-224. DOI 10.1016/j.tig.2013.02.002.

McLaren W., Gil L., Hunt S., Riat H., Ritchie G., Thormann A., Flicek P., Cunningham F. The Ensembl Variant Effect Predictor. Genome Biol. 2016;17(1):122. DOI 10.1186/s13059-016-0974-4.

Nyachoti C., Kiarie E., Bhandari S., Zhang G., Krause D. Weaned pig responses to *Escherichia coli* K88 oral challenge when receiving a lysozyme supplement. J. Anim. Sci. 2012;90(1):252-260. DOI 10.2527/jas.2010-3596.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M., Bender D., Maller J., Sklar P., Bakker P., Daly M., Sham P. PLINK: a toolset for whole-genome association and population-based linkage analysis. Am. J. Hum. Genet. 2007;81:559-575. DOI 10.1086/519795.

Peterson R.A. Estimating normalization transformations with bestNormalize. 2017. Available at: https://github.com/petersonR/bestNormalize

Salas R., Mingala C. Genetic factors affecting pork quality: halothane and rendement napole genes. Anim. Biotechnol. 2017;28(2):148-155. DOI 10.1080/10495398.2016.1243550. Epub 2016 Nov. 17.

See T., Max F., Rothschild C., Christains J. Swine Genetic Abnormalities. Pork Information Gateway. 2006; PIG 06-06-01.

Sermyagin A., Gladyr E., Plemyashov K., Kudinov A., Dotsev A., Deniskova T., Zinovieva N. Genome-wide association studies for milk production traits in Russian population of Holstein and black-and-white cattle. In: Anisimov K.V. et al. (Eds.). Proc. of the Sci.-Pract. Conf. “Research and Development – 2016”, 14–15 Dec. 2016, Moscow, Russia. Springer Open, 2018;591-599. DOI 10.1007/978-3-319-62870-7_62.

Sermyagin A., Kovalyuk N., Ermilov A., Yanchukov I., Satsuk V., Dotsev A., Deniskova T., Brem G., Zinovieva N.A. Associations of *Bola-drb3* genotypes with breeding values for milk production traits in Russian dairy cattle population. Selskokhozyaystvennaya Biologiya = Agricultural Biology. 2016;51(6):775-781. DOI 10.15389/agrobiology.2016.6.775eng.

Storey J., Bass A., Dabney A., Robinson D. qvalue: Q-value estimation for false discovery rate control. R package version 2.10.1. 2017. http://github.com/StoreyLab/qvalue

Turner S. qqman: Q-Q and Manhattan Plots for GWAS Data. R package version 0.1.4. 2017. https://CRAN.R-project.org/package = qqman

Zhou X., Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 2012;44:821-824. DOI 10.1038/ng.2310.

Acknowledgements. This work was supported by the Russian Science Foundation, project 18-76-00034.
Conflict of interest. The authors declare no conflict of interest.
Received April 24, 2019. Revised November 05, 2019. Accepted November 06, 2019.

Вавиловский журнал генетики и селекции 2020;24(2):191-199
DOI 10.18699/VJ20.42-o

Коэкспрессия глутаматергических генов и генов аутистического спектра в гиппокампе у самцов мышей с нарушением социального поведения

И.Л. Коваленко ✉, А.Г. Галямина, Д.А. Смагин, Н.Н. Кудрявцева

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: koir@bionet.nsc.ru

Аннотация. В настоящее время существует представление о вовлеченности глутаматергической системы (ГГ) в механизмы развития аутизма. В предыдущих исследованиях нами было показано, что негативный социальный опыт, приобретенный в ежедневных межсамцовых конфронтациях, приводит к нарушениям в социальном поведении: снижению коммуникативности, нарушению социализации, появлению стереотипных форм поведения, которые могут рассматриваться как симптомы аутистического спектра. В связи с этим целью нашей работы было изучение с помощью транскриптомного анализа изменений экспрессии генов, кодирующих белки, вовлеченные в функционирование глутаматергической системы, и генов, связанных с патологией аутизма (ГА), в гиппокампе. В эксперименте использовали животных с нарушениями социального поведения, вызванными повторным опытом социальных побед или поражений в ежедневных агонистических взаимодействиях. Для формирования групп животных с контрастными типами поведения использовали модель сенсорного контакта (хронического социального стресса). Полученные образцы мозга были секвенированы в ЗАО «Геноаналитика» (http://genoanalytica.ru/, Россия, Москва). Транскриптомный анализ показал, что у агрессивных животных снижается экспрессия генов *Shank3*, *Auts2*, *Ctnnd2*, *Nrxn2*, для которых показано участие в развитии аутизма, а также глутаматергического гена G*rm4*. В то же время у животных с негативным социальным опытом экспрессия ГА *Shank2*, *Nlgn2*, *Ptcdh10*, *Reln*, *Arx* возрастает. При этом ГГ (*Grik3*, *Grm2*, *Grm4*, *Slc17a7*, *Slc1a4*, *Slc25a22*), за исключением гена *Grin2a*, повышают свою экспрессию. Корреляционный анализ выявил статистически значимую взаимосвязь измененной экспрессии ГГ и ГА. Полученные результаты, с одной стороны, могут служить подтверждением участия ГГ в патофизиологии развития симптомов аутистического спектра, с другой – свидетельствовать о коэкспрессии ГГ и ГА в гиппокампе, развивающейся под влиянием социальной среды. Так как большинство ГА, изменивших экспрессию в настоящем исследовании, являются генами, связанными с клеточным скелетом и внеклеточным матриксом, в частности участвующими в формировании синапсов, а ГГ, изменившие свою экспрессию, – генами, кодирующими субъединицы рецепторов, то можно предположить, что вовлечение ГГ в патофизиологию аутизма происходит на уровне рецепторов.

Ключевые слова: RNA-seq; аутизм; гиппокамп; гены аутизма; глутаматергические гены; социальный опыт.

Для цитирования: Коваленко И.Л., Галямина А.Г., Смагин Д.А. , Кудрявцева Н.Н. Коэкспрессия глутаматергических генов и генов аутистического спектра в гиппокампе у самцов мышей с нарушением социального поведения. Вавиловский журнал генетики и селекции. 2020;24(2):191-199. DOI 10.18699/VJ20.42-o

Co-expression of glutamatergic and autism-related genes in the hippocampus of male mice with disturbances of social behavior

I.L. Kovalenko ✉, A.G. Galyamina, D.A. Smagin, N.N. Kudryavtseva

Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ e-mail: koir@bionet.nsc.ru

Abstract. There is a hypothesis of the involvement of the glutamatergic system in the development of autism. It has been shown that the chronic experience in daily intermale confrontations leads to disturbances in social behavior: a decrease in communicativeness, disturbances of socialization, emergence of stereotypical behaviors that can be considered as symptoms of the autistic spectrum disorders. So, the aim of this study was to investigate changes in the expression of glutamatergic (GG) and autism-related (GA) genes in the hippocampus of animals with impaired social behavior caused by repeated experience of social defeat or aggression in daily agonistic confrontations. To form groups of animals with contrasting behaviors, a model of sensory contact (chronic social stress) was used. The collected brain samples were sequenced at JSC Genoanalytica (http://genoanalytica.ru/, Moscow, Russia). Transcriptomic analysis re-

vealed a down-regulation of autism-related (*Shank3*, *Auts2*, *Ctnnd2*, *Nrxn2*) and glutamatergic (*Grm4*) genes in aggressive mice. At the same time, the expression of GA-related genes (*Shank2*, *Nlgn2*, *Ptcdh10*, *Reln*, *Arx*) and GG genes (*Grik3*, *Grm2*, *Grm4*, *Slc17a7*, *Slc1a4*, *Slc25a22*) excluding *Grin2a* was increased in defeated mice. Correlative analysis revealed a statistically significant association between GG and GA expression. These results can serve as a confirmation of the participation of the glutamatergic system in the pathophysiology of the autistic spectrum disorder.
Key words: RNA-seq; autism; hippocampus; glutamatergic genes; autism-related genes; social experience.

For citation: Kovalenko I.L., Galyamina A.G., Smagin D.A., Kudryavtseva N.N. Co-expression of glutamatergic and autism-related genes in the hippocampus of male mice with disturbances of social behavior. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):191-199. DOI 10.18699/VJ20.42-o (in Russian)

Введение

Полагают, что выраженные нарушения в социальном поведении могут свидетельствовать об аутизме, который проявляется в детском возрасте и представляет собой нарушения развития нервной системы (Bauman, Kemper, 2005; Zablotsky et al., 2015). Согласно DSM-IV (American Psychiatric Association…, 1993), диагностические критерии аутизма включают триаду основных признаков: ухудшение социализации, под которой можно понимать способность индивида адекватно встраиваться в социум в новой обстановке, низкий уровень общительности и повторяющееся/стереотипное поведение. Хотя изучение близнецов свидетельствует о высокой наследуемости аутизма (Hallmayer et al., 2011), ни один ген не определен как единственная причина развития этого заболевания. Недавние геномные и генетические исследования показали, что сотни генетических вариантов, включающие общие и редкие взаимодействия генов, способствуют возникновению аутизма (Miles, 2011). Согласно различным базам данных (http://omim.org/, http://www.genecards.org/, http://autism.mindspec.org/autdb/Welcome.dohttp://www.malacards.org), в базе генов аутизма насчитывается около 1500 генов, которые в той или иной мере вовлечены в механизмы аутизма. Многие гены, связанные с развитием головного мозга, – потенциальные гены-кандидаты аутизма, к ним относятся гены и семейства генов *Shank* (*1–3*), *Nlgn*, *Reln*, *Arx*, *Pcdh*, *Mecp2*, *Auts2* (Kleijer et al., 2014; Liu et al., 2015).

Литературные данные свидетельствуют об участии различных нейромедиаторных систем (например, серотонергической, дофаминергической) в развитии аутизма (Pavăl, 2017). В последнее время большое внимание уделяется глутаматергической гипотезе аутизма (Rojas, 2014). В пользу этой теории говорит то, что глутамат является одним из наиболее распространенных нейротрансмиттеров в мозге млекопитающих, его рецепторы сосредоточены в областях мозга (мозжечок, гиппокамп, префронтальная кора), в которых обнаружены нейропатологические изменения при аутизме. Во взрослом мозге глутаматные рецепторы участвуют в обучении и памяти (Riedel et al., 2003; Simonyi, 2010). Глутамат-опосредованные межнейронные взаимодействия также играют роль в формировании эмоционального поведения (Morgane et al., 2005; Faure et al., 2010). Показано, что для больных аутизмом характерен повышенный уровень глутамата в плазме крови, который может даже служить одним из биомаркеров этого заболевания (Zheng et al., 2016). У пациентов с аутизмом обнаруживаются и молекулярно-генетические повреждения глутаматергической системы (ГГ). Так, например, есть сообщения о мутациях в генах глутаматного рецептора *GluR6* (Jamain et al., 2002), а также митохондриального переносчика глутамата (AGC1, ген *Slc25a12*) (Ramoz et al., 2004). Изменения наблюдаются на уровне м-РНК, белков-транспортеров и рецепторов ГГ в посмертных образцах мозга больных аутизмом (Purcell et al., 2001). В целом некоторые исследователи связывают аутизм с дефицитом глутаматергической системы в мозге (Carlsson, 2015).

Ранее нами было показано, что в условиях хронического социального конфликта, вызванного повторным опытом побед или поражений в ежедневных межсамцовых конфронтациях (Kudryavtseva, 1991), у самцов мышей формируются не только повышенная тревожность, агрессивность или депрессивноподобное поведение, но и нарушения коммуникативного поведения и социального взаимодействия. Такие животные демонстрировали сниженную коммуникативность, а также стереотипные формы поведения (аутогруминг, вставание на задние лапы, кружение, разрывание и разбрасывание подстилки, повороты в прыжке и др.), что позволяет рассматривать данную модель как модель, воспроизводящую некоторые черты аутистического поведения (Коваленко, Кудрявцева, 2010), формирующиеся под влиянием социальной среды. При этом наши предыдущие исследования показали, что в этот процесс были вовлечены гены аутистического спектра, экспрессия которых изменялась под влиянием агонистических взаимодействий (Kudryavtseva et al., 2017).

В связи с этим целью настоящей работы было изучить у животных с альтернативным опытом социального поведения, позитивным и негативным, взаимосвязь изменения экспрессии генов, кодирующих белки, вовлеченные в функционирование ГГ, и генов, связанных с патологией аутизма (ГА) в гиппокампе. Выбор этой структуры головного мозга основан на многочисленных литературных данных о том, что гиппокамп непосредственно участвует в патогенетических процессах аутизма (DeLong, 1992), в развитии тревожных расстройств (Irle et al., 2010) и депрессии (Savitz, Drevets, 2009). Гиппокамп отвечает за эмоциональную саморегуляцию, обучаемость и память. А, как известно, нарушение именно этих функций часто наблюдается у людей, демонстрирующих симптомы аутистического спектра.

Материалы и методы

Животные. Эксперименты проводили на самцах мышей линии C57BL/6J в возрасте 2.5 мес. и массой тела 26–28 г. Животные были привезены из питомника лабораторных животных Института биоорганической химии Российской академии наук (Пущино, Московская область). Экспе-

римент проводили в виварии конвенциональных животных Института цитологии и генетики Сибирского отделения РАН. Воду и корм (гранулы) животные получали в достаточном количестве. Световой режим был 12С:12Т. Все процедуры осуществляли в соответствии с международными правилами проведения экспериментов с животными (Directive 2010/63/EU of the European Parliament and of the Council on the Protection of Animals Used for Scientific Purposes). Применяемые методики для изучения поведения у мышей были одобрены Научной комиссией (№ 9) Института цитологии и генетики СО РАН (Протокол № 613 от 24.03.2010).

Поведенческие исследования. Для формирования альтернативного опыта социального поведения у самцов мышей использовали модель сенсорного контакта (Kudryavtseva et al., 2014). Животных попарно помещали в экспериментальные клетки, разделенные пополам прозрачной перегородкой с отверстиями, позволявшей мышам видеть, слышать, воспринимать запахи друг друга (сенсорный контакт), но предотвращавшей физическое взаимодействие. Ежедневно во второй половине дня (15:00–17:00) убирали перегородку, что приводило к агонистическим взаимодействиям. В течение первых двух-трех дней тестов выявляли победителей (агрессоров/агрессивных животных) и особей, терпящих поражения (жертв) при взаимодействии с одним и тем же партнером. В дальнейшем ежедневно после теста побежденного самца пересаживали в новую клетку к незнакомому агрессивному партнеру, сидящему за перегородкой. Взаимодействие самцов прекращали, если интенсивные атаки со стороны нападающей особи во время агрессивных столкновений длились более трех минут, устанавливая между ними перегородку. В исследование были взяты агрессивные животные с 20-дневным опытом побед (агрессоры) и с 20-дневным опытом поражений (жертвы). В качестве контроля использовали самцов, не имевших последовательного опыта агонистических взаимодействий. В каждой группе было по 14–16 животных.

Количественную оценку реакции экспериментальных животных на незнакомого партнера на нейтральной территории проводили при помощи теста «социальные взаимодействия» (Kudryavtseva et al., 2017). Как параметры социального поведения мы рассматривали *избегание* незнакомого партнера или же замирание при его подходе и *приближение* к партнеру: подходы, обнюхивания, следование за партнером; как параметры индивидуального поведения выделяли с*тойки* – вставания на задние лапы, являющиеся показателем исследовательской активности, *аутогруминг*, служащий показателем смещенной активности и неадекватного стереотипного поведения, и *двигательную активность*, оценивающую интенсивность передвижений по клетке.

Статистическую обработку полученных данных выполняли с использованием пакета программ STATISTICA (ver. 8.0; StatSoft, Inc., 2001). Проверка нормальности распределения количественных признаков была проведена с использованием критерия Шапиро-Уилка (Shapiro-Wilk's W-test). Поскольку выборки исследованных параметров поведения удовлетворяли гипотезе о нормальном распределении, были использованы методы параметрической статистики: однофакторный дисперсионный анализ (ANOVA) с фактором «группа» (контроль, агрессивные самцы, жертвы); последующее попарное сравнение показателей осуществляли с помощью LSD-теста Фишера.

При помощи корреляционного анализа по методу Пирсона мы исследовали взаимосвязь экспрессии ГГ и ГА. В экспериментальных поведенческих группах было по 10–12 животных.

RNA-Seq-анализ проводили с помощью ЗАО «Геноаналитика» (http://genoanalytica.ru, Москва). Методика анализа подробно описана ранее (Галямина и др., 2017).

Мы провели проверку результатов, сравнив их с данными B.M. Kadakkuzha (Kadakkuzha et al., 2015), представившими полный транскриптомный анализ генов в гиппокампе интактных мышей. Обнаружен высокий уровень корреляции (0.74 по Спирмену) между экспрессией генов аутизма и ГГ у контрольных особей в нашем эксперименте и у интактных животных в работе (Kaddakuzha et al., 2015), что может быть дополнительным доказательством адекватности применяемого метода. Кроме того, осуществляли кросс-верификацию результатов (Babenko et al., 2017) с данными, полученными в Стэндфордском университете (Zhang et al., 2014), и обнаружили значительную корреляцию между ними. Это доказывает, что метод транскриптомного анализа позволяет выявить происходящие в мозге молекулярные процессы.

В качестве верификации данных настоящего эксперимента использовали результаты, полученные нами ранее при помощи RT-PCR (Smagin et al., 2013), экспрессия которых, по методу RNA-Seq, была изменена. Было показано совпадение направленности и выраженности изменений экспрессии для генов, кодирующих белки серотонергической системы в ядрах шва среднего мозга, полученных при применении этих методов, что позволяет говорить о высокой надежности результатов этого исследования и о стабильности примененного метода.

Декапитацию всех трех групп экспериментальных животных проводили одновременно на следующий день после последней конфронтации. Гиппокамп извлекался одним исследователем в соответствии с анатомическим атласом мозга (Allen Mouse Brain Atlas; http://mouse.brainmap.org/static/atlas). Все образцы помещали в раствор RNAlater (LifeTechnologies, США) и хранили при температуре –70 °С до секвенирования. Проводили два типа сравнения: контроль–агрессоры и контроль–жертвы.

Для изучения изменений глутаматергической системы в гиппокампе животных с нарушенным социальным поведением были исследованы: гены переносчиков глутамата: *Slc17a6*, *Slc17a7* и *Slc17a8*; гены, кодирующие 1–4-ю субъединицы ионотропного AMPA-рецептора: *Gria1*, *Gria2*, *Gria3*, *Gria4*; гены, кодирующие 1–5-ю субъединицы ионотропного каинатного глутаматного рецептора: *Grik1*, *Grik2*, *Grik3*, *Grik4*, *Grik5*; гены, кодирующие 1, 2a, *2b*, *2c*, *2d*, *3a*, *3b* субъединицы ионотропного NMDA-рецептора: *Grin1*, *Grin2a*, *Grin2b*, *Grin2c*, *Grin2d*, *Grin3a*, *Grin3b*; гены, кодирующие метаботропные рецепторы 1–8-го подтипов: *Grm1*, *Grm2*, *Grm3*, *Grm4*, *Grm5*, *Grm6*, *Grm7*, *Grm8*; гены, кодирующие 1 и 2-ю субъединицы глутаматного ионотропного рецептора дельта *Grid1* и *Grid2*; GRID2IP – белок, взаимодействующий с *Grid2*; гены

фермента глутаматдекарбоксилазы, метаболизирующей глутамат в ГАМК: *Gad1* и *Gad2*.

По генетическим базам данных OMIM (http://omim.org/), GeneCards (http://www.genecards.org/), MalaCards (http://www.malacards.org/) из 1.5 тыс. аннотированных генов было выбрано около 80 основных генов-кандидатов аутизма, которые в дальнейшем были просмотрены в гиппокампе у контрольных особей и животных с нарушенным социальным поведением. При анализе сравнивали по три пробы от каждой группы животных.

Результаты

Исследование нарушений социального поведения у самцов мышей под влиянием хронического социального стресса. ANOVA выявил достоверное влияние фактора «группа» (контроль, агрессоры, жертвы) на избегание партнера ($F(2.29) = 52.30$, $p < 0.001$), приближение ($F(2.28) = 1097.4$, $p < 0.001$), время стоек ($F(2.29) = 661.7$, $p < 0.001$), время двигательной активности ($F(2.29) = 2549$, $p < 0.001$). Сравнение групп LSD-тестом Фишера (рис. 1) выявило увеличение времени избегания у жертв, по сравнению с контролем и агрессивными животными ($p < 0.001$ для обеих групп). По сравнению с контролем у агрессоров и жертв также было показано снижение времени стоек ($p < 0.001$ для обеих групп) и времени приближения к партнеру ($p < 0.001$ для обеих групп). Кроме того, у жертв время приближения к партнеру было снижено и по сравнению с агрессивными животными ($p < 0.006$). Выявлены увеличение времени двигательной активности у агрессоров ($p < 0.001$) и снижение времени у жертв ($p < 0.042$), по сравнению с контрольными животными. Время двигательной активности было значительно ниже у жертв, по сравнению с агрессорами ($p < 0.001$). Кроме того, у жертв были увеличены число и время аутогруминга, по сравнению с контролем ($p < 0.041$, $p < 0.034$ соответственно).

Таким образом, мы видим, что в тесте «социальные взаимодействия» жертвы активно избегали незнакомого партнера (тестера) на нейтральной территории. Они редко первыми подходили и проявляли интерес к незнакомому партнеру, в отличие от контроля. Контрольные животные целенаправленно следовали за партнером, обнюхивали его. У жертв также было снижено время стоек, рассматриваемое нами как показатель исследовательской активности. Кроме того, у них было увеличено время демонстрации аутогруминга, который может оцениваться как признак стереотипного поведения. При этом отмечено, что у агрессивных животных были снижены параметры коммуникативности, что может свидетельствовать о нарушениях социального поведения. Большую часть 10-минутного теста (около 5–6 мин) агрессоры хаотично перемещались по клетке, не обращая внимания на партнера: общее время двигательной активности было значительно больше, чем у жертв, что может отражать развитие гиперактивности и, по-видимому, дефицит внимания. Ранее нами было показано, что в агонистических взаимодействиях агрессоры часто демонстрируют быстрые повороты в прыжке или же повторяющиеся повороты вокруг оси тела в тесте «перегородка» (Kudryavtseva, 2006), т. е. стереотипии у агрессивных мышей этой линии могут проявляться в другом тесте и другой форме. Таким образом, и у агрессоров, и у жертв в результате 20-дневных агонистических взаимодействий развиваются симптомы аутистического спектра.

Исследование экспрессии ГГ и ГА у самцов мышей с контрастными типами социального поведения. В результате анализа данных RNA-Seq в гиппокампе обнаружено изменение экспрессии девяти генов-кандидатов аутизма (рис. 2, табл. 1). Так, у жертв выявлено увеличение экспрессии генов *Shank2* ($p < 0.040$), *Nlgn2* ($p < 0.047$), *Pcdh10* ($p < 0.011$), *Reln* ($p < 0.026$) и *Arx* ($p < 0.0002$), по сравнению с уровнем экспрессии у контрольных животных. У агрессивных мышей в гиппокампе под влиянием повторного опыта агрессии в межсамцовых конфронтациях обнаружено снижение экспрессии генов *Shank3* ($p < 0.010$), *Auts2* ($p < 0.023$), *Ctnnd2* ($p < 0.020$), *Nrxn2* ($p < 0.010$).

Были проанализированы изменения уровня экспрессии генов, кодирующих белки, вовлеченные в функционирование ГГ в гиппокампе мышей (см. табл. 1, рис. 3). Показано, что у мышей в гиппокампе под влиянием повторного опыта агрессии в межсамцовых конфронтациях снижена экспрессия гена *Grm4* ($p < 0.023$), кодирующего метаботропный рецептор 4-го подтипа. У жертв было обнаружено снижение экспрессии гена *Grin2a* ($p < 0.01$), кодирующего субъединицу 2a ионотропного NMDA-ре-

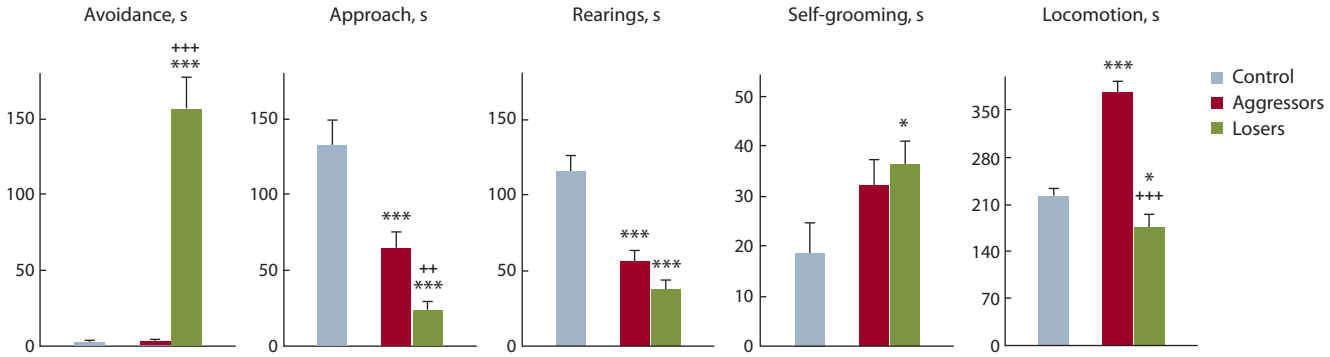


Fig. 1. Behavior of the aggressors and losers in the social interactions test.
* $p < 0.05$; *** $p < 0.001$ vs control; ++ $p < 0.01$; +++ $p < 0.001$ – losers vs aggressors.

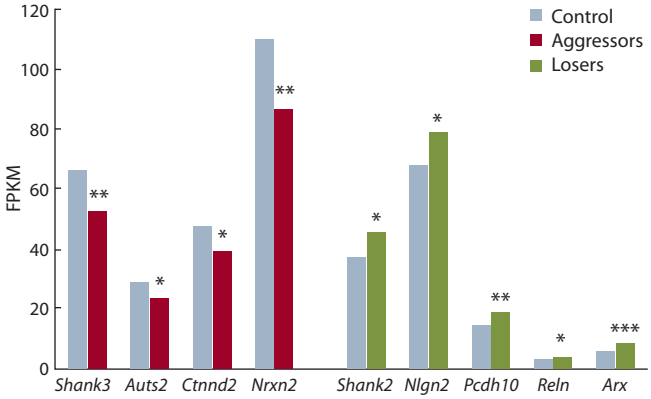


Fig. 2. Change in GA expression in the hippocampus of aggressors and losers.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ – vs control.

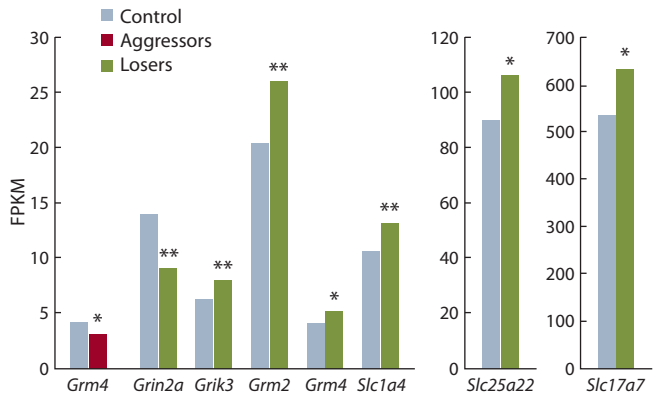


Fig. 3. Change in expression of GG in the hippocampus in mice with a disturbances of social behavior.

* $p < 0.05$, ** $p < 0.01$ – vs control.

Table 1. Differential expression GG and GA in hippocampus of aggressors and losers

Genes	Agressors	Losers
Associated with autistic spectrum disorders		
Shank3, SH3 and Multiple Ankyrin Repeat Domains 3	▽▽	
Auts2, Activator of Transcription and Developmental Regulator	▽	
Ctnnd2, Catenin Delta 2	▽	
Nrxn2, Neurexin 2	▽	
Shank2, SH3 and Multiple Ankyrin Repeat Domains 2		△
Nlgn2, Neuroligin 2		△
Pcdh10, Protocadherin 10		△△
Reln, Reelin		△
Arx, Aristaless Related Homeobox		△△△
Involved in the functioning of the glutaminergic system		
Grm4, Glutamate Metabotropic Receptor 4	▽	△
Grin2a, Glutamate Ionotropic Receptor NMDA Type Subunit 2A		▽▽
Grik3, Glutamate Ionotropic Receptor Kainate Type Subunit 3		△△
Grm2, Glutamate Metabotropic Receptor 2		△△
SLc17a7, Solute Carrier Family 17 Member 7		△
Slc1a4, Solute Carrier Family 1 Member 4		△△
Slc25a22, Solute Carrier Family 25 Member 22		△

Note: Decreased expression: ▽ – $p < 0.05$; ▽▽ – $p < 0.01$. Increased expression: △ – $p < 0.05$; △△ – $p < 0.01$; △△△ – $p < 0.001$ vs control.

цептора, при этом повышалась экспрессия генов *Grm2* ($p < 0.004$) и *Grm4* ($p < 0.02$), кодирующих метаботропные рецепторы 2 и 4-го подтипов, гена *Grik3* ($p < 0.003$), кодирующего субъединицу 3 ионотропного каинатного глутаматного рецептора, генов переносчиков глутамата *Slc17a7* ($p < 0.051$), *Slc1a4* ($p < 0.01$), *Slc25a22* ($p < 0.028$).

Результаты корреляционного анализа продемонстрировали высокий уровень корреляционной взаимосвязи между экспрессией ГГ и ГА у животных исследуемых групп (см. табл. 1 и 2). Корреляция между уровнями экспрессии ГА и ГГ показывает возможное участие глутаматэргической системы в механизмах этого заболевания. Из ГА наибольшее число корреляций обнаружено для генов *Nlgn2*, *Pcdh10*, *Arx*, *Ctnnd2*, *Nrxn2*. В то же время уровень экспрессии гена *Reln* коррелирует только с уровнем экспрессии гена *Slc1a4*.

Значения экспрессии некоторых ГГ и ГА в единицах FPKM, между которыми установлена статистически значимая коррелятивная взаимосвязь, показаны на рис. 4. Большая часть генов демонстрирует положительную кор-

Table 2. Co-expression of GG and GA (Pearson coefficient value)

Gene	*Nlgn2*	*Pcdh10*	*Reln*	*Shank2*	*Shank3*	*Arx*	*Auts2*	*Ctnnd2*	*Nrxn2*
Grm2	**0.761**	0.581	0.138	0.644	0.657	0.538	0.547	**0.692**	**0.668**
Grm4	**0.894**	**0.915**	0.537	**0.827**	**0.810**	**0.862**	0.602	**0.868**	**0.823**
Grin2a	**–0.825**	**–0.870**	–0.397	–0.643	–0.521	**–0.823**	–0.581	**–0.733**	**–0.867**
Grik3	**0.847**	**0.946**	0.597	**0.851**	**0.708**	**0.956**	**0.693**	**0.854**	**0.862**
Slc1a4	**0.677**	0.640	**0.696**	0.515	0.437	**0.714**	0.178	0.632	0.561
Slc17a7	**0.878**	**0.696**	0.231	**0.738**	**0.691**	**0.727**	**0.679**	**0.864**	**0.875**
Slc25a22	**0.922**	**0.918**	0.457	**0.906**	**0.792**	**0.940**	**0.802**	**0.933**	**0.935**

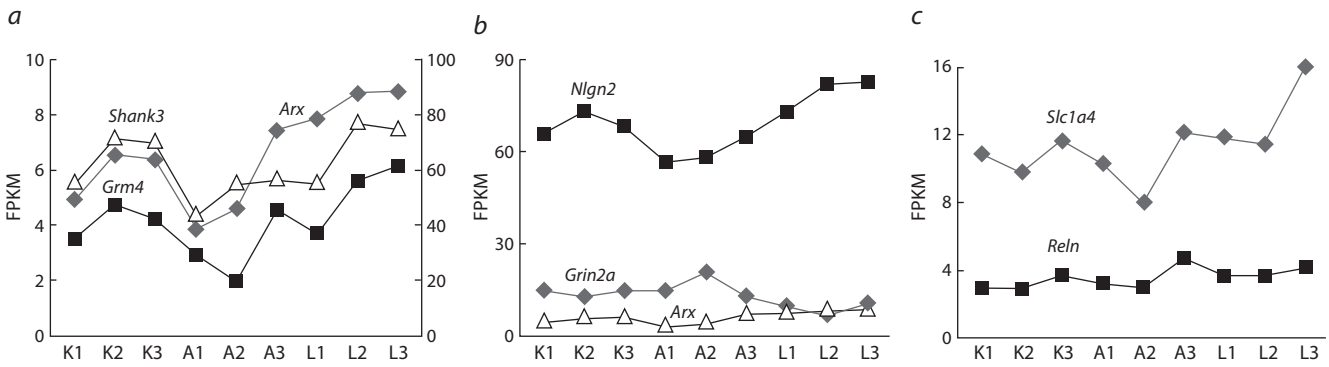


Fig. 4. Correlating GG and GA.

K1–3 – control animals; A1–3 – aggressors; L1–3 – losers. *a*, The right axis of ordinates is for the *Shank3* gene, the left one is *Arx* and *Grm4*.

реляцию друг с другом. Например, экспрессия *Grm4*, единственного гена, который изменил экспрессию и у жертв, и у агрессоров, положительно коррелирует с экспрессией ГА: *Shank3* ($r = 0.810$, $p < 0.01$) и *Arx* ($r = 0.862$, $p < 0.01$) (см. рис. 4, *а*). Экспрессия гена *Grin2a* демонстрирует отрицательную корреляцию с экспрессией ГА, например генов *Nlgn2* ($r = –0.823$, $p < 0.01$) и *Arx* ($r = –0.823$, $p < 0.01$) (см. рис. 4, *б*). Единственный ген, с которым коррелирует уровень экспрессии гена *Reln*, – это *Slc1a4* ($r = 0.696$, $p < 0.05$) (см. рис. 4, *в*).

Обсуждение

В наших предыдущих работах было обнаружено, что после 20 дней агонистических взаимодействий и проживания в условиях хронического социального конфликта у экспериментальных самцов нарушаются многие параметры социального поведения (Коваленко, Кудрявцева, 2010), схожие по симптоматике с заболеваниями аутистического спектра. Мы считаем, что в данном случае коморбидные аутистические симптомы развиваются на фоне других заболеваний. Это может быть тревожно-депрессивное состояние, возникающее вследствие хронического социального стресса у самцов мышей C57Bl/6 (Berton et al., 2006; Kudryavtseva et al., 2006) или каталепсии, развивающейся у животных CBA/Lac (Kudryavtseva et al., 2006), или же патологической агрессии (Kudryavtseva, 2006), формирующейся под влиянием повторного опыта агрессии. В литературе существуют данные, свидетельствующие, что коморбидность аутистических симптомов и расстройств настроения может быть следствием работы одних и тех же генов, продукты которых вовлекаются в патофизиологические механизмы развития этих заболеваний (Ragunath et al., 2011).

В связи с этим помимо генов, для которых показано участие в развитии симптомов аутизма, мы исследовали также изменения экспрессии ГГ (транспортеров, рецепторов, ферментов катаболизма), так как известно, что эта нейромедиаторная система вовлечена в развитие данной патологии (Carlsson, 2015).

Мы выявили в гиппокампе изменение экспрессии семи ГГ: ген *Grin2a* кодирует $NR2_A$-субъединицу NMDA-рецептора; *Grm2* и *Grm4* – метаботропные глутаматные рецепторы, mGluR2 и mGluR4 соответственно; *Grik3* – субъединицу каинатного рецептора, *Slc1a4* и *Slc17a7* – транспортеры глутамата Vglut1; *Slc25a22* – белок митохондриального переносчика глутамата. Изменили свою экспрессию 9 генов, ассоциированных с заболеваниями аутистического спектра: *Shank2*, *Shank3*, *Nlgn2*, *Reln*, *Arx*, *Nrxn2*, *Auts2*, *Ctnnd2*, *Pcdh10*.

Белки семейства SHANK, известные также как ProSAP, являются строительными белками на возбуждающих глутаматергических синапсах. Так, ген *Shank2* кодирует белок, взаимодействующий на постсинаптической мембране с NMDA-рецептором, участвуя в синаптической глутаматергической передаче (Carlsson, 2015). Вовлечение в процесс развития аутизма генов *Shank* было впервые описано для S*hank3*. Например, было отмечено, что мыши, гетерозиготные по белку Shank3 (Shank3+/delta-C), демонстрировали низкий уровень социального распознавания и коммуникативности (Wang et al., 2011). В литературе было

показано, что для аутистических симптомов характерно снижение функции *Shank2* (Won et al., 2012). Нарушения в социальном поведении характерны как для пониженного (Peça et al., 2011), так и для повышенного (Moessner et al., 2007) уровня белка SHANK3. В нашей работе также обнаружены повышенный уровень экспрессии гена *Shank2* и сниженный уровень гена *Shank3*. Вероятно, нарушения в социальном поведении могут быть связаны с любыми изменениями в функционировании этих белков. Известно, что ProSAP2/Shank3 влияет на работу глутаматергических синапсов, взаимодействуя с внеклеточным С-концом нейролигинов (Meyer et al., 2004). Мутации в этом гене ухудшают синаптическую передачу (Arons et al., 2012). Следовательно, можно предположить, что влияние белков SHANK на развитие аутистических симптомов осуществляется через изменение функционирования глутаматергической системы.

Значительную часть ГА, изменивших свой уровень экспрессии в нашей работе, составляют гены клеточной адгезии: *Nlgn2* – нейролигин 2-го типа; *Pcdh10* – протокадерин; *Ctnnd2* – δ-катенин; *Nrxn3* – нейрексин 3. Для всех этих белков многократными исследованиями было показано участие в формировании социального поведения и расстройств аутистического спектра. Известно, что изменение уровня белка NLGN2 может влиять на социальное и эмоциональное поведение (Maćkowiak et al., 2014). У аутистических больных обнаружены мутации в генах *Pcdh10* (Anitha et al., 2013), *Ctnnd2* (Turner et al., 2015), *Nrxn3* (Vaags et al., 2012). Кроме того, продукты этих генов взаимодействуют с глутаматергической системой. Так, например, нейролигины вызывают кластеризацию Vglut-положительных синаптических пузырьков (Graf et al., 2004), без чего невозможно созревание и функционирование синапса. Правда, там же отмечено, что это характерно для нейролигина 1, а нейролигин 2 не располагается на одном синапсе с транспортером глутамата. Обнаружено, что в отсутствие нейролигинов снижается число Vglut1-положительных терминалей (Chih et al., 2005). Можно предположить, что на самом деле нейролигин взаимодействует непосредственно с NMDA-рецепторами, а переносчик Vglut1 выступает только в качестве маркера этих рецепторов. Что касается продукта гена *Pcdh10*, то установлено, что протокадерины локализуются на нейронах, формирующих глутаматергические AMPA и каинатные рецепторы (Puller, Haverkamp, 2011). Кроме того, показано, что у мышей, нокаутных по генам протокадеринов, снижено число глутаматных транспортеров Vglut 1–2 (Chen et al., 2012), что также свидетельствует о взаимосвязи между протокадеринами и глутаматергической системой.

Продукты генов *Arx* и *Reln* (белок рилин) играют важную роль в механизмах пре- и постнатального нейрогенеза. Участие этих генов в патофизиологии аутизма является спорным: часть авторов приводят данные в пользу этого факта (Wall et al., 2009), в то время как другие опровергают его (Persico et al., 2001). В то же время обнаружена связь этих генов с функционированием глутаматергической системы. Так, известно, что мутации в гене *Arx* связаны с изменениями в глутаматергической системе у больных эпилепсией (Beguin et al., 2013), а белок рилин способен повышать мобильность рецепторов, содержащих $NR2_B$-субъединицу, снижая время ее пребывания в синапсе, таким образом, изменяя состав рецепторов в сторону преобладания $NR2_A$-субъединицы (Groc et al., 2007).

Ген *Auts2* кодирует экспрессирующийся в головном мозге белок с неизвестной функцией, хотя предполагают, что он вовлечен в механизмы развития нервной системы (Oksenberg et al., 2013). Мутации в этом гене обнаружены у пациентов с аутистическими расстройствами (Liu et al., 2015). Показано, что *Auts2* экспрессируется на глутаматергических нейронах гиппокампа (Hori et al., 2014).

Так как большинство ГА, изменивших экспрессию в нашем исследовании, – гены, связанные с клеточным скелетом и внеклеточным матриксом, в частности участвующие в формировании синапсов, а ГГ – гены, кодирующие субъединицы рецепторов, то можно полагать, что изменение активности глутаматергической системы проявляется в аутистических симптомах.

Корреляционный анализ наших экспериментальных данных в целом выявил высокий уровень корреляции между экспрессией ГГ и ГА. Исключение составил ген *Reln*, коррелирующий только с геном *Slc1a4*. Характерно также, что везде обнаруживается положительная корреляция между генами, в то время как ген *Grin2a* отрицательно коррелирует с ГА. Однако снижение экспрессии *Grin2a*, продукт которого является субъединицей NMDA-рецептора, хорошо согласуется с литературными данными, связывающими развитие аутистических симптомов с дефицитом этих рецепторов (Lee et al., 2015). Экспрессия гена *Nlgn2* коррелирует с экспрессией всех ГГ, экспрессия генов *Pcdh10*, *Arx*, *Cttnd2*, *Nrxn2* – с экспрессией практически всех генов. Это позволяет предположить, что именно при помощи данных генов ГГ вовлекается в механизмы развития аутистического поведения. Можно видеть, что все наиболее коррелирующие гены, кроме *Arx*, – гены клеточной адгезии. Мы предполагаем, что продукты этих генов участвуют в соединении синапсов, в том числе глутаматергических, с клеточной мембраной. Таким образом, они вовлекаются в функционирование глутаматергической системы. Следовательно, изменение экспрессии этих генов изменяет уровень активности ГГ, что, в свою очередь, приводит к изменениям в социальном поведении, хотя нельзя исключить и обратный вариант взаимодействия.

Заключение

Проведенное исследование подтвердило полученные ранее результаты, свидетельствующие о том, что под влиянием хронического социального конфликта у самцов мышей могут развиваться нарушения социального поведения, причем в этом процессе участвуют как гены, кодирующие белки, вовлеченные в функционирование глутаматергической системы, так и гены, связанные с патологией аутизма. Коэкспрессия этих генов позволяет говорить о вовлечении глутаматергической системы головного мозга в развитие патологии социального поведения. Полученные результаты могут служить доказательством того, что аутистические симптомы могут не только быть следствием генетических нарушений, но и развиваться в течение жизни индивидуума под влиянием стрессорных воздействий.

Список литературы / References

Галямина А.Г., Коваленко И.Л., Смагин Д.А., Кудрявцева Н.Н. Изменение экспрессии генов нейромедиаторных систем в вентральной тегментальной области депрессивных мышей: данные RNA-Seq. Журн. высш. нервн. деят. им. И.П. Павлова. 2017; 67(1):113-128. DOI 10.7868/S0044467717010063.
[Galyamina A.G., Kovalenko I.L., Smagin D.A., Kudryavtseva N.N. Altered expression of neurotransmitters systems' genes in the ventral tegmental area of depressive male mice: data of RNA-Seq. Zhurnal Vysshey Nervnoy Deyatel'nosti imeni I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity. 2017;67(1):113-128. DOI 10.7868/S0044467717010063. (in Russian)]

Коваленко И.Л., Кудрявцева Н.Н. Развитие симптомов аутистического спектра под влиянием хронического социального стресса у тревожных самцов мышей: эффекты диазепама. Психофармакол. биол. наркология. 2010;10(1-2):2624-2635.
[Kovalenko I.L., Kudryvtseva N.N. Development of autistic spectrum symptoms under chronic social defeat stress in anxious male mice: effects of diazepam. Psikhopharmakologiya and Biologicheskaya Narkologiya = Psychopharmacology and Biological Narcology. 2010;10(1-2):2624-2635.

American Psychiatric Association. DSM-IV Draft Criteria. Washington, DC, American Psychiatric Press, 1993.

Anitha A., Thanseem I., Nakamura K., Yamada K., Iwayama Y., Toyota T., Iwata Y., Suzuki K., Sugiyama T., Tsujii M., Yoshikawa T., Mori N. Protocadherin α (PCDHA) as a novel susceptibility gene for autism. J. Psychiatry Neurosci. 2013;38(3):192-198. DOI 10.1503/jpn.120058.

Arons M.H., Thynne C.J., Grabrucker A.M., Li D., Schoen M., Cheyne J.E., Boeckers T.M., Montgomery J.M., Garner C.C.J. Autism-associated mutations in ProSAP2/Shank3 impair synaptic transmission and neurexin-neuroligin-mediated transsynaptic signaling. J. Neurosci. 2012;32(43):14966-14978. DOI 10.1523/JNEUROSCI.2215-12.2012.

Babenko V.N., Smagin D.A., Kudryavtseva N.N. RNA-Seq mouse brain regions expression data analysis: Focus on *ApoE* functional network. J. Integr. Bioinform. 2017;14(3):20170024. DOI 10.1515/jib-2017-0024.

Bauman M.L., Kemper T.L. Neuroanatomic observations of the brain in autism: a review and future directions. Int. J. Dev. Neurosci. 2005; 23(2-3):183-187. DOI 10.1016/j.ijdevneu.2004.09.006.

Beguin S., Crépel V., Aniksztejn L., Becq H., Pelosi B., Pallesi-Pocachard E., Bouamrane L., Pasqualetti M., Kitamura K., Cardoso C., Represa A. An epilepsy-related ARX polyalanine expansion modifies glutamatergic neurons excitability and morphology without affecting GABAergic neurons development. Cereb. Cortex. 2013; 23(6):1484-1494. DOI 10.1093/cercor/bhs138.

Berton O., McClung C.A., Dileone R.J., Krishnan V., Renthal W., Russo S.J., Graham D., Tsankova N.M., Bolanos C.A., Rios M., Monteggia L.M., Self D.W., Nestler E.J. Essential role of BDNF in the mesolimbic dopamine pathway in social defeat stress. Science. 2006;311:864-886. DOI 10.1126/science.1120972.

Carlsson M. Autism and glutamate. In: Fatemi S.H. (Ed.). The Molecular Basis of Autism. New York: Springer, 2015;243-256.

Chen W.V., Alvarez F.J., Lefebvre J.L., Friedman B., Nwakeze C., Geiman E., Smith C., Thu C.A., Tapia J.C., Tasic B., Sanes J.R., Maniatis T. Functional significance of isoform diversification in the protocadherin-gamma gene cluster. Neuron. 2012;75(3):402-409. DOI 10.1016/j.neuron.2012.06.039.

Chih B., Engelman H., Scheiffele P. Control of excitatory and inhibitory synapse formation by neuroligins. Science. 2005;307:1324-1328. DOI 10.1126/science.1107470.

DeLong G.R. Autism, amnesia, hippocampus, and learning. Neurosci. Biobehav. Rev. 1992;16:63-70.

Faure A., Richard J.M., Berridge K.C. Desire and dread from the nucleus accumbens: cortical glutamate and subcortical GABA differentially generate motivation and hedonic impact in the rat. PLoS One. 2010;5(6):e11223. DOI 10.1371/journal.pone.0011223.

Graf E.R., Zhang X., Jin S.X., Linhoff M.W., Craig A.M. Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. Cell. 2004;119(7):1013-1026. DOI 10.1016/j.cell.2004.11.035.

Groc L., Choquet D., Stephenson F.A., Verrier D., Manzoni O.J., Chavis P. NMDA receptor surface trafficking and synaptic subunit composition are developmentally regulated by the extracellular matrix protein reelin. J. Neurosci. 2007;27(38):10165-10175. DOI 10.1523/JNEUROSCI.1772-07.2007.

Hallmayer J., Cleveland S., Torres A., Phillips J., Cohen B., Torigoe T., Miller J., Fedele A., Collins J., Smith K., Lotspeich L., Croen L.A., Ozonoff S., Lajonchere C., Grether J.K., Risch N. Genetic heritability and shared environmental factors among twin pairs with autism. Arch. Gen. Psychiatry. 2011;68(11):1095-1102. DOI 10.1001/archgenpsychiatry.2011.76.

Hori K., Nagai T., Shan W., Sakamoto A., Taya S., Hashimoto R., Hayashi T., Abe M., Yamazaki M., Nakao K., Nishioka T., Sakimura K., Yamada K., Kaibuchi K., Hoshino M. Cytoskeletal regulation by AUTS2 in neuronal migration and neuritogenesis. Cell Rep. 2014;9(6):2166-2179. DOI 10.1016/j.celrep.2014.11.045.

Irle E., Ruhleder M., Lange C., Seidler-Brandler U., Salzer S., Dechent P., Weniger G., Leibing E., Leichsenring F. Reduced amygdalar and hippocampal size in adults with generalized social phobia. J. Psychiatry Neurosci. 2010;35:126-131.

Jamain S., Betancur C., Quash H. Linkage and association of the glutamate receptor 6 gene with autism. Paris autism research international sibpair [PARIS] study. Mol. Psychiatry. 2002;7(3):302-310.

Kadakkuzha B.M., Liu X.A., McCrate J., Shankar G., Rizzo V., Afinogenova A., Young B., Fallahi M., Carvalloza A.C., Raveendra B., Puthanveettil S.V. Transcriptome analyses of adult mouse brain reveal enrichment of lncRNAs in specific brain regions and neuronal populations. Front. Cell. Neurosci. 2015;9:63. DOI 10.3389/fncel.2015.00063.

Kleijer K.T., Schmeisser M.J., Krueger D.D., Boeckers T.M., Scheiffele P., Bourgeron T., Brose N., Burbach J.P. Neurobiology of autism gene products: towards pathogenesis and drug targets. Psychopharmacology (Berl.). 2014;231(6):1037-1062. DOI 10.1007/s00213-013-3403-3.

Kudryavtseva N.N. A sensory contact model for the study of aggressive and submissive behavior in male mice. Aggress. Behav. 1991; 17(5):285-291. DOI 10.1007/s00210-018-1580-7.

Kudryavtseva N.N. The psychopathology of repeated aggression: a neurobiological aspect. In: Morgan J.P. (Ed.). Perspectives on the Psychology of Aggression. New York: Nova Science Publishers, 2006;35-64.

Kudryavtseva N.N., Avgustinovich D.F., Bakshtanovskaya I.V., Koryakina L.A., Alekseyenko O.V., Lipina T.V., Bondar N.P. Experimental studies of hereditary predisposition to the development of depression. In: Kalueff A. (Ed.). Animal Models of Biological Psychiatry. New York: Nova Science Publishers, 2006;75-95.

Kudryavtseva N.N., Kovalenko I.L., Smagin D.A., Galyamina A.G., Babenko V.N. Abnormality of social behavior and dysfunction of autism related gene expression developing under chronic social defeat stress in male mice. Eur. Neuropsychopharm. 2017;27(4):S678-S679.

Kudryavtseva N.N., Smagin D.A., Kovalenko I.L., Vishnivetskaya G.B. Protocol for repeated positive fighting experience in male mice: the tool for the study of movementdisorders. Nature Protoc. 2014;9(11): 2705-2717.

Lee E.J., Choi S.Y., Kim E.NMDA receptor dysfunction in autism spectrum disorders. Curr. Opin. Pharmacol. 2015;20:8-13. DOI 10.1016/j.coph.2014.10.007.

Liu Y., Zhao D., Dong R., Yang X., Zhang Y., Tammimies K., Uddin M., Scherer S.W., Gai Z. De novo exon1 deletion of *AUTS2* gene in a patient with autism spectrum disorder and developmental delay: a case report and a brief literature review. Am. J. Med. Genet. A. 2015;167(6):1381-1385. DOI 10.1002/ajmg.a.37050.

Maćkowiak M., Mordalska P., Wędzony K. Neuroligins, synapse balance and neuropsychiatric disorders. Pharmacol. Rep. 2014;66: 830-835. DOI 10.1016/j.pharep.2014.04.011.

Meyer G., Varoqueaux F., Neeb A., Oschlies M., Brose N. The complexity of PDZ domain-mediated interactions at glutamatergic synapses: a case study on neuroligin. Neuropharmacology. 2004;47:724-733. DOI 10.1016/j.neuropharm.2004.06.023.

Miles J.H. Autism spectrum disorders, a genetics review. Genet. Med. 2011;13(4):278-294. DOI 10.1097/GIM.0b013e3181ff67ba.

Moessner R., Marshall C.R., Sutcliffe J.S., Skaug J., Pinto D., Vincent J., Zwaigenbaum L., Fernandez B., Roberts W., Szatmari P., Scherer S.W. Contribution of SHANK3 mutations to autism spectrum disorder. Am. J. Hum. Genet. 2007;81:1289-1297. DOI 10.1086/522590.

Morgane P.J., Galler J.R., Mokler D.J. A review of systems and networks of the limbicforebrain/limbicmidbrain. Prog. Neurobiol. 2005;75(2):143-160. DOI 10.1016/j.pneurobio.2005.01.001.

Oksenberg N., Stevison L., Wall J.D., Ahituv N. Function and regulation of *AUTS2*, a gene implicated in autism and human evolution. PLoS Genet. 2013;9(1):e1003221. DOI 10.1371/journal.pgen.1003221.

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

Peça J., Feliciano C., Ting J.T., Wang W., Wells M.F., Venkatraman T.N., Lascola C.D., Fu Z., Feng G. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. Nature. 2011; 472(7344):437-442. DOI 10.1038/nature09965.

Persico A.M., D'Agruma L., Maiorano N., Totaro A., Militerni R., Bravaccio C., Wassink T.H., Schneider C., Melmed R., Trillo S., Montecchi F., Palermo M., Pascucci T., Puglisi-Allegra S., Reichelt K.L., Conciatori M., Marino R., Quattrocchi C.C., Baldi A., Zelante L., Gasparini P., Keller F.; Collaborative Linkage Study of Autism. Reelin gene alleles and haplotypes as a fact or predisposing to autistic disorder. Mol. Psychiatry. 2001;6(2):150-159. DOI 10.1038/sj.mp.4000850.

Puller C., Haverkamp S. Cell-type-specific localization of protocadherin β16 at AMPA and AMPA/Kainate receptor-containing synapses in the primate retina. J. Comp. Neurol. 2011;519(3):467-479. DOI 10.3233/DMA-2012-0917.

Purcell A.E., Jeon O.H., Pevsner J. The abnormal regulation of gene expression in autistic brain tissue. J. Autism Dev. Disord. 2001;31(6): 545-549.

Ragunath P.K., Chitra R., Mohammad S., Abhinand P.A. A systems biological study on the comorbidity of autism spectrum disorders and bipolar disorder. Bioinformation. 2011;7(3):102-106.

Ramoz N., Reichert J.G., Smith C.J., Silverman J.M., Bespalova I.N., Davis K.L., Buxbaum J.D. Linkage and association of the mitochondrial aspartate/glutamate carrier SLC25A12 gene with autism. Am. J. Psychiatry. 2004;161(4):662-669. DOI 10.1176/appi.ajp.161.4.662.

Riedel G., Platt B., Micheau J. Glutamate receptorfunction in learning and memory. Behav. Brain Res. 2003;140(1-2):1-47.

Røjas D.C. The role of glutamate and its receptors in autism and the use of glutamate receptor antagonists in treatment. J. Neural. Transm. (Vienna). 2014;121(8):891-905. DOI 10.1007/s00702-014-1216-0.

Savitz J.B., Drevets W.C. Imaging phenotypes of major depressive disorder: genetic correlates. Neuroscience. 2009;164:300-330. DOI 10.1016/j.neuroscience.2009.03.082.

Simonyi A., Schachtman T.R., Christoffersen G.R. Metabotropic glutamate receptor subtype 5 antagonism in learning and memory. Eur. J. Pharmacol. 2010;639(1-3):17-25. DOI 10.1016/j.ejphar.2009.12.039.

Smagin D.A., Boyarskikh U.A., Bondar N.P., Filipenko M.L., Kudryavtseva N.N. Reduction of serotonergic gene expression in the midbrain raphe nuclei under positive fighting experience. Adv. Biosci. Biotechnol. 2013;4(10B):36-44. DOI 10.4236/abb.2013.410A3005.

Turner T.N., Sharma K., Oh E.C., Liu Y.P., Collins R.L., Sosa M.X., Auer D.R., Brand H., Sanders S.J., Moreno-De-Luca D., Pihur V., Plona T., Pike K., Soppet D.R., Smith M.W., Cheung S.W., Martin C.L., State M.W., Talkowski M.E., Cook E., Huganir R., Katsanis N., Chakravarti A. Loss of δ-catenin function in severe autism. Nature. 2015;520(7545):51-56. DOI 10.1038/nature14186.

Vaags A.K., Lionel A.C., Sato D., Goodenberger M., Stein Q.P., Curran S., Ogilvie C., Ahn J.W., Drmic I., Senman L., Chrysler C., Thompson A., Russell C., Prasad A., Walker S., Pinto D., Marshall C.R., Stavropoulos D.J., Zwaigenbaum L., Fernandez B.A., Fombonne E., Bolton P.F., Collier D.A., Hodge J., Roberts W., Szatmari P., Scherer S.W. Rare deletions at the neurexin 3 locus in autism spectrum disorder. Am. J. Hum. Genet. 2012;90(1):133-141. DOI 10.1016/j.ajhg.2011.11.025.

Wall D.P., Esteban F.J., Deluca T.F., Huyck M., Monaghan T., Velez de Mendizabal N., Goñí J., Kohane I.S. Comparative analysis of neurological disorders focuses genome-wide search for autism genes. Genomics. 2009;93(2):120-129. DOI 10.1016/j.ygeno.2008.09.015.

Wang X., McCoy P.A., Rodriguiz R.M., Pan Y., Je H.S., Roberts A.C., Kim C.J., Berrios J., Colvin J.S., Bousquet-Moore D., Lorenzo I., Wu G., Weinberg R.J., Ehlers M.D., Philpot B.D., Beaudet A.L., Wetsel W.C., Jiang Y.H. Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. Hum. Mol. Genet. 2011;20:3093-3108. DOI 10.1093/hmg/ddr212.

Won H., Lee H.R., Gee H.Y., Mah W., Kim J.I., Lee J., Ha S., Chung C., Jung E.S., Cho Y.S., Park S.G., Lee J.S., Lee K., Kim D., Bae Y.C., Kaang B.K., Lee M.G., Kim E. Autistic-like social behaviour in Shank2-mutant mice improved by restoring NMDA receptor function. Nature. 2012;486:261-265. DOI 10.1038/nature11208.

Zablotsky B., Black L.I., Maenner M.J., Schieve L.A., Blumberg S.J. Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 national health interview survey. Natl. Health Stat. Report. 2015;87:1-20.

Zhang Y., Chen K., Sloan S.A., Bennett M.L., Scholze A.R., O'Keeffe S., Phatnani H.P., Guarnieri P., Caneda C., Ruderisch N., Deng S., Liddelow S.A., Zhang C., Daneman R., Maniatis T., Barres B.A., Wu J.Q. An RNA-sequence and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J. Neurosci. 2014;34(36): 11929-11947. DOI 10.1523/JNEUROSCI.1860-14.2014.

Zheng C., Zhu Q., Liu X., Huang X., He C., Jiang L., Quan D., Zhou X., Zhu Z. Effect of platelet-rich plasma (PRP) concentration on proliferation, neurotrophic function and migration of Schwann cells *in vitro*. J. Tissue Eng. Regen. Med. 2016;10(5):428-436. DOI 10.1002/term.175.

ORCID ID

I.L. Kovalenko orcid.org/0000-0002-6672-1287
A.G. Galyamina orcid.org/0000-0001-6101-2521
D.A. Smagin orcid.org/0000-0002-4017-7287
N.N. Kudryavtseva orcid.org/0000-0003-2418-1750

Acknowledgements. This work was supported by the Russian Foundation for Basic Research, project 17-04-00140, and State Budgeted Projects 0324-2019-0041 and 0324-2019-0041-C-01.

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

Received July 5, 2019. Revised October 17, 2019. Accepted November 8, 2019. Published online February 19, 2020.

Вавиловский журнал генетики и селекции 2020;24(2):200-208
DOI 10.18699/VJ20.40-o

Pharmacological effects of fibroblast growth factor 21 are sex-specific in mice with the *lethal yellow* (A^y) mutation

E.N. Makarova[1] ✉, T.V. Yakovleva[1], N.Yu. Balyibina[2], K.O. Baranov[3], E.I. Denisova[1], A.D. Dubinina[1], N.A. Feofanova[4], N.M. Bazhan[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
[3] The Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[4] Research Institute of Fundamental and Clinical Immunology, Novosibirsk, Russia
✉ e-mail: enmakarova@gmail.com

Abstract. Hypothalamic melanocortin 4 receptors (MC4R) regulate energy balance. Mutations in the *MC4R* gene are the most common cause of monogenic obesity in humans. Fibroblast growth factor 21 (FGF21) is a promising anti-obesity agent, but its effects on melanocortin obesity are unknown. Sex is an important biological variable that must be considered when conducting preclinical studies; however, in laboratory animal models, the pharmacological effects of FGF21 are well documented only for male mice. We aimed at investigating whether FGF21 affects metabolism in male and female mice with the *lethal yellow* (A^y) mutation, which results in MC4R blockage and obesity development. Obese C57Bl-A^y male and female mice were administered subcutaneously for 10 days with vehicle or FGF21 (1 mg per 1 kg). Food intake (FI), body weight (BW), blood parameters, and gene expression in the liver, muscles, brown adipose tissue, subcutaneous and visceral white adipose tissues, and hypothalamus were measured. FGF21 action strongly depended on the sex of the animals. In the males, FGF21 decreased BW and insulin blood levels without affecting FI. In the females, FGF21 increased FI and liver weight, but did not affect BW. In control A^y-mice, expression of genes involved in lipid and glucose metabolism (*Ppargc1a*, *Cpt1*, *Pck1*, *G6p*, *Slc2a2*) in the liver and genes involved in lipogenesis (*Pparg*, *Lpl*, *Slc2a4*) in visceral adipose tissue was higher in females than in males, and FGF21 administration inhibited the expression of these genes in females. FGF21 administration decreased hypothalamic POMC mRNA only in males. Thus, the pharmacological effect of FGF21 were significantly different in male and female A^y-mice; unlike males, females were resistant to catabolic effects of FGF21.
Key words: FGF21; A^y-mice; melanocortin obesity; sex differences; liver; hypothalamus.

For citation: Makarova E.N., Yakovleva T.V., Balyibina N.Yu., Baranov K.O., Denisova E.I., Dubinina A.D., Feofanova N.A., Bazhan N.M. Pharmacological effects of fibroblast growth factor 21 are sex-specific in mice with the *lethal yellow* (A^y) mutation. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):200-208. DOI 10.18699/VJ20.40-o

У мышей с мутацией *lethal yellow* (A^y) фармакологические эффекты фактора роста фибробластов 21 зависят от пола

Е.Н. Макарова[1] ✉, Т.В. Яковлева[1], Н.Ю. Балыбина[2], К.О. Баранов[3], Е.И. Денисова[1], А.Д. Дубинина[1], Н.А. Феофанова[4], Н.М. Бажан[1, 2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
[4] Научно-исследовательский институт фундаментальной и клинической иммунологии, Новосибирск, Россия
✉ e-mail: enmakarova@gmail.com

Аннотация. Гипоталамические меланокортиновые рецепторы 4-го типа (МК4Р) принимают участие в поддержании баланса энергии. Мутации в гене, кодирующем МК4Р, – наиболее распространенная причина моногенного ожирения у людей. Фактор роста фибробластов 21 (FGF21) рассматривают в качестве перспективного кандидата для медикаментозного лечения ожирения, однако неизвестно, влияет ли он на меланокортиновое ожирение. Пол особей необходимо учитывать при проведении доклинических исследований, но у лабораторных животных фармакологические эффекты FGF21 изучали только на самцах. В настоящей работе исследовано влияние FGF21 на метаболизм у самцов и самок мышей с мутацией *lethal yellow* в локусе агути (A^y), которая приводит к блокаде МК4Р в гипоталамусе и развитию ожирения. Самцам и самкам мышей A^y линии C57Bl с развитым ожирением вводили подкожно в течение 10 дней физиологический раствор или FGF21 (1 мг/кг). Измеряли потребление пищи (ПП), массу тела (МТ), показатели крови и экспрессию генов в печени, мышцах, бурой жировой ткани, подкожной и висцеральной белой жировой ткани и гипоталамусе. Эффекты FGF21 зависели от пола животных. У самцов FGF21 снижал МТ и уровень инсулина в крови и не влиял на ПП. У самок FGF21 увеличивал ПП и массу печени, но не влиял на МТ. У контрольных самок A^y экспрессия генов углеводно-жирового обмена (*Ppargc1a*, *Cpt1*, *Pck1*, *G6p*, *Slc2a2*) в печени и генов липогенеза (*Pparg*, *Lpl*, *Slc2a4*) в висцеральной жировой ткани была выше, чем у самцов,

и введение FGF21 снижало экспрессию этих генов только у самок. Введение FGF21 уменьшало уровень мРНК ПОМК в гипоталамусе только у самцов. Полученные результаты демонстрируют, что фармакологический эффект FGF21 значительно различается у самцов и самок мышей с мутацией A^y: в отличие от самцов, самки проявляют устойчивость к катаболическому действию FGF21.
Ключевые слова: FGF21; мыши A^y; меланокортиновое ожирение; половые различия; печень; гипоталамус.

Introduction

Obesity is a serious problem in modern society, this being the reason why various methods of combating obesity (medicinal, non-medicinal, preventive, etc.) are under intensive investigation. Hypothalamus plays a critical role in coordination of energy homeostasis, and mutations in various hypothalamic genes responsible for controlling appetite and metabolism lead to obesity (Singh et al., 2017). Melanocortin (MC) obesity, caused by mutations in the melanocortin system of the brain, is the most common genetic form of obesity in humans (Farooqi et al., 2003; Girardet, Butler, 2014). Melanocortin system regulates energy intake and expenditure. Activation of type 4 melanocortin receptors (MC4R) in the hypothalamic neurons reduces food consumption and increases energy expenditure, while their blockade or loss (knockout) is associated with hyperphagia, gradual development of obesity, and insulin resistance (Tao, 2010). In humans, the loss of MC4R functions causes severe obesity (Farooqi et al., 2003), but intensive search for therapeutic options of MC obesity correction has yet not identified an efficient drug (Fani et al., 2014).

Fibroblast Growth Factor 21 (FGF21) is assumed to be one of the most promising candidates for obesity treatment, because administration of FGF21 or its analogs was shown to reduce body weight in laboratory rodents, monkeys, and humans (Jackson et al., 2015). In rodents, it is efficient against both diet-induced and genetic forms of obesity (leptin *ob*/*ob* or its receptor *db*/*db* deficiency) (Kharitonenkov et al., 2005; Coskun et al., 2008). FGF21 is an atypical member of the fibroblast growth factor family; it possesses a hormone-like activity and is involved in maintaining energy homeostasis, regulation of carbohydrate and lipid metabolism, and adaptation to various stresses, including metabolic, such as nutritional deficiencies and calorie overload (Xie, Leung, 2017). FGF21 induces weight loss through its effects on the central nervous system (Lan et al., 2017). It is not known whether melanocortin system is involved in signal transmission from FGF21 to the CNS. If melanocortin signaling pathways are involved in the central action of FGF21, loss of function of the melanocortin system could reduce or eliminate the beneficial effect of FGF21 on metabolism and body weight. However, the pharmacological effects of FGF21 have not been studied in melanocortin obesity models.

Most animal studies of physiological and pharmacological effects of FGF21 have been made on males (Kharitonenkov et al., 2005; Coskun et al., 2008; Xu et al., 2009; Camporez et al., 2013; Markan et al., 2014). However, sex steroids have such a significant effect on the regulation of metabolic processes that National Institutes of Health (NIH) recognized sex as an important biological variable that must be considered when conducting preclinical studies (Mauvais-Jarvis et al., 2017; Clayton, 2018). In a few studies performed on rats and mice of both sexes, sex differences in the expression of FGF21 in liver (Lee et al., 2016; Chukijrungroat et al., 2017) and other tissues (Gasparin et al., 2018) were observed, exhibiting differential manifestation in obesity and starvation (Bazhan et al., 2019). These data suggest that the physiological and pharmacological effects of FGF21 may vary in individuals of different sexes.

The objective of this study was to investigate the pharmacological effects of FGF21 in male and female mice with melanocortin obesity. As a model of melanocortin obesity, we used mice with the *lethal yellow* mutation at the *agouti* locus (A^y). In mice, A^y mutation causes ectopic overexpression of the *agouti* gene (Bultman et al., 1992). A^y-mice have yellow coat color and develop obesity and non-insulin-dependent diabetes with age (Wolff et al., 1999), due to ectopic expression of *agouti* gene in the hypothalamus, which evokes chronic blockage of MC4Rs by the agouti protein (Michaud et al., 1997).

We found that therapeutic effects of FGF21 in A^y-mice strongly depended on the sex. In male A^y-mice, the blockage of MC4Rs did not prevent anti-obesity effect of FGF21, and its administration resulted in weight loss and decreased blood insulin levels. In females, FGF21 administration increased food intake without reducing body weight and glucose and insulin concentrations in blood, but inhibited the expression of genes related to glucose and lipid turnover in liver and increased liver weight. Thus, female A^y-mice were resistant to anti-obesity effects of FGF21.

Materials and methods

Ethical approval. All experiments were performed according to the Guide for the Care and Use of Laboratory Animals (1996) and the Russian National Instructions for the Care and Use of Laboratory Animals. The protocols were approved by the Independent Ethics Committee of the Institute of Cytology and Genetics (Siberian Branch of the Russian Academy of Sciences).

Animals. C57Bl and C57Bl-A^y mice were bred in the vivarium of the Institute of Cytology and Genetics in reciprocal crosses. The mice were separated from their mothers at the age of 4 weeks and housed in groups of 5–6 per cage. At the age of 30 weeks, each mouse was placed into a separate cage and housed individually until the beginning of the experiment. The mice were housed under a 12/12-h light-dark regime (light from 07:30 to 19:30) at an ambient temperature of 22–24 °C. The mice were provided *ad libitum* access to commercial mouse chow (Assortiment Agro, Turakovo Village, Moscow region, Russia) and water.

FGF21 (1 mg per 1 kg) or PBS were administered subcutaneously at the end of the light period (17:00–17:30) for 10 days. We have chosen this dose based on literature data. T. Coskun et al. (2008) showed that daily FGF21administration at this dose reduced body weight and blood glucose concentrations in male mice. To reveal the effect of FGF21 on glycaemia, fasted blood glucose was measured before and during the experiment. The mice were fasted overnight for two days before the first injection and at the seventh day of the experiment (after seven injections of FGF21 or PBS), and blood glucose was measured at the end of fasting. Glucose

concentrations were measured using a Lifescan One Touch Basic Plus glucometer. Body weight and food intake were measured daily for 6 days prior to fasting and within 24 h of refeeding after fasting.

On the last day, the animals were sacrificed by decapitation (an hour after the injection), and samples of trunk blood were collected; liver, brown adipose tissue (BAT), and subcutaneous and abdominal white adipose tissues (WAT) were weighed, and the tissues were collected and snap-frozen in liquid nitrogen to evaluate gene expression. Seven male and six female mice received PBS (control); six male and five female mice received FGF21.

Plasma assays. Concentrations of insulin, leptin, and adiponectin were measured, respectively, using Rat/Mouse Insulin ELISA Kit, Mouse Leptin ELISA Kit (EMD Millipore, St. Charles, Missouri, USA), and Mouse Adiponectin ELISA Kit (EMD Millipore, Billerica, MA, USA). Concentrations of glucose, triglycerides and cholesterol were measured colorimetrically using, respectively, Fluitest GLU, Fluitest TG, and Fluitest CHOL (Analyticon® Biotechnologies AG Am Mühlenberg 10, 35104 Lichtenfels, Germany). Concentrations of free fatty acids were measured using NEFA FS DiaSys kits (Diagnostic Systems GmbH, Holzheim, Germany).

Expression and purification of mouse FGF21. Mouse FGF21 coding sequence (aa 29 to 210) was optimized for *Escherichia coli* expression and synthesized by Genewiz (South Plainfield, NJ, USA). This DNA sequence was subcloned into the expression vector pE-SUMOpro (LifeSensors Inc., USA). This construct was used for induction of fusion 6xHis-SUMO-fgf21 protein in *E. coli* BL21 (DE3) cells. The purified 6xHis-SUMO-fgf21 was cleaved using SUMO protease 1 and loaded onto a column with Ni-NTA resin. The FGF21 protein (aa 29 to 210) was in the flow-through fractions. Size exclusion chromatography on a Superdex 200 10/300 GL column was used as a final purification step. The absence of bacterial endotoxins in FGF21 protein sample was confirmed by LAL-test (< 0.2 U/μg protein).

Relative quantitation real-time PCR. Total RNA was isolated from tissue samples using ExtractRNA kit (Evrogen, Moscow, Russia) according to the manufacturer's instructions. First-strand cDNA was synthesized using Moloney murine leukemia virus (MMLV) reverse transcriptase (Evrogen) and oligo(dT) as a primer. TaqMan gene expression assays (Applied Biosystems) listed in Table 1 were used for relative quantitation real-time PCR with β-actin as an endogenous control according to manufacturer's manual. Sequence amplification and fluorescence detection were performed on an Applied Biosystems ViiA 7 Real-Time PCR System. Relative quantification was performed by the comparative threshold cycle (ΔΔCT) method.

Statistical analysis. Each result is presented as an arithmetic mean ± SE for a sample size (i. e., number of mice) indicated. Three-way ANOVA with factors "sex" (male, female), "experimental group" (PBS, FGF21 administration), and "day of experiment" (1–6) was used to analyze FGF21 effects on food intake and body weight; with factors "sex", "experimental group", and "eating" (daily FI before and after fasting), to analyze FGF21 effect on FI after fasting. Two-way ANOVA with factors "sex" and "experimental group" was used to analyze FGF21 effects on blood parameters and gene expression with multiple comparisons using the post hoc Duncan test. Significance was determined as $p \leq 0.05$. The STATISTICA 6 software package (StatSoft) was used for analysis.

Results

Biochemical characteristics of blood and plasma. No differences between male and female A^y-mice were observed in concentrations of blood glucose, either before or during the experiment. Only the plasma level of adiponectin was higher in females than in males. FGF21 administration significantly decreased plasma insulin specifically in male mice. Both male and female mice tended to respond to FGF21 administration by decreased plasma levels of free fatty acids (FFA) (Table 2).

Body weight (BW), weights of fat and liver, and food intake (FI). FGF21 administration exerted differential effects on BW in male and female mice ($p < 0.001$, $F_{1.118} \pm 12.2$, "sex" × "experimental group", three-way ANOVA, Fig. 1). In males, FGF21 contributed to weight loss, and significant differences in BW between male mice treated with PBS and FGF21 were observed from day 5 of the experiment (see Fig. 1). In females, FGF21 administration did not affect BW.

The weights of both subcutaneous and abdominal WAT were higher in female mice, whereas no such differences could be detected in the case of BAT, FGF21 administration did not affect fat weights in either males or females (Fig. 2). Liver weight was lower in females, and FGF21 administration increased this parameter in females, without affecting it in males (see Fig. 2).

The effect of FGF21 on FI was also sex-dependent ($p < 0.0001$, $F_{1.113} \pm 20.7$, "sex" × "experimental group", three-way ANOVA, Fig. 3, *a*). FGF21 administration did not affect FI in males, but significantly increased it in females from the first day of the experiment (see Fig. 3, *a*).

We also assessed the effect of FGF21 on FI during the 24-h period of refeeding after overnight fasting. Fasting taken alone stimulated FI, and FGF21 administration further increased it during refeeding in both males and females ($p < 0.001$, $F_{1.39} \pm 14.3$, "experimental group", three-way ANOVA, see Fig. 3, *b*).

Gene expression. In mice that received PBS, liver expression of most of the genes studied was sex-dependent (Fig. 4).

Higher levels of mRNA of genes involved in beta-oxidation (*Ppara*, *Pgc1*, *Cpt1*), glucose metabolism (*Insr*, *Slc2a2*), glycolysis (*G6pc*, *Pklr*), and gluconeogenesis (*Gck*, *Pck1*), as well as lipolysis (*Pnpla2*) and lipogenesis (*Lpl*, *Acaca*, *Acacb*) were observed in females, as compared to males. FGF21 administration inhibited liver expression of the genes, and this inhibition was more pronounced in females than in males. Thus, FGF21 administration effectively eliminated sex differences in liver expression of the genes, observed in the control (see Fig. 4). Liver expression of *Fgf21*, exhibiting no differences between males and females, was reduced in response to exogenous FGF21. Sex differences in gene expression were also observed in abdominal fat. In mice treated with PBS, mRNA levels of genes encoding PPARg, LPL, and GLUT4 were higher in females than in males (Fig. 5). FGF21 administration reduced the expression of these genes in females to a greater extent than in males, thereby eliminating the sex-dependent differences observed in the control (see Fig. 5).

Table 1. Gene expression assays used for relative quantitation real-time PCR

Protein	Gene	Gene expression assay
Acetyl-coenzyme A carboxylase alpha	*Acaca*	Mm01304285_m1
Acetyl-coenzyme A carboxylase beta	*Acacb*	Mm01204683_m1
Agouti related neuropeptide	*Agrp*	Mm00475829_g1
Carnitine palmitoyltransferase 1a	*Cpt1a*	Mm01231183_m1
Carnitine palmitoyltransferase 1b	*Cpt1b*	Mm00487191_g1
Deiodinase, iodothyronine, type II	*Dio2*	Mm00515664_m1
Fatty acid synthase	*Fasn*	Mm00662319_m1
Fibroblast growth factor 21	*Fgf21*	Mm00840165_g1
Glucose-6-phosphatase, catalytic	*G6pc*	Mm00839363_m1
Glucokinase	*Gck*	Mm00439129_m1
Insulin receptor	*Insr*	Mm01211875_m1
Klotho beta	*Klb*	Mm00473122_m1
Leptin receptor	*Lepr*	Mm00440181_m1
Lipase, hormone sensitive	*Lipe*	Mm00495359_m1
Lipoprotein lipase	*Lpl*	Mm00434764_m1
Neuropeptide Y	*Npy*	Mm01410146_m1
Patatin-like phospholipase domain containing 2 (adipocyte triglyceride lipase (ATGL))	*Pnpla2*	Mm00503040_m1
Peroxisome proliferative activated receptor, gamma, coactivator 1 alpha	*Ppargc1a*	Mm01208835_m1
Peroxisome proliferator activated receptor alpha	*Ppara*	Mm0040939_m1
Peroxisome proliferator activated receptor gamma	*Pparg*	Mm00440940_m1
Phosphoenolpyruvate carboxykinase 1, cytosolic	*Pck1*	Mm01247058_m1
Pro-opiomelanocortin	*Pomc*	Mm00435874_m1
Pyruvate kinase liver and red blood cell	*Pklr*	Mm00443090_m1
Solute carrier family 2 (facilitated glucose transporter), member 1 (GLUT1)	*Slc2a1*	Mm00441480_m1
Solute carrier family 2 (facilitated glucose transporter), member 2 (GLUT2)	*Slc2a2*	Mm00446229_m1
Solute carrier family 2 (facilitated glucose transporter), member 4 (GLUT4)	*Slc2a4*	Mm00436615_m1
Uncoupling protein 1 (mitochondrial, proton carrier)	*Ucp1*	Mm01244861_m1
Uncoupling protein 3	*Ucp3*	Mm01163394_m1
Beta-actin	*Actb*	Mm00607939_s1

Table 2. Effects of FGF21 administration on biochemical parameters of blood

Biochemical parameter	Males		Females		*p* (ANOVA)
	PBS *n* = 7	FGF21 *n* = 6	PBS *n* = 6	FGF21 *n* = 5	
Glucose fasted initial, mM	8.17 ± 0.51	8.08 ± 0.75	7.90 ± 1.28	8.30 ± 1.02	NS
Glucose fasted day 7, mM	7.73 ± 1.51	6.68 ± 0.70	5.73 ± 1.11	8.30 ± 0.93	NS
Glucose, mM	9.64 ± 0.71	8.64 ± 0.30	9.27 ± 0.50	10.48 ± 0.49	NS
FFA, mM	0.19 ± 0.09	0.05 ± 0.04	0.27 ± 0.13	0.05 ± 0.05	0.07, E
Triglycerides, mM	1.04 ± 0.10	0.99 ± 0.07	0.99 ± 0.07	0.79 ± 0.11	NS
Cholesterol, mM	3.06 ± 0.12	3.25 ± 0.41	2.70 ± 0.22	2.83 ± 0.17	NS
Insulin, ng/ml	12.06 ± 0.89	7.56 ± 1.29*	8.60 ± 1.48	7.42 ± 1.45	< 0.05, E
Adiponectin, μg/ml	3.90 ± 0.37	4.12 ± 0.46	5.44 ± 0.54	5.67 ± 0.67	< 0.01, S
Leptin, ng/ml	21.75 ± 3.13	22.28 ± 6.86	21.15 ± 5.85	23.08 ± 4.72	NS

Male and female A^y-mice were administered with FGF21 (1 mg per 1 kg) or PBS for 10 days, and blood samples to measure plasma parameters were collected 1 h after the last injection. Fasted glucose was measured in blood before (initial) and on day 7 of the experiment using a glucometer. *$p < 0.05$, males, FGF21 vs. PBS, post hoc Duncan test. NS – non-significant; S – "sex"; E – "experimental group".

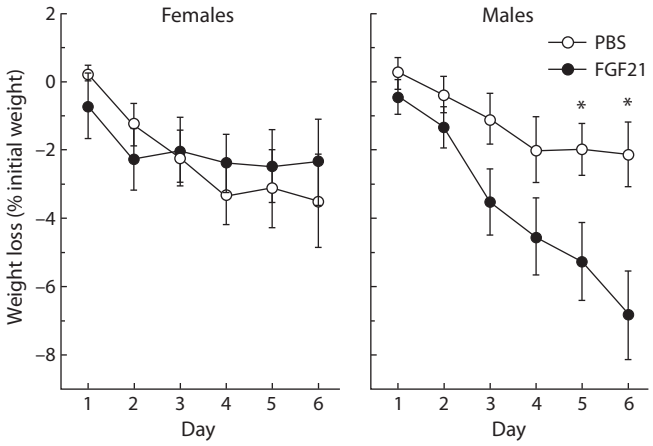


Fig. 1. Effect of FGF21 administration on BW in male and female A^y-mice.

Weight loss was calculated as the difference between the weight on the day of injection and the initial weight, related to the initial weight (%).
*$p < 0.05$, post hoc Duncan test.

For the majority of the genes studied, the inhibitory effects of FGF21 administration on the expression were more pronounced in females. This was confirmed for genes involved in lipogenesis (*Pparg*, *Lpl*, *Acacb*), lipolysis (*Pnpla2*), and transcription coactivation (*Ppargc1a*). In a single case of *Acaca* gene, the pattern was reversed: FGF21 administration decreased the level of mRNA in males while increasing it in females. The expression of *Insr* and *Lipe* was higher in females, and FGF21 administration did not affect their mRNA levels (see Fig. 5).

In subcutaneous fat of the controls, there were no sex-dependent expression differences for any of the genes studied. FGF21 administration affected only *Cpt1b*, decreasing its expression both in males and females.

In brown fat of the controls, there were no sex-dependent expression differences for any of the genes studied, and FGF21 administration did not affect the expression.

In muscle tissue of the controls, there were no sex-dependent expression differences for any of the genes studied.

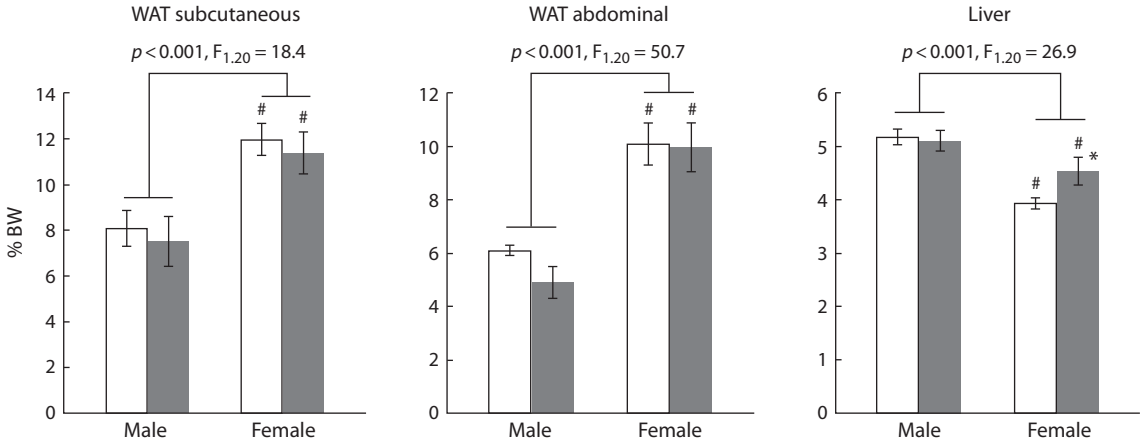


Fig. 2. FGF21 influence on weights of fat tissues and liver in male and female A^y-mice. Organ weights were calculated as percentages of BW. Sex dependence (two-way ANOVA) is indicated in the graph.

#$p < 0.05$ for males vs. females; *$p < 0.05$ for FGF21 vs. PBS in females, post hoc Duncan test.

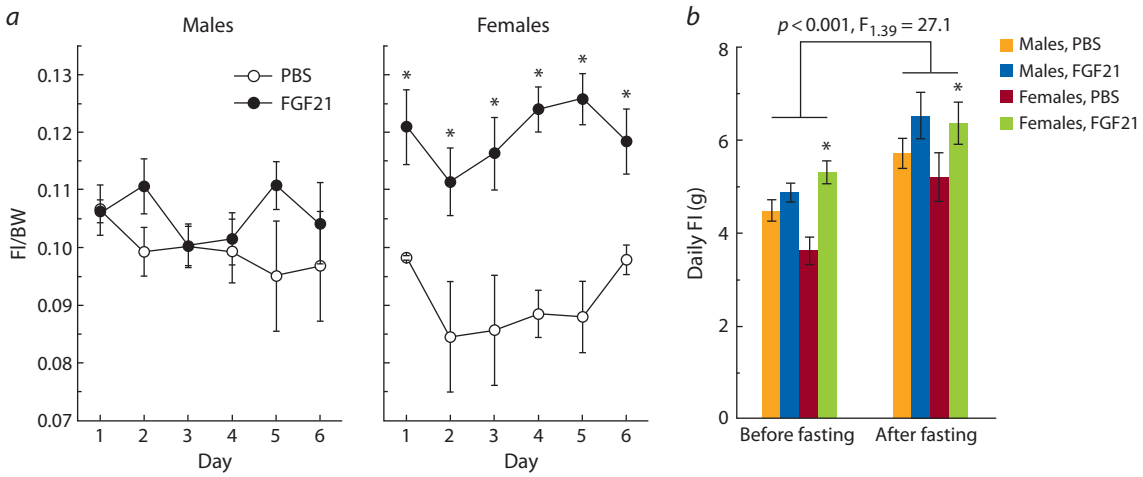


Fig. 3. FGF21 influence on FI *ad libitum* (*a*) and after fasting (*b*) in male and female A^y-mice.

a – FI/BW ratio; *b* – mice were fasted overnight after seven injections of FGF21 or PBS. FI "before fasting" was calculated as an arithmetic mean of six daily FIs measured for 6 days prior to fasting. FI "after fasting" represents the amount of food consumed during 24 h of refeeding. The effect of fasting on FI (three-way ANOVA) is indicated in the graph.
*$p \le 0.05$ for FGF21 vs. PBS in females, post hoc Duncan test.

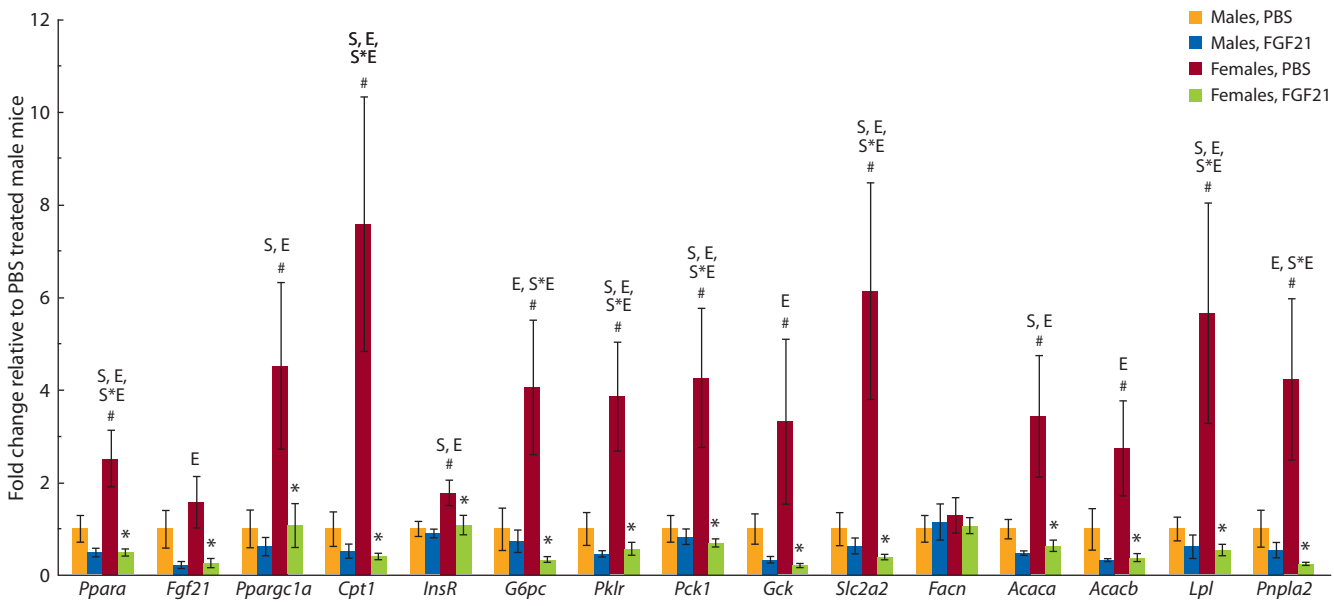


Fig. 4. FGF21 influence on gene expression in the liver of male and female A^y-mice.

Effects of factors "S" (sex) or "E" (experimental group), or interactions thereof ("S*E") at the level of significance $p < 0.05$ (two-way ANOVA) are indicated in the graph; $^{\#}p < 0.05$ for males vs. females; $^{*}p < 0.05$ for FGF21 vs. PBS in females; post hoc Duncan test.

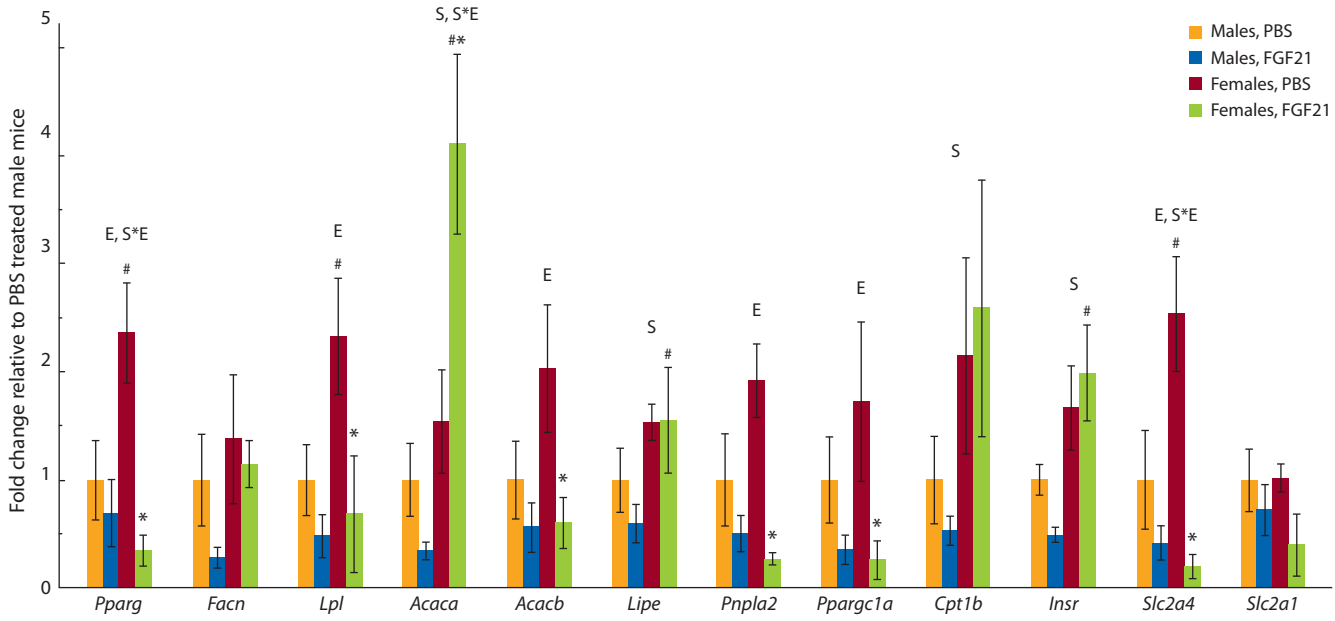


Fig. 5. FGF21 influence on gene expression in the abdominal WAT of male and female A^y-mice.

Effects of factors "S" (sex) or "E" (experimental group), or interactions thereof ("S*E") at the level of significance $p < 0.05$ (two-way ANOVA) are indicated in the graph. $^{\#}p < 0.05$ for males vs. females; $^{*}p < 0.05$ for FGF21 vs. PBS in females, post hoc Duncan test.

FGF21 did not affect the expression significantly, although the level of *Cpt1b* mRNA tended to increase after FGF21 administration ($p \pm 0.06$, two-way ANOVA).

Mice treated with PBS exhibited no differences between males and females in the hypothalamic mRNA levels of the genes studied. After FGF21 administration, the expression of *Klb* was decreased in mice of both sexes ($p < 0.05$, $F_{1.20} \pm 6.2$), whereas that of *Pomc* was decreased in males only ($p < 0.05$, Student's *t*-test, 1.0 ± 0.3, $n \pm 6$, control males vs. 0.29 ± 0.10, $n \pm 6$, FGF21 treated males).

Discussion

In the present study, we assessed the pharmacological effects of FGF21 in male and female mice with A^y mutation, which evokes MC4R blockage. We found that these effects were strongly dependent on the sex of the animals.

FGF21 administered to animals with diabetes and obesity decreases their BW and blood glucose/insulin levels, improves blood lipid profile, and increases insulin sensitivity (Bon-Durant, Potthoff, 2018). In our study, responses to FGF21 administration in male A^y-mice were largely consistent with

its expected impact: their BWs and blood insulin decreased, blood levels of FFA tended to decrease, and blood glucose and adiponectin were not affected. Our observation that FGF21 decreased blood insulin without affecting glucose, which retained normal levels, suggests that insulin sensitivity in male A^y-mice was possibly improved. Metabolic effect of FGF21 administration was not associated with changes in blood adiponectin, since adiponectin is not required for the chronic effects of FGF21 to reduce body weight or its effects on glucose homeostasis (BonDurant et al., 2017).

Although we did not monitor energy expenditure, it was most likely increased, since weight loss in male mice took place in the absence of changes in FI. The ability of FGF21 to induce weight loss without affecting FI was reported previously for male mice with diet-induced and genetic obesity (*db*/*db*, *ob*/*ob*) (Kharitonenkov et al., 2005; Xu et al., 2009; Camporez et al., 2013). This effect was arguably due to increased energy expenditure caused by intense locomotor activity (Xu et al., 2009) and elevated metabolic rate (Coskun et al., 2008; Xu et al., 2009). FGF21-induced growth of metabolic rates is associated with augmentation of fatty acid oxidation in liver and adipose tissue (BAT and WAT), caused in turn by increased expression of genes encoding CPT-1 (in liver only), PGC-1 (in liver and adipose tissue), and UCP-1 (in liver and adipose tissue) (Camporez et al., 2013). No such changes in gene expression have been observed in this study; on the contrary, FGF21 administration to males reduced the expression of *Ppargc1a* and *Cpt1* in liver and *Ppargc1a* in abdominal WAT ($p < 0.05$, two-way ANOVA).

Our results are consistent with those reported by (Coskun et al., 2008), demonstrating that weight loss and improved glucose metabolism in FGF21-treated C57Bl males were accompanied by a decrease in liver expression of *Cpt1* (with no changes of *Ppargc1a* expression). However, T. Coskun et al. (2008) found increased expression of genes related to (1) thermogenesis (*Ucp1*) in WAT and BAT, (2) lipolysis (*Lipe*, *Pnpla2*) in WAT, and (3) lipogenesis (*Acaca*, *Acacb*) in WAT, suggesting that energy expenditure is achieved via thermogenesis and induction of a state of increased futile cycling. We have not observed induction of *Ucp1* expression either in BAT or WAT; likewise, there have been no increase in the expression of lipolysis/lipogenesis genes in liver or adipose tissue. Thus, our experimental data provide no indication of futile cycling activation. FGF21-induced activation of energy expenditure involves, in addition to central mechanisms, the sympathetic nervous system (Owen et al., 2014; Lan et al., 2017). Of note, the latter also mediates regulation of energy expenditure by the melanocortin system (Rossi et al., 2011; Berglund et al., 2014). Therefore, FGF21 may act, at least in certain cases, via melanocortin signaling pathways. If so, the blockade of MC4R in A^y-mice can interfere with the activating effect of FGF21 on the expression of the genes we studied. Weight loss in male A^y-mice is indicative of yet other mechanisms whereby energy expenditure may be increased, which remain to be explored.

In females with A^y mutation, the effect of FGF21 differed dramatically from that observed in males. In females unlike males, FGF21 administration did not induce weight loss, but increased liver weight. A^y-mice are hyperphagic (Wolff et al., 1999), and FGF21 exacerbated hyperphagia specifically in females. It is not known whether this difference is due to A^y mutation or, rather, we are dealing with a general sex-linked discrepancy in responses of female and male mice to FGF21 treatment (the pharmacological effects of FGF21 have thus far not been studied in female mice).

Resistance of female A^y-mice to catabolic effects of FGF21 can be explained by increase of food intake. FGF21 influence on FI in female rodent was not studied, but in males, FGF21 was shown previously to increase FI in rats (Recinella et al., 2017), mice with diet-induced obesity (Coskun et al., 2008), and to increase protein intake while reducing carbohydrate intake in normal mice, and the later effect was mediated via CNS (Larson et al., 2019). However, FGF21-dependent signaling pathways that regulate eating behavior have not been identified. The melanocortin system is involved in the regulation of the response to protein deficiency, and FGF21 may act as the sensor triggering that response. In our experiment, male mice demonstrated no increase in FI, although FGF21 administration decreased hypothalamic expression of *Pomc*. POMC is a precursor of the anorexigenic neuropeptide MSH, and a decrease in *Pomc* expression (limiting MC4R activation by MSH) is associated with increased FI. A^y mutation leads to MC4R blockade, which disrupts MSH regulation of FI and may be the reason why males did not respond to FGF21 by increased FI.

In females, FGF21 administration caused a considerable increase in FI without affecting the expression of the melanocortin system genes. In A^y-females, FGF21 influence on FI may be mediated via estradiol-sensitive mechanism that is not altered by impaired melanocortin signaling (Morton et al., 2004). It is possible that the orexigenic effects of FGF21 in males and females involve distinct neuronal signaling pathways.

According to the data obtained in the control groups, a different metabolic response to exogenous FGF21 in males and females was induced at the background of significant sex differences in the mass of the liver and adipose tissue and in the expression of the liver and fat genes. Sex differences in gene expression in adipose tissue differed depending on localization; in our experiment, we found them only in abdominal adipose tissue. These data are consistent with the results obtained when assessing gene expression in adipose tissue in mice: the number of differentially expressed genes in males and females was significantly higher in abdominal than in subcutaneous fat (Grove et al., 2010). Increased weight of abdominal fat in control A^y-females, compared to males, was associated with increased expression of genes upregulating lipogenesis in adipose tissue: *Ppar*, *Lpl*, *Insr* and *Slc2a4*. FGF21 administration eliminated differences in the expression of these genes, however, it did not lead to significant changes in the mass of adipose tissue. It is possible that more prolonged FGF21 administration is required to detect changes at the level of adipose tissue weight.

The most striking sex differences in gene expression in the control groups were observed in liver. In controls, the expression of most genes studied (except for *Fgf21* and *Fasn*) was higher in females than in males. These results are consistent with transcriptome analysis data demonstrating differential expression rates for 72 % active liver genes in male and female mice (Yang et al., 2006). Sex differences in expression of genes

related to lipid metabolism were found in the liver of mice and rats on high-calorie diet: compared to males, females had enhanced mRNA levels of *Acc1*, *Pparg*, *Cpt1* and other genes in mice (Gasparin et al., 2018), and enhanced protein levels of ACC, PPARs and FAS in rats (Chukijrungroat et al., 2017). As shown by S. Della Torre et al. (2017), liver is the major target for estrogens, and estrogen receptor alpha (ERα) has a direct effect on the regulation of the hepatic genes relevant for energy metabolism.

Increased expression of genes involved in fatty acid β-oxidation (*Ppara*, *Ppargc1a*, *Cpt1*), glycolysis (*G6p*, *Pklr*), gluconeogenesis (*Pck1*, *Gck*), glucose transport *(Slc2a2*), lipogenesis and lipolysis (*Acaca*, *Acacb*, *Lpl*, *Pnpla2*) possibly indicates that the rate of glucose and fat metabolism is higher in the liver of A^y-females than A^y-males. FGF21 administration reduced liver expression of all the genes mentioned, thereby eliminating entirely any sex-dependent differences, and this effect was associated with the increase of female liver weight. Enlarged liver under the action of FGF21 in female A^y-mice may indicate the development of steatosis. Simultaneous decrease in catabolic (lipolysis, glycolysis) and anabolic (lipogenesis, gluconeogenesis) processes can suggest decreased metabolism in the liver of A^y-females. Decreased liver metabolism in females could reduce liver energy expenditure and contribute, together with increased food intake, to liver weight gain. Additional morphological and biochemical studies are needed to find out whether FGF21 initiates (or promotes) steatosis in A^y-female mice.

Taken together, our findings indicate that therapeutic effects of FGF21 in mice with disrupted melanocortin signaling are strongly sex-dependent. In male A^y-mice, the blockage of MC4Rs does not prevent the anti-obesity effect of FGF21: its administration results in weight loss and blood insulin decrease. However, obese A^y-females exerted resistance to catabolic and antidiabetic effects of FGF21. In females, exogenous FGF21 stimulates FI without reducing BW and blood glucose/insulin, inhibits liver expression of genes related to glucose and lipid turnover, and increases liver weight. The contribution of estrogens and the melanocortin system to those effects remains to be elucidated.

Conclusion

The pharmacological effect of FGF21 may depend on the animal sex and etiology of obesity. Although an immediate translation to humans of findings obtained in experiments with mice is not possible, our results suggest that detailed preclinical studies of the pharmacological effects of FGF21 are required, taking into account the sex of individuals under study and the genesis of obesity.

References

Bazhan N., Jakovleva T., Balyibina N., Dubinina A., Denisova E., Feofanova N., Makarova E. Sex dimorphism in the Fgf21 gene expression in liver and adipose tissues is dependent on the metabolic condition. Online J. Biol. Sci. 2019;19(1):28-36. DOI 10.3844/ojbsci.2019.28.36.

Berglund E.D., Liu T., Kong X., Sohn J.-W., Vong L., Deng Z., Lee C.E., Lee S., Williams K.W., Olson D.P., Scherer P.E., Lowell B.B., Elmquist J.K. Melanocortin 4 receptors in autonomic neurons regulate thermogenesis and glycemia. Nat. Neurosci. 2014;17(7):911-913. DOI 10.1038/nn.3737.

BonDurant L.D., Ameka M., Naber M.C., Markan K.R., Idiga S.O., Acevedo M.R., Walsh S.A., Ornitz D.M., Potthoff M.J. FGF21 regulates metabolism through adipose-dependent and -independent mechanisms. Cell Metab. 2017;25(4):935-944.e4. DOI 10.1016/j.cmet.2017.03.005.

BonDurant L.D., Potthoff M.J. Fibroblast growth factor 21: a versatile regulator of metabolic homeostasis. Annu. Rev. Nutr. 2018;38(1):173-196. DOI 10.1146/annurev-nutr-071816-064800.

Bultman S.J., Michaud E.J., Woychik R.P. Molecular characterization of the mouse agouti locus. Cell. 1992;71(7):1195-1204.

Camporez J.P.G., Jornayvaz F.R., Petersen M.C., Pesta D., Guigni B.A., Serr J., Zhang D., Kahn M., Samuel V.T., Jurczak M.J., Shulman G.I. Cellular mechanisms by which FGF21 improves insulin sensitivity in male mice. Endocrinology. 2013;154(9):3099-3109. DOI 10.1210/en.2013-1191.

Chukijrungroat N., Khamphaya T., Weerachayaphorn J., Songserm T., Saengsirisuwan V. Hepatic FGF21 mediates sex differences in high-fat high-fructose diet-induced fatty liver. Am. J. Physiol. Endocrinol. Metab. 2017;313(2):E203-E212. DOI 10.1152/ajpendo.00076.2017.

Clayton J.A. Applying the new SABV (sex as a biological variable) policy to research and clinical care. Physiol. Behav. 2018;1872-1875. DOI 10.1016/j.physbeh.2017.08.012.

Coskun T., Bina H.A., Schneider M.A., Dunbar J.D., Hu C.C., Chen Y., Moller D.E., Kharitonenkov A. Fibroblast growth factor 21 corrects obesity in mice. Endocrinology. 2008;149(12):6018-6027. DOI 10.1210/en.2008-0816.

Della Torre S., Lolli F., Ciana P., Maggi A. Sexual dimorphism and estrogen action in mouse liver. Adv. Exp. Med. Biol. 2017;1043:141-151.

Fani L., Bak S., Delhanty P., van Rossum E.F.C., van den Akker E.L.T. The melanocortin-4 receptor as target for obesity treatment: a systematic review of emerging pharmacological therapeutic options. Int. J. Obes. (Lond.) 2014;38(2):163-169. DOI 10.1038/ijo.2013.80.

Farooqi I.S., Keogh J.M., Yeo G.S.H., Lank E.J., Cheetham T., O'Rahilly S. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. N. Engl. J. Med. 2003;348(12):1085-1095. DOI 10.1056/NEJMoa022050.

Gasparin F.R.S., Carreño F.O., Mewes J.M., Gilglioni E.H., Pagadigorria C.L.S., Natali M.R.M., Utsunomiya K.S., Constantin R.P., Ouchida A.T., Curti C., Gaemers I.C., Elferink R.P.J.O., Constantin J., Ishii-Iwamoto E.L. Sex differences in the development of hepatic steatosis in cafeteria diet-induced obesity in young mice. Biochim. Biophys. Acta Mol. Basis Dis. 2018;1864(7):2495-2509. DOI 10.1016/j.bbadis.2018.04.004.

Girardet C., Butler A.A. Neural melanocortin receptors in obesity and related metabolic disorders. Biochim. Biophys. Acta. 2014;1842(3):482-494. DOI 10.1016/j.bbadis.2013.05.004.

Grove K.L., Fried S.K., Greenberg A.S., Xiao X.Q., Clegg D.J. A microarray analysis of sexual dimorphism of adipose tissues in high-fat-diet-induced obese mice. Int. J. Obes. (Lond.). 2010;34(6):989-1000. DOI 10.1038/ijo.2010.12.

Jackson V.M., Breen D.M., Fortin J.-P., Liou A., Kuzmiski J.B., Loomis A.K., Rives M.-L., Shah B., Carpino P.A. Latest approaches for the treatment of obesity. Expert Opin. Drug Discov. 2015;10(8):825-839. DOI 10.1517/17460441.2015.1044966.

Kharitonenkov A., Shiyanova T.L., Koester A., Ford A.M., Micanovic R., Galbreath E.J., Sandusky G.E., Hammond L.J., Moyers J.S., Owens R.A., Gromada J., Brozinick J.T., Hawkins E.D., Wroblewski V.J., Li D.-S., Mehrbod F., Jaskunas S.R., Shanafelt A.B. FGF-21 as a novel metabolic regulator. J. Clin. Invest. 2005;115(6):1627-1835. DOI 10.1172/JCI23606.

Lan T., Morgan D.A., Rahmouni K., Sonoda J., Fu X., Burgess S.C., Holland W.L., Kliewer S.A., Mangelsdorf D.J. FGF19, FGF21, and an FGFR1/β-Klotho-activating antibody act on the nervous system to regulate body weight and glycemia. Cell Metab. 2017;26(5):709-718.e3. DOI 10.1016/j.cmet.2017.09.005.

Larson K.R., Chaffin A.T.-B., Goodson M.L., Fang Y., Ryan K.K. Fibroblast growth factor-21 controls dietary protein intake in male

mice. Endocrinology. 2019;160(5):1069-1080. DOI 10.1210/en.2018-01056.

Lee Y.-H., Kim S.H., Kim S.-N., Kwon H.-J., Kim J.-D., Oh J.Y., Jung Y.-S. Sex-specific metabolic interactions between liver and adipose tissue in MCD diet-induced non-alcoholic fatty liver disease. Oncotarget. 2016;7(30):46959-46971. DOI 10.18632/oncotarget.10506.

Markan K.R., Naber M.C., Ameka M.K., Anderegg M.D., Mangelsdorf D.J., Kliewer S.A., Mohammadi M., Potthoff M.J. Circulating FGF21 is liver derived and enhances glucose uptake during refeeding and overfeeding. Diabetes. 2014;63(12):4057-4063. DOI 10.2337/db14-0595.

Mauvais-Jarvis F., Arnold A.P., Reue K. A guide for the design of pre-clinical studies on sex differences in metabolism. Cell Metab. 2017;25(6):1216-1230. DOI 10.1016/j.cmet.2017.04.033.

Michaud E.J., Mynatt R.L., Miltenberger R.J., Klebig M.L., Wilkinson J.E., Zemel M.B., Wilkison W.O., Woychik R.P. Role of the agouti gene in obesity. J. Endocrinol. 1997;155(2):207-209.

Morton G.J., Mystkowski P., Matsumoto A.M., Schwartz M.W. Increased hypothalamic melanin concentrating hormone gene expression during energy restriction involves a melanocortin-independent, estrogen-sensitive mechanism. Peptides. 2004;25(4):667-674. DOI 10.1016/j.peptides.2004.02.007.

Owen B.M., Ding X., Morgan D.A., Coate K.C., Bookout A.L., Rahmouni K., Kliewer S.A., Mangelsdorf D.J. FGF21 acts centrally to induce sympathetic nerve activity, energy expenditure, and weight loss. Cell Metab. 2014;20(4):670-677. DOI 10.1016/j.cmet.2014.07.012.

Recinella L., Leone S., Ferrante C., Chiavaroli A., Di Nisio C., Martinotti S., Vacca M., Brunetti L., Orlando G. Effects of central fibroblast growth factor 21 (FGF21) in energy balance. J. Biol. Regul. Homeost. Agents. 2017;31(3):603-613.

Rossi J., Balthasar N., Olson D., Scott M., Berglund E., Lee C.E., Choi M.J., Lauzon D., Lowell B.B., Elmquist J.K. Melanocortin-4 receptors expressed by cholinergic neurons regulate energy balance and glucose homeostasis. Cell Metab. 2011;13(2):195-204. DOI 10.1016/j.cmet.2011.01.010.

Singh R.K., Kumar P., Mahalingam K. Molecular genetics of human obesity: A comprehensive review. C. R. Biol. 2017;340(2):87-108. DOI 10.1016/j.crvi.2016.11.007.

Tao Y.-X. The melanocortin-4 receptor: physiology, pharmacology, and pathophysiology. Endocr. Rev. 2010;31(4):506-543. DOI 10.1210/er.2009-0037.

Wolff G.L., Roberts D.W., Mountjoy K.G. Physiological consequences of ectopic agouti gene expression: the yellow obese mouse syndrome. Physiol. Genomics. 1999;1(3):151-163. DOI 10.1152/physiolgenomics.1999.1.3.151.

Xie T., Leung P.S. Fibroblast growth factor 21: a regulator of metabolic disease and health span. Am. J. Physiol. Endocrinol. Metab. 2017;313(3):E292-E302. DOI 10.1152/ajpendo.00101.2017.

Xu J., Lloyd D.J., Hale C., Stanislaus S., Chen M., Sivits G., Vonderfecht S., Hecht R., Li Y.-S., Lindberg R.A., Chen J.-L., Jung D.Y., Zhang Z., Ko H.-J., Kim J.K., Véniant M.M. Fibroblast growth factor 21 reverses hepatic steatosis, increases energy expenditure, and improves insulin sensitivity in diet-induced obese mice. Diabetes. 2009;58(1):250-259. DOI 10.2337/db08-0392.

Yang X., Schadt E.E., Wang S., Wang H., Arnold A.P., Ingram-Drake L., Drake T.A., Lusis A.J. Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res. 2006;16(8):995-1004. DOI 10.1101/gr.5217506.

ORCID ID

E.N. Makarova orcid.org/0000-0002-6417-9893
T.V. Yakovleva orcid.org/0000-0001-7628-5856
N.Yu. Balyibina orcid.org/0000-0001-7168-1438
K.O. Baranov orcid.org/0000-0002-4029-6518
E.I. Denisova orcid.org/0000-0001-8696-8781
A.D. Dubinina orcid.org/0000-0003-3907-3488
N.A. Feofanova orcid.org/0000-0002-2591-1623
N.M. Bazhan orcid.org/0000-0002-7246-4758

Acknowledgements. This research was funded by the Russian Science Foundation, Grant No. 17-15-01036.

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

Received October 7, 2019. Revised November 28, 2019. Accepted November 29, 2019. Published online February 3, 2020.

Вавиловский журнал генетики и селекции. 2020;24(2):209-218
DOI 10.18699/VJ20.613

Involvement of transposable elements in neurogenesis

R.N. Mustafin[1] ✉, E.K. Khusnutdinova[2]

[1] Bashkir State Medical University, Ufa, Russia
[2] Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia
✉ e-mail: ruji79@mail.ru

Abstract. The article is about the role of transposons in the regulation of functioning of neuronal stem cells and mature neurons of the human brain. Starting from the first division of the zygote, embryonic development is governed by regular activations of transposable elements, which are necessary for the sequential regulation of the expression of genes specific for each cell type. These processes include differentiation of neuronal stem cells, which requires the finest tuning of expression of neuron genes in various regions of the brain. Therefore, in the hippocampus, the center of human neurogenesis, the highest transposon activity has been identified, which causes somatic mosaicism of cells during the formation of specific brain structures. Similar data were obtained in studies on experimental animals. Mobile genetic elements are the most important sources of long non-coding RNAs that are coexpressed with important brain protein-coding genes. Significant activity of long non-coding RNA was detected in the hippocampus, which confirms the role of transposons in the regulation of brain function. MicroRNAs, many of which arise from transposon transcripts, also play an important role in regulating the differentiation of neuronal stem cells. Therefore, transposons, through their own processed transcripts, take an active part in the epigenetic regulation of differentiation of neurons. The global regulatory role of transposons in the human brain is due to the emergence of protein-coding genes in evolution by their exonization, duplication and domestication. These genes are involved in an epigenetic regulatory network with the participation of transposons, since they contain nucleotide sequences complementary to miRNA and long non-coding RNA formed from transposons. In the memory formation, the role of the exchange of virus-like mRNA with the help of the Arc protein of endogenous retroviruses HERV between neurons has been revealed. A possible mechanism for the implementation of this mechanism may be reverse transcription of mRNA and site-specific insertion into the genome with a regulatory effect on the genes involved in the memory.
Key words: brain; differentiation; noncoding RNA; retroelements; neuronal stem cells; transposable elements.

For citation: Mustafin R.N., Khusnutdinova E.K. Involvement of transposable elements in neurogenesis. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):209-218. DOI 10.18699/VJ20.613

Участие мобильных элементов в нейрогенезе

Р.Н. Мустафин[1] ✉, Э.К. Хуснутдинова[2]

[1] Башкирский государственный медицинский университет, Уфа, Россия
[2] Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
✉ e-mail: ruji79@mail.ru

Аннотация. В обзоре представлены накопленные в научной литературе данные об участии мобильных генетических элементов в регуляции дифференцировки нейрональных стволовых клеток и функционирования зрелых нейронов головного мозга. Начиная с первого деления зиготы, эмбриональное развитие управляется закономерными активациями транспозонов, необходимыми для последовательного изменения экспрессии специфических для каждого типа клеток генов. Частным отражением этих процессов может быть дифференцировка нейрональных стволовых клеток – процесс, в ходе которого необходима наиболее тонкая настройка экспрессии генов в нейронах различных областей головного мозга. Доказательствами этого предположения являются данные о высокой активности транспозонов в центре нейрогенеза, зубчатой извилине гиппокампа. Кроме того, мобильные элементы – источники возникновения и эволюции длинных некодирующих РНК, которые коэкспрессируются с необходимыми для работы головного мозга белок-кодирующими генами. Наибольшая активность длинных некодирующих РНК, так же как и транспозонов, обнаружена в центре нейрогенеза человека, что позволяет предположить их участие в управлении работой головного мозга. В регуляции дифференцировкой нейрональных стволовых клеток используются также микроРНК, многие из которых возникают из транскриптов мобильных элементов. Транспозоны посредством собственных процессированных транскриптов играют роль в эпигенетической регуляции дифференцировки нейронов. Объяснением глобальной регуляторной функции мобильных элементов в головном мозге человека может служить их значение в возникновении белок-кодирующих генов в эволюции путем экзонизации, дупликации и доместикации. Эти гены вовлечены в эпигенетическую регу-

ляторную сеть с участием транспозонов, так как содержат нуклеотидные последовательности, комплементарные микроРНК и длинным некодирующим РНК, образуемым из транскриптов мобильных элементов. В формировании памяти выявлена роль обмена вирусоподобными частицами мРНК при помощи белка Arc эндогенных ретровирусов HERV между нейронами. Возможными способами реализации этого механизма могут быть обратная транскрипция мРНК и сайт-специфическая интеграция в геном с регуляторным воздействием на гены, участвующие в консолидации информации.
Ключевые слова: головной мозг; дифференцировка; некодирующие РНК; ретроэлементы; стволовые нервные клетки; транспозоны.

Introduction

Transposable elements (TE) make up 69 % of the human genome (de Koning et al., 2011). In the course of evolution, many protein-coding genes (Joly-Lopez, Bureau, 2018), regulatory nucleotide sequences (Ito et al., 2017; Schrader, Schmitz, 2018), and telomeres (Kopera et al., 2011), non-coding RNAs (ncRNAs), including microRNAs (Piriyapongsa et al., 2007; Yuan et al., 2010, 2011; Qin et al., 2015) and long human ncRNAs (Johnson, Guigo, 2014) originating from TE. Over millions of years of evolution, cells have developed various defense systems against TE insertion into their genomes, including DNA methylation, heterochromatin formation, and RNA interference (RNAi). These epigenetic mechanisms have made a significant contribution to the regulation of specific gene expression and cell differentiation (Habibi et al., 2015).

Transposable elements are divided into two main classes, in accordance with the mechanisms of their transposition. DNA-TEs are transposed by "cut and paste" or "rolling circle". Retroelements (REs) are integrated into new genome sites using "copy and paste". REs are classified into those containing long terminal repeats (LTR REs) (Fig. 1) and those not containing them (non-LTR REs) (Fig. 2). The latter are divided into autonomous (LINE, long interspersed nuclear elements) and non-autonomous (SINE, short interspersed nuclear elements) and SVA (SINE-VNTR-Alu) (Fig. 3) (Klein, O'Neill, 2018).

The human genome contains more than 500,000 copies of LINE1 (L1), which make up 17 % of all nucleotide sequences. Among them, only about 100 L1 are active, since they contain the full length of 6000 bp. Among non-autonomous REs, the human genome contains more than 2700 copies of SVA (Hancks, Kazazian, 2012). One of the important factors for the development of the human brain is considered the waves of L1 retrotranspositions, as well as the birth of new TEs, such as SINE, Alu and SVA in the evolution of primates (Linker et al., 2017).

Human endogenous retroviruses (HERV) belong to LTR-RE. They occupy about 8 % of the entire genome and serve as sources of a huge number (794,972) of binding sites with specific transcription factors (TFs), the activation of which plays a role in embryogenesis. For example, in the mesoderm, LTRs interact with SOX17, FOXA1, GATA4; in pluripotent cells, with SOX2, NANOG, POU5F1; in hematopoietic cells, with TAL1, GATA1, PU1 (Ito et al., 2017). Mammalian-wide interspersed repeats (MIRs), which belong to the ancient SINE family descended from tRNA, are also associated with tissue-specific gene expression (Jjingo et al., 2014).

Transposable elements are characterized by nonrandom activation, depending on the tissue and stage of development. High-throughput profiling of integration sites by next-generation sequencing, combined with large-scale genomic data mining and cellular or biochemical approaches, has revealed that the insertions are usually non-random (Sultana et al., 2017). Programmed activation of TE in individual cells during neurogenesis leads to a change in the expression of certain genes necessary for differentiation into specific types of neurons for the formation and functioning of brain structures (Coufal et al., 2009; Bailie et al., 2011; Thomas, Muotri, 2012; Richardson et al., 2014; Evrony et al., 2015; Upton et al., 2015; Muotri, 2016; Suarez et al., 2018). In accordance with this, somatic mosaicism of neurons detected by insertions of TEs (Richardson et al., 2014; Upton et al., 2015; Bachiller et al., 2017; Paquola et al., 2017; Rohrback et al., 2018; Suarez et al., 2018) can reflect the programmed regulatory pattern of the genome necessary for the maturation of specific structures of the central nervous system (Paquola et al., 2017; Rohrback et al., 2018). Somatic mosaicism means the presence, in the same organism, of cells with different genomes as a result of *de novo* DNA changes. These structural variations may be due to CNV, insertions of REs, deletions under the influence of TEs, and SNV (Paquola et al., 2017). This means that in different cells of one organism, not only the genotype, but also the whole genome changes. This is due to the occurrence of mutations in exons of protein-coding genes, intergenic regulatory regions and introns, which is accompanied by a specific expression of certain genes specific for each cell type.

The role of transposable elements in neuronal differentiation

The human brain contains an average of 86.1 billion neurons. Moreover, each of the neurons forms from 5,000 to 20,000 synaptic connections, creating a complex network with a variety of cell types and subtypes. The number of subtypes of neurons is so large that it does not lend itself to modern methods for their description. There must be mechanisms to ensure such a diversity of neurons with their specific temporal and spatial features of functioning (Thomas, Muotri, 2012). The sources of these mechanisms can be TEs, combinations of movements of which can become sources of countless variety. An example of this is the molecular mechanism for generating antibodies by the mammalian immune system (V(D)J recombination), derived from TEs (Lapp, Hunter, 2016). TEs played a role in the development of the central nervous system. In evolution, they turned out to be sources of the formation of regulatory structures and genes involved in the formation of the brain. Non-autonomous TEs MER130 were preserved in the genomes during evolution due to their location near the neocortex genes as a necessary link for their regulation. The experiments showed the activation of MER130 in mouse

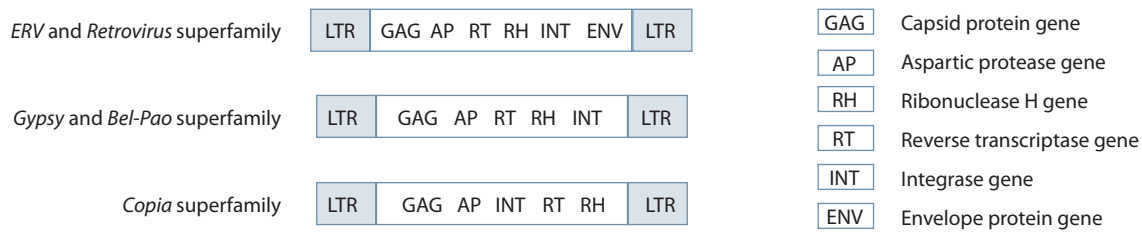


Fig. 1. Scheme of the structure of the genes of LTR-containing retroelements.

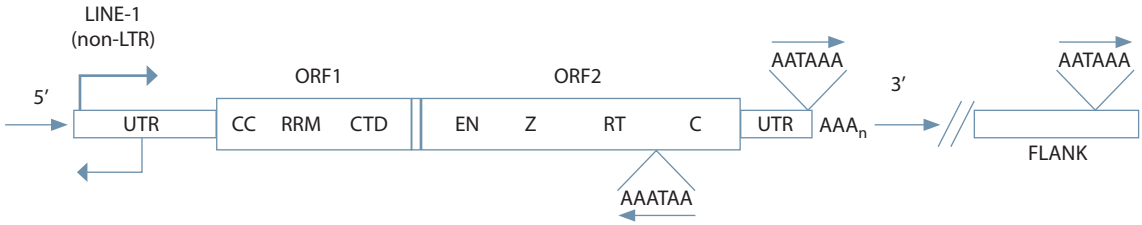


Fig. 2. Scheme of the structure of the gene of non-LTR retroelements (LINE-1).

UTR – untranslated region; ORF – open reading frame; CC – coiled-coiled; RRM – RNA recognition motif; CTD – C-terminal domain; EN – endonuclease; Z – Z-domain; RT – reverse transcriptase; C – cysteine-rich domain.

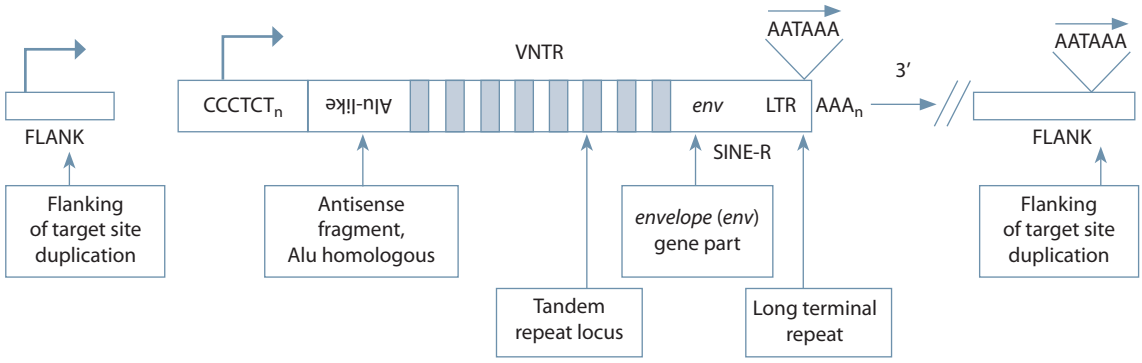


Fig. 3. Scheme of the structure of SVA elements.

embryos on the fourteenth day of development as gene enhancers for the development of the neocortex (Notwell et al., 2015). Among 11 *sushi-ichi*-specific placental animal genes derived from REs, the *SIRH11*/*ZCCHC16* gene encoding zinc finger CCHC protein contributed to the evolution of the brain. This domesticated gene is involved in the development of cognitive functions of placental animals (Irie et al., 2016).

In 2009, in neuronal stem cells isolated from the brain of a human embryo, L1s retrotranspositions were detected, as well as an increase (in comparison with the liver and heart of the same individual) in the number of copies of endogenous L1s in the adult hippocampus (Coufal et al., 2009). In addition to L1 (7743 insertions), a large number of somatic transpositions Alu (13,692 insertions) and SVA (1350 insertions) were found in the hippocampus of adults (Bailie et al., 2011). These *de novo* integrations can affect the expression of certain genes, creating unique transcriptomes of individual neurons (Muotri, 2016). This may be due to the genome-programmed TE ability for their regular site-specific insertions (Sultana et al., 2017). In 2009, of 19 retrotranspositions, 16 were found at a distance of less than 100 kilobases from genes expressed in neurons (Coufal et al., 2009). In 2015, in a study of the somatic mosaicism of the human hippocampus K.R. Upton et al. revealed, out of 20 identified L1 transpositions, 2 functionally significant insertions into the introns of the *ZFAND3* and *USP33* genes functioning in the brain (Upton et al., 2015). A.A. Kurnosov et al., when studying human brain samples, showed that out of 3100 transpositions of L1 in neurons of the dentate gyrus of the hippocampus, 50.26 % of insertions are located in the genes, and out of 2984 Alu, 49.1 % (Kurnosov et al., 2015). In 2016, J.A. Erwin et al. revealed that in the brain of healthy people 44–63 % of neurons undergo somatic mosaicism at the loci of genes that are important for the functioning of the nervous system. For example, a high insertion frequency of L1-RE is shown for the *DLG2* gene, which affects cognitive flexibility, attention, and learning. Mutations in *DLG2* are associated with the development of schizophrenia (Erwin et al., 2016).

Somatic retrotranspositions, unlike germinal ones, cannot be inherited by future generations. However, the programmed ability for specific insertions, depending on the composition and location of TEs in the genome, can be inherited. An explanation of the ability of TEs to be inserted in a site-specific manner in the region of genes expressed in the brain may be the evolutionary relationship of protein-coding genes and their regulatory sequences with TEs (Gianfrancesco et al., 2017; Ito et al., 2017; Joly-Lopez, Bureau, 2018). The insertions specific for humans and chimpanzees were revealed near the promoters of the tachycin receptor genes *TACR3*, cation channels *TRPV1* and *TRPV3*, oxytocin *OXT*. These genes are associated with the functioning of neuropeptides. Analysis of the genomes of various mammals showed that the neural enhancer nPE2, which regulates the expression of the *POMC* gene in the hypothalamus, evolved from SINE in evolution (Gianfrancesco et al., 2017).

Transpositions and expression of TEs can vary depending on the area of the brain and change under environmental influences, as they can perform a number of adaptive functions (Lapp, Hunter, 2016). More active are L1, which retained the ability to transpose, causing somatic mosaicism (Suarez et al., 2018). In 2005, A.R. Muotri et al. suggested that L1 using somatic transpositions can actively create mosaicism of neuron genomes (Muotri et al., 2005). In the brain, somatic mosaicism plays an important role in the regulation of cognition and behavior. The consequences of somatic mosaicism encompass vast changes – from a variant at a single locus, to genes in neuronal networks (Paquola et al., 2017; Rohrback et al., 2018). Moreover, the features of somatic mosaicism differ between neurons of various regions of the brain. For example, in the cerebral cortex, only 0.6 insertions of L1-RE are observed, while in the hippocampus, from 80 to 800 inserts per neuron (Lapp, Hunter, 2016). Somatic mosaicism due to retrotranspositions is a source of phenotypic diversity between neurons during development. In the brain of an adult under the influence of various environmental factors, L1 expression can affect the functioning of neurons during the formation of long-term memory (Bachiller et al., 2017).

The hippocampus is the center of human neurogenesis, where many insertions affect transcriptional expression, creating unique transcriptomes in neurons. In addition, transcriptional activation of L1 is similar to that for the *NeuroD1* gene. This may indicate the effect of L1 expression on neurogenesis, since stimulation of Wnt3a in neuronal stem cells increases L1 expression 10-fold along the beta-catenin pathway, similarly activating transcription of the *NeuroD1* gene. This gene encodes the transcription factor that activates the genes involved in neurogenesis. The *NeuroD1* promoter region contains a Sox/LEF site similar to the 5′UTR of the L1 element, and the pattern of time expression of the *NeuroD1* and L1 genes during differentiation of neurons is similar (Thomas, Muotri, 2012).

Genetic variations between neurons due to L1 retrotranspositions may be associated with specific enrichment of neuronal stem cell enhancers. It was shown that specific enhancers for certain types of neurons (determined using FANTOM5) correspond to the coordinates in the genome for insertions L1, which are within 100 bp from the enhancer. These patterns have not been identified for astrocytes and hepatocytes (Upton et al., 2015). When studying the features of L1 retrotranspositions in more than 30 regions of the brain, a lot of L1 insertion-specific cell lines were found (Evrony et al., 2015). In experiments on mice, specific L1 expression was also shown depending on the area of the brain and the age of the animal (Cappucci et al., 2018).

In addition to L1 elements, LTR-REs are also involved in the regulation of neurogenesis. For example, in mice, the region where the full-length ERVmch8 on chromosome 8 was located was comparatively less methylated in the cerebellum, due to its specific expression depending on the stage of development (Lee et al., 2011). In accordance with these data, it can be assumed that the features of TEs activation observed in neuronal stem cells can naturally alter the expression of specific genes necessary for differentiation of neurons during the formation of specific brain structures. The reason for the activation of TEs in the neuronal stem cells of the hippocampus and the reason for their importance in memory consolidation may be the sensitivity of TEs to stressful environmental influences (Mustafin, Khusnutdinova, 2019). These mechanisms are a particular reflection of the general pattern of epigenetic control of the development of the whole organism, starting from the first division of the zygote, under the regulatory influence of TEs (Mustafin, Khusnutdinova, 2018). To understand the role of TEs in these processes, it is necessary to consider their participation in embryogenesis.

The role of transposable elements in embryogenesis

To initiate the development of the body after fertilization, gametes are reprogrammed to totipotency. During this reprogramming, TEs activation is observed. Previously, this phenomenon was believed to be a side effect of extensive chromatin remodeling at the basis of epigenetic reprogramming of gametes. However, a targeted epigenomic approach has been performed to determine whether TEs directly affect chromatin organization and body development. It was found that silencing of L1 elements reduces the availability of chromatin, and prolonged activation of L1 prevents its gradual compaction, which occurs naturally during development. That is, L1 activation is an integral part of the development program (Jachowicz et al., 2017). In experiments on mice, the role of LTR-REs as a necessary control element for early embryogenesis was proved (Wang et al., 2016).

For the *cis*-regulatory activity of the LTR retroelements ERVK, MERVL and GLN, a complex of RNA and proteins is required, formed using the long ncRNA LincGET. Artificial silencing of LincGET expression in the embryo at the bicellular stage leads to a complete halt to further development due to disruption of *cis*-regulation of the genes necessary for proliferation under the influence of LTR-REs driven by LincGET (Wang et al., 2016). It has also been shown that HERVs are activated in all types of human cells with characteristic features for certain tissues and organs (Seifarth et al., 2005). In the study of the association of 112 TE families in 24 human tissues, tissue-specific enrichment of active regions of LTR-REs was noted, which indicates the involvement of LTE-REs in the regulation of gene expression for differentiation of cells depending on their functional purpose in ontogenesis. This is due to the presence, in the TEs sequences,

of transcription factors binding sites (TFBSs) that regulate the development of the corresponding tissue. TE enrichment characteristic of certain cells in intron enhancers correlates with tissue-specific variations in the expression of nearby genes (Trizzino et al., 2018).

The genetic program in the 2-cell stage of embryogenesis in mice and humans is largely controlled by transcription factors of the DUX family, which are key inducers of zygote genome activation in placental mammals (De Laco et al., 2017). L1 transcripts in embryos are necessary for Dux silencing, rRNA synthesis and exit from the 2-cell stage. M. Percharde et al. in their article showed that L1 expression is required for preimplantation development (Percharde et al., 2018). In embryonic cells, L1 transcripts act as a nuclear RNA scaffold that recruits Nucleolin and Kap1/Trim28 factors for Dux repression. In parallel, L1 products mediate the binding of Nucleolin and Kap1 to rDNA, contributing to the synthesis of rRNA and self-renewal of embryonic stem cells (Percharde et al., 2018). The role of L1 in the repression of the transcriptional program of a 2-cell embryo indicates their participation in the development-specific regulation of gene expression necessary for cell differentiation and body development (Jachowicz et al., 2017). It can be assumed that the activity of REs in neuronal stem cells indicates their use as switches of transcription programs in the specific functionalization of neurons. That is, TEs are involved in the management of both the differentiation of embryonic cells and postnatal stem cells. Regulation is carried out by implementing information encoded in the features of the composition and distribution of TEs in the genome, through the sequential activation of strictly defined TEs in each new cell, specific for the tissue and stage of development. The greatest role is played by this species-specific "coding" in the brain, where neurons are distinguished by higher activity of REs. This is reflected in the structural and functional complexity of the brain compared to other organs. The use of TEs as sources of ncRNAs plays an important role in these processes.

The relationship of transposable elements with non-coding RNAs in the brain

According to recent data, from 75 to 85 % of the human genome is transcribed into primary transcripts, while only 1.2 % of the genome is translated into proteins. Most transcripts are registered as ncRNAs that are involved in the regulation of the genome (Djebali et al., 2012). In humans, 13,000 genes of long ncRNAs have been identified, for the occurrence of which HERVs are responsible by insertion of promoters. HERV-stimulated long ncRNAs are characterized by specific transcription in different types of pluripotent cells, which is consistent with the over-expression of these HERVs in human embryonic stem cells (Johnson, Guigo, 2014). Transcription of most long ncRNAs is associated with the expression of protein coding genes according to the type of neurons and a specific region of the brain. For example, according to Allen Brain Atlas *in situ* hybridization data, out of 1328 known long mouse ncRNAs, 849 are expressed in their brain and are associated with cell types and subcellular structures. The biological significance of these ncRNAs in the functioning of neurons and their relationship with protein-coding genes has been shown (Mercer et al., 2008).

Long ncRNAs expressed in the brain, such as *Miat*, *Rmst*, *Gm17566*, *Gm14207*, *Gm16758*, *2610307P16Rik*, *C230034O21Rik*, *9930014A18Rik*, share a similar expression model with neurogenesis genes and overlap these genes, which proves the role of long ncRNAs in neurogenesis (Aprea et al., 2013). These data are consistent with the role of TEs in neurogenesis (Coufal et al., 2009; Kurnosov et al., 2015; Erwin et al., 2016; Muotri, 2016) and regulation of brain function (Thomas, Muotri, 2012; Upton et al., 2015; Rohrback et al., 2018). This is because TEs are the main sources of the emergence and evolution of long ncRNAs, forming their functional domains and making up more than 2/3 of their mature transcripts in humans (Kapusta, Feschotte, 2014). REs can serve as genes for long ncRNAs (Lu et al., 2014). L1s have a function similar to lncRNA in regulating the expression of genes necessary for self-renewal of stem cells and for preimplantation development (Honson, Macfarlan, 2018).

In a number of studies, the role of miRNAs in controlling the differentiation of neurons, switching expression profiles of genes important for cell function in time and space has been proved (Stappert et al., 2015). About 40 % of all known human miRNAs are expressed in the human brain. The specific expression of many of them differs in different types of cells and is important in the regulation of differentiation, which is necessary for a huge variety of phenotypes of neurons in the brain (Smirnova et al., 2005). The accumulation of certain miRNAs in various structures of neurons (axons, dendrites, synapses) was revealed. For example, in experiments in mice, the role of miR-134 in the regulation of specific mRNAs of the *LIMK1* gene for the growth of dendritic spins was shown, and the accumulation of miR-99a, 124a1-3, 125b1, 125b2, 134, 339 was noted in synaptosomes (Lugli et al., 2008). The formation of neurites is promoted by miR-21 (the target is the mRNA of the *SPRY2* gene), miR-431 is involved in the regeneration of axons (the target is the *Kremen-1* gene), differentiation of neurons occurs under the influence of miR-34a (the targets are *Tap73*, *synaptotagmin-1*, *syntaxin-1A*) and miR-137 (targets are the *Mib1*, *Ezh2* genes). Enhanced expression of miR-9 promotes branching and reduced axon growth by repressing microtubule-associated Map1b protein. Axon growth depends on the effect of miR-431, as well as miR-17-92, which interacts with PTEN (phosphate tensin homolog) in neurons of the cerebral cortex of the embryo. The regulatory role of differential expression of miR-221 and miR-222 in neurogenesis has also been proven (Nampoothiri, Rajanikant, 2017).

In 2007, J. Piriyapongsa et al. found that in humans TEs can be sources of microRNAs (Piriyapongsa et al., 2007), which was confirmed by other researchers (Yuan et al., 2010, 2011; Qin et al., 2015). The key role of TEs in the formation of microRNAs and long ncRNAs (Johnson, Guigo, 2014; Kapusta, Feschotte, 2014) indicates that the maximum activity of TEs at the center of human neurogenesis (Kurnosov et al., 2015) as a natural phenomenon is necessary for epigenetic control of differentiation of neuronal stem cells. Another mechanism of TE participation in the regulation of gene expression necessary for the specific work of neurons is the *cis*- and *trans*-effects of TEs (Garcia-Perez et al., 2016). This confirms the nonrandom activations of TEs as sources of heterogeneous subpopulations of neurons (Fig. 4) (Faulkner, 2011).

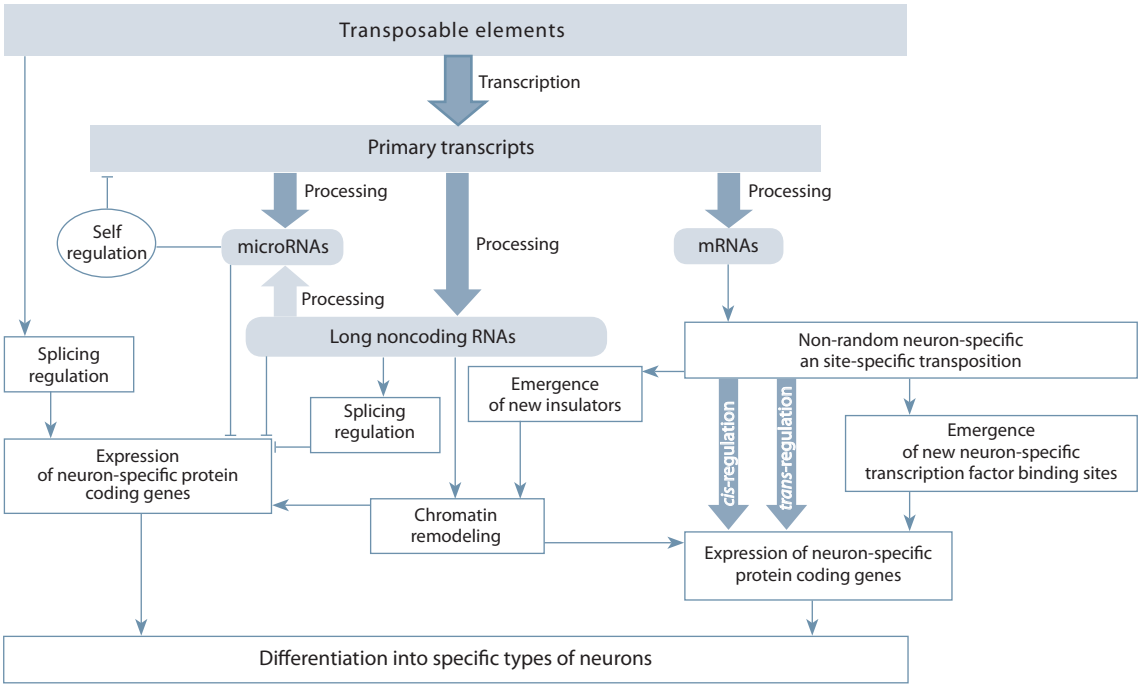


Fig. 4. Scheme of TE involvement in neurogenesis.

The role of retroelements in interactions between neurons

For the development and functioning of the brain, intercellular interactions are necessary, the study of the regulation mechanisms of which is promising for therapeutic targeted exposure to the work of the brain. For this, it is important to identify drivers for gene expression and post-transcriptional epigenetic regulation of the structural components of neurons. Based on the analysis of the accumulated data on the role of TEs in controlling the functioning of the genome in embryonic development (Garcia-Perez et al., 2007; Van den Hurk et al., 2007; Macia et al., 2011; Kurnosov et al., 2015; Percharde et al., 2018) and the physiological functioning of the human brain (Coufal et al., 2009; Bailie et al., 2011; Thomas, Muotri, 2012; Richardson et al., 2014; Evrony et al., 2015; Upton et al., 2015; Muotri, 2016; Suarez et al., 2018), it was concluded that TEs are regulators of epigenetic control for gene function in ontogenesis (Mustafin, Khusnutdinova, 2017, 2018). Despite the lack of mitotic activity of mature neurons, the specific expression of TEs in them is important in controlling both interneuronal interactions and the structural and functional characteristics of neurons (Bailie et al., 2011; Richardson et al., 2014; Erwin et al., 2016). These properties may be due to processing from transcripts of transposons of specific long ncRNAs (Lu et al., 2014; Honson, Macfarlan, 2018) and microRNAs (Piriyapongsa et al., 2007; Yuan et al., 2010, 2011; Qin et al., 2015). Indeed, in experiments on laboratory animals, the enrichment of specific miRNAs in certain structures and regions of neurons was revealed. For example, an abundance of miR-15b, miR-16, miR-204, miR-221 was found in the distal axons compared to neuron bodies (Natera-Naranjo et al., 2010). Enrichment of specific miRNAs in synapses was detected. This suggests a local post-transcriptional regulation of the expression of neuron-specific genes (Lugli et al., 2008). The role of miRNAs in intercellular interactions in the brain was shown, as well as the value of the electrical activity of neurons for the secretion of miR-124 and miR-9, which can penetrate microglia and change the phenotype of its cells (Veremeyko et al., 2019).

Transposable elements regulate brain function through expression into specific microRNAs that regulate gene expression in neurons and in intercellular interactions in the brain. The role of ERV in transferring information between neurons for memory consolidation has also been identified. In the human genome, the full-length HERV-K (about 10,000 bp) consists of the remains of ancient retroviruses and includes LTR-flanked regions, including three retroviral ORFs: *pol-pro* (encodes protease, RT and integrase enzymes), *env* (encodes horizontal transfer proteins) and *gag* (encodes structural proteins of the retroviral capsid) (Klein, O'Neill, 2018). In the course of evolution, the specific ERV *Ty3/gypsy* has become the source of Arc protein. This protein is similar in biological properties to the *gag* retroviral gene expression product (Pastuzyn et al., 2018).

Since domestication and use for the needs of the host, the *Arc* gene has become highly conserved for vertebrates, playing a role in the functioning of their brain. Expression of *Arc* is highly dynamic in the brain in accordance with the encoding of information in neural networks. *Arc* gene transcript is transported to dendrites and accumulates in areas of local synaptic activity, where translation into protein occurs (Shepherd, 2018). In neurons, the Arc protein forms spatial structures resembling viral capsids that encapsulate cell mRNA. The resulting virus-like elements in the composition of extracellular vesicles are transmitted to neighboring neurons, where they are able to translate. This phenomenon is used to consolidate

long-term memory (Pastuzyn et al., 2018), in the formation of which the hippocampus is involved, where the maximum activity of TEs is detected (Coufal et al., 2009; Bailie et al., 2011; Thomas, Muotri, 2012; Bachiller et al., 2017).

Based on the data listed above, it can be concluded that the observed phenomenon of intercellular neuronal interconnection using Arc has developed in evolution as a reflection of the adaptive value of the TE transcript transfer phenomenon between postmitotic cells. It is possible that when neurons exchange virus-like mRNA particles between neurons, the ability of TEs to be integrated in a site-specific manner (Sultana et al., 2017) with a change in the expression of neuron-specific genes is used to form long-term memory. As a result, the functioning of neurons and the storage of information in the brain change (Bachiller et al., 2017).

Other functions of transposable elements

Transposable elements transpositions affect gene expression in various ways. Insertions within a gene can cause frameshift mutations, premature stop codons, or exon skipping. In the transcribed portion of the gene, TEs can reduce mRNA levels by slowing transcription due to the high A/T content in ORF2 of TEs such as L1 RE (Thomas, Muotri, 2012). However, despite the potentially mutagenic effect TEs play a role in the evolution of the genomes of all eukaryotes through the use of TE sequences to form host adaptive abilities (Mustafin, Khusnutdinova, 2019). TEs are involved in controlling the expression of protein-coding genes, many of which (Joly-Lopez, Bureau, 2018), including transcription factors (Ito et al., 2017), originated from TEs. In addition to the direct domestication of TEs, new protein-coding genes were formed due to exonization and duplication of genes using TEs (Thomas, Muotri, 2012; Joly-Lopez, Bureau, 2018; Mustafin, Khusnutdinova, 2018).

Mechanisms derived from TEs are used by the mammalian immune system to generate antibodies using the V(D)J recombination system. TEs are the source of most steroid receptors, participating in the global regulation of cell function by the hormonal system (Lapp, Hunter, 2016). Regulatory sequences, silencers, and insulators evolved from TEs (Jjingo et al., 2014; Ito et al., 2017; Schrader, Schmitz, 2018). If TEs are inserted into non-coding regions of genomes, they are used as alternative promoters, enhancers, and polyadenylation signals of genes. For example, L1s are found in non-coding regions of 80 % of human genes, the expression pattern of which depends on the density of these REs (Klein, O'Neill, 2018).

About 60 % of all SVAs in the human genome are located in the genes or flank them within 10 kb. These SVAs are characterized as mobile CpG islands capable of upstream or downstream regulation of gene expression by recruiting transcription factors. In addition, due to the high GC content, SVAs can form alternative DNA structures, such as the G-quadruplex (characteristic of promoters of 40 % of human genes), which affects transcription (Gianfrancesco et al., 2017). Many transcription factors are immediately directed to the relationship with TEs, forming and maintaining heterochromatin (Lapp, Hunter, 2016). TEs serve as sources of *cis*- and *trans*-regulatory elements that coordinate the expression of groups of genes. In addition to acting as promoters that control the expression of alternative host gene isoforms, TFBS within TEs can act as enhancers in certain tissues and at certain stages of development (Garcia-Perez et al., 2016).

In evolution, TEs were the sources of a significant part of the specific sequences of the genome, as well as transcripts and proteins interacting with them. This indicates a global regulatory role of TEs, necessary for both mitosis and meiosis, and for controlling the work of cells in interphase. For example, not only spliceosomal introns (Kubiak, Makalowska, 2017), but also the Prp8 spliceosome component originated from TEs (Galej et al., 2013). Splicing enhancers and silencers are 10-nucleotide-long ncRNAs that interact with SR proteins and snRNAs. They are formed by processing transcripts of Alu retroelements (Pastor et al., 2009). TEs turned out to be sources of satellites due to the capability of site-specific insertions (McGurk, Barbash, 2018) and illegitimate recombination, followed by amplification by gene conversion (Han et al., 2016). In evolution, TEs have become sources of telomerase and telomeres (Kopera et al., 2011), as well as centromeres (Cheng, Murata, 2003; Sharma et al., 2013; Han et al., 2016) and the protein CENP/CENH3 interacting with them (Lopez-Flores et al., 2004; Volff, 2006). Small ncRNAs formed upon transcription of centromeric REs are involved in the regulation of these interactions (Carone et al., 2013).

Conclusion

Less than 1.2 % of the human genome is responsible for the coding of proteins. The remaining non-coding part of the genome was largely formed due to TEs. The data on the participation of TEs in the regulation of gene switching during cell differentiation in embryogenesis, starting with the first zygote division, suggest that somatic mosaicism observed in neurons reflects the active role of TEs in neurogenesis. A number of papers have been published proving the participation of TEs in the control of differentiation of neurons. Transposable elements are sources of ncRNA, which are also important in gene switching in brain cells. The revealed role of LTR-containing REs in the exchange of transcripts between neurons may reflect the general principle of the participation of TEs in the regulation of gene expression for the development and maintenance of brain function. The use of Arc protein for the formation of virus-like particles in the transfer of information between cells indicates the evolutionary mechanisms of TE conversion into viruses for the formation of adaptive functions. This mechanism is associated with the use of TEs to ensure the dynamism of the genomes of postmitotic cells with the possibility of their adaptive changes in response to environmental influences. The realization of this phenomenon is possible due to reverse transcription of mRNA transported between cells with site-specific insertions, the formation of somatic mosaicism of mature neurons, and a change in gene expression for memory consolidation.

Since somatic mosaicism cannot be inherited, the functional role of TE insertions in neurogenesis is difficult to prove. Moreover, these changes can be characterized as random events that are more important for the development of neurological disorders. However, the data presented in the review prove the importance of TE transpositions into functionally significant regions of the genome, which are necessary for differentiation of neuronal stem cells and the response to environmental influences. The explanation of this

regular phenomenon is the capability of TEs to be inserted in a site-specific manner programmed by their own position in the genome. These nonrandom events are selected during the evolution of multicellular organisms, promoting regulatory regulation of gene expression during cell differentiation.

The results obtained on the importance of TE transpositions in neurogenesis reflect one of the stages of regulation of gene expression in successive cell divisions during differentiation of tissues and organs of the whole organism. Somatic mosaicism in neurons and stem cells is in favor of this assumption, since the brain is characterized by a pronounced variety of cell types, for the specific tuning of gene expression of which universal combinatorial units, such as TEs, are required.

References

Allen Brain Atlas. Available at: www.brain-map.org

Aprea J., Prenninger S., Dori M., Ghosh T., Monasor L.S., Wessendorf E., Zocher S., Massalini S., Alexopoulou D., Lesche M., Dahl A., Groszer M., Hiller M., Calegari F. Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment. EMBO J. 2013;32(24):3145-3160.

Bachiller S., Del-Pozo-Martin Y., Carrio A.M. L1 retrotransposition alters the hippocampal genomic landscape enabling memory formation. Brain. Behav. Immun. 2017;64:65-70.

Bailie J.K., Barnett M.W., Upton K.R., Gerhardt D.J., Richmond T.A., De Sapio F., Brennan P.M., Rizzu P., Smith S., Fell M., Talbot R.T., Gustinicich S., Freeman T.C., Mattick J.S., Hume D.A., Heutink D.A., Carninci P., Jeddeloh J.A., Faulkner G.J. Somatic retrotransposition alters the genetic landscape of the human brain. Nature. 2011;479:534-537.

Cappucci U., Torromino G., Casale A.M., Camon J., Capitano F., Berloco M., Mele A., Pimpinelli S., Rinaldi A., Piacentini L. Stress-induced strain and brain region-specific activation of LINE-1 transposons in adult mice. Stress. 2018;21:575-579.

Carone D.M., Zhang C., Hall L.E., Obergfell C., Carone B.R., O'Neill M.J., O'Neill R.J. Hypermorphic expression of centromeric retroelement-encoded small RNAs impairs CENP-A loading. Chromosome Res. 2013;21:49-62.

Cheng Z.J., Murata M. A centromeric tandem repeat family originating from a part of Ty3/gypsy-retroelement in wheat and its relatives. Genetics. 2003;164:665-672.

Coufal N.G., Garcia-Perez J.L., Peng G.E., Yeo G.W., Mu Y., Lovci M.T., Morell M., O'Shea K.S., Moran J.V., Gage F.H. L1 retrotransposition in human neural progenitor cells. Nature. 2009;460: 1127-1131.

De Koning A.P., Gu W., Castoe T.A., Batzer M.A., Pollock D.D. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet. 2011;7:e1002384.

De Laco A., Planet E., Coluccio A., Verp S., Cuc J., Trono D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. Nat. Genet. 2017;49:941-945.

Djebali S., Davis C.A., Merkel A., Dobin A., Lassmann T., Mortazavi A., Tanzer A., Lagarde J., Lin W., Schlesinger F., Xue C., Marinov G.K., Khatun J., Williams B.A., Zaleski C., Rozowsky J., Roder M., Kokocinski F., Abdelhamid R.F., Alioto T., Antoshechkin I., Carninci P., Guigo R., Gingeras T.R. Landscape of transcription in human cells. Nature. 2012;489(7414):101-108.

Erwin J.A., Paquola A.C., Singer T., Gallina I., Novotny M., Quayle C., Bedrosian T., Ivanio F., Butcher C.R., Herdy J.R., Sarkar A., Lasken R.S., Muotri A.R., Gage F.H. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. Nat. Neurosci. 2016;19:1583-1591.

Evrony G.D., Lee E., Mehta B.K., Benjamini Y., Johnson R.M., Cai X., Yang L., Haseley P., Lehmann H.S., Park P.J., Walsh C.A. Cell lineage analysis in human brain using endogenous retroelements. Neuron. 2015;85:49-59.

Faulkner G.J. Retrotransposons: mobile and mutagenic from conception to death. FEBS Lett. 2011;585(11):1589-1594.

Galej W.P., Oubridge C., Newman A.J., Nagai K. Crystal structure of Prp8 reveals active site cavity of the spliceosome. Nature. 2013; 493:638-643.

Garcia-Perez J.L., Marchetto M.C., Muotri A.R., Coufal N.G., Gage F.H., O'Shea K.S., Moran J.V. LINE-1 retrotransposition in human embryonic stem cells. Hum. Mol. Genet. 2007;16(13):1569-1577.

Garcia-Perez J.L., Widmann T.J., Adams I.R. The impact of transposable elements on mammalian development. Development. 2016; 143:4101-4114.

Gianfrancesco O., Bubb V.J., Quinn J.P. SVA retrotransposons as potential modulators of neuropeptide gene expression. Neuropeptides. 2017;64:3-7.

Habibi L., Pedram M., AmirPhirozy A., Bonyadi K. Mobile DNA elements: the seeds of organic complexity on earth. DNA Cell Biol. 2015;34:597-609.

Han J., Masonbrink R.E., Shan W., Song F., Zhang J., Yu W., Wang K., Wu Y., Tang H., Wendel J.F., Wang K. Rapid proliferation and nucleolar organizer targeting centromeric retrotransposons in cotton. Plant J. 2016;88:992-1005.

Hancks D.C., Kazazian H.H. Jr. Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev. 2012;22(3):191-203. DOI 10.1016/j.gde.2012.02.006.

Honson D.D., Macfarlan T.S. A lncRNA-like role for LINE1s in development. Dev. Cell. 2018;46:132-134.

Irie M., Koga A., Kaneko-Ishino T., Ishino F. An LTR retrotransposon-derived gene displays lineage-specific structural and putative species-specific functional variations in eutherians. Front. Chem. 2016; 4:26.

Ito J., Suqimoto R., Nakaoka H., Yamada S., Kimura T., Hayano T., Inoue I. Systematic identification and characterization of regulatory elements derived from human endogenous retroviruses. PLoS Genet. 2017;13(7):e1006883.

Jachowicz J.W., Bing X., Pontabry J., Boskovic A., Rando O.J., Torres-Padilla M.E. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. Nat. Genet. 2017;49:1502-1510.

Jjingo D., Conley A.B., Wang J., Marino-Ramirez L., Lunyak V.V., Jordan I.K. Mammalian-wide interspersed repeat (MIR)-derived enhancers and the regulation of human gene expression. Mob. DNA. 2014;5:5-14.

Johnson R., Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. RNA. 2014;20(7): 959-976.

Joly-Lopez Z., Bureau T.E. Exaptation of transposable element coding sequences. Curr. Opin. Genet. Dev. 2018;49:34-42.

Kapusta A., Feschotte C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. Trends Genet. 2014;30(10):439-452.

Klein S.J., O'Neill R.J. Transposable elements: genome innovation, chromosome diversity, and centromere conflict. Chromosome Res. 2018;26:5-23.

Kopera H.C., Moldovan J.B., Morrish T.A., Garcia-Perez J.L., Moran J.V. Similarities between long interspersed element-1 (LINE-1)

reverse transcriptase and telomerase. Proc. Natl. Acad. Sci. USA. 2011;108:20345-20350.

Kubiak M.R., Makalowska I. Protein-coding genes' retrocopies and their functions. Viruses. 2017;9(4):pii: E80.

Kurnosov A.A., Ustyugova S.V., Nazarov V., Minervina A.A., Komkov A.Y., Shugay M., Pogorelyy M.V., Khodosevich K.V., Mamedov I.Z., Lebedev Y.B. The evidence for increased L1 activity in the site of human adult brain neurogenesis. PLoS One. 2015;10(2): e0117854.

Lapp H.E., Hunter R.G. The dynamic genome: transposons and environmental adaptation in the nervous system. Epigenomics. 2016;8:237.

Lee K.H., Horiuchi M., Itoh T., Greenhalgh D.G., Cho K. Cerebellum-specific and age-dependent expression of an endogenous retrovirus with intact coding potential. Retrovirology. 2011;8:82.

Linker S.B., Marchetto M.C., Narvaiza I., Denli A.M., Gage F.H. Examining non-LTR retrotransposons in the context of the evolving primate brain. BMC Biol. 2017;15:68.

Lopez-Flores I., de la Herran R., Garrido-Ramos M.A., Boudry P., Ruiz-Rejon C., Ruiz-Rejon M. The molecular phylogeny of oysters based on a satellite DNA related to transposons. Gene. 2004;339: 181-188.

Lu X., Sachs F., Ramsay L., Jacques P.E., Goke J., Bourque G., Ng H.H. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. Nat. Struct. Mol. Biol. 2014; 21(4):423-425.

Lugli G., Torvik V.L., Larson J., Smalheiser N.R. Expression of microRNAs and their precursors in synaptic fractions of adult mouse forebrain. J. Neurochem. 2008;106:650-661.

Macia A., Munoz-Lopez M., Cortes J.L., Hastings R.K., Morell S., Lucena-Aguilar G., Marchal J.A., Badge R.M., Garcia-Perez J.L. Epigenetic control of retrotransposons expression in human embryonic stem cells. Mol. Cell. Biol. 2011;31(2):300-316.

McGurk M.P., Barbash D.A. Double insertion of transposable elements provides a substrate for the evolution of satellite DNA. Genome Res. 2018;28(5):714-725.

Mercer T.R., Dinger M.E., Sunkin S.M., Mehler M.F., Mattick J.S. Specific expression of long noncoding RNAs in the mouse brain. Proc. Natl. Acad. Sci. USA. 2008;105(2):716-721.

Muotri A.R. L1 retrotransposition in neural progenitor cells. Methods Mol. Biol. 2016;1400:157-163.

Muotri A.R., Chu V.T., Marchetto M.C., Deng W., Moran J.V., Gage F.H. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature. 2005;435:903-910.

Mustafin R.N., Khusnutdinova E.K. Non-coding parts of genomes as the basis of epigenetic heredity. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2017;21(6): 742-749. DOI 10.18699/VJ17.30-o. (in Russian)

Mustafin R.N., Khusnutdinova E.K. The role of transposons in epigenetic regulation of ontogenesis. Russian Journal of Developmental Biology. 2018;49(2):61-78.

Mustafin R.N., Khusnutdinova E.K. The role of transposable elements in the ecological morphogenesis under the influence of stress. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(4):380-389. DOI 10.18699/VJ19.506. (in Russian)

Nampoothiri S.S., Rajanikant G.K. Decoding the ubiquitous role of microRNAs in neurogenesis. Mol. Neurobiol. 2017;54:2003-2011.

Natera-Naranjo O., Aschrafi A., Gioio A.E., Kaplan B.B. Identification and quantitative analyses of microRNAs located in the distal axons of sympathetic neurons. RNA. 2010;16:1516-1529.

Notwell J.H., Chung T., Heavner W., Bejerano G. A family of transposable elements co-opted into developmental enhancers in the mouse neocortex. Nat. Commun. 2015;6:6644.

Paquola A.C.M., Erwin J.A., Gage F.H. Insight into the role of somatic mosaicism in the brain. Curr. Opin. Syst. Biol. 2017;1:90-94.

Pastor T., Talotti G., Lewandowska M.A., Pagani F. An Alu-derived intronic splicing enhancer facilitates intronic processing and modulates aberrant splicing in ATM. Nucleic Acids Res. 2009;37:7258-7267.

Pastuzyn E.D., Day C.E., Kearns R.B., Kyrke-Smith M., Taibi A.V., McCormick J., Yoder N., Belnap D.M., Erlendsson S., Morado D.R., Briggs J.A.G., Feschotte C., Shephered J.D. The neuronal gene *Arc* encodes a repurposed retrotransposon Gag protein that mediates intercellular RNA transfer. Cell. 2018;172:275-278.

Percharde M., Lin C.J., Yin Y., Guan J., Peixoto G.A., Bulut-Rarslioglu A., Biechele S., Huang B., Shen X., Ramalho-Santos M. A LINE1-nucleolin partnership regulates early development and ESC identity. Cell. 2018;174:391.

Piriyapongsa J., Marino-Ramirez L., Jordan I.K. Origin and evolution of human microRNAs from transposable elements. Genetics. 2007; 176(2):1323-1337.

Qin S., Jin P., Zhou X., Chen L., Ma F. The role of transposable elements in the origin and evolution of microRNAs in human. PLoS One. 2015;10(6):e0131365.

Richardson S.R., Morell S., Faulkner G.J. L1 retrotransposons and somatic mosaicism in the brain. Annu. Rev. Genet. 2014;48:1-27.

Rohrback S., Siddoway B., Liu C.S., Chun J. Genomic mosaicism in the developing and adult brain. Dev. Neurobiol. 2018;78:1026-1048.

Schrader L., Schmitz J. The impact of transposable elements in adaptive evolution. Mol. Ecol. 2018(6):1537-1549. DOI 10.111/mec.14794.

Seifarth W., Frank O., Zeilfelder U., Spiess B., Greenwood A.D., Hehlmann R., Leib-Mosch C. Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray. J. Virol. 2005;79(1):341-352.

Sharma A., Wolfgruber T.K., Presting G.G. Tandem repeats derived from centromeric retrotransposons. BMC Genomics. 2013;14:142.

Shepherd J.D. Arc – An endogenous neuronal retrovirus? Semin. Cell Dev. Biol. 2018;77:73-78.

Smirnova L., Grafe A., Seiler A., Schumacher S., Nitsch R., Wulczyn F.G. Regulation of miRNA expression during neural cell specification. Eur. J. Neurosci. 2005;21:1469-1477.

Stappert L., Roese-Koerner B., Brustle O. The role of microRNAs in human neural stem cells, neuronal differentiation and subtype specification. Cell Tissue Res. 2015;359:47-64.

Suarez N.A., Macia A., Muotri A.R. LINE-1 retrotransposons in healthy and diseased human brain. Dev. Neurobiol. 2018;78:434-455.

Sultana T., Zamborlini A., Cristofari G., Lesage P. Integration site selection by retroviruses and transposable elements in eukaryotes. Nat. Rev. Genet. 2017;18(5):292-308.

Thomas C.A., Muotri A.R. LINE-1: creators of neuronal diversity. Front. Biosci. (Elite Ed). 2012;4:1663-1668.

Trizzino M., Kapusta A., Brown C.D. Transposable elements generate regulatory novelty in a tissue-specific fashion. BMC Genomics. 2018;19:468.

Upton K.R., Gerhardt D.J., Jesuadian J.S., Richardson S.R., Sanchez-Luque F.J., Bodea G.O., Ewing A.D., Salvador-Palomeque C., van der Knaap M.S., Brennan P.M., Vanderver A., Faulkner G.J. Ubiquitous L1 mosaicism in hippocampal neurons. Cell. 2015;161(2): 228-239.

Van den Hurk J.A., Meij I.C., Seleme M.C. L1 retrotransposition can occur early in human embryonic development. Hum. Mol. Genet. 2007;16(13):1587-1592.

Veremeyko T., Kuznetsova I.S., Dukhinova M., Yung A., Kopeikina E., Barteneva N.S., Ponomarev E.D. Neuronal extracellular microRNAs miR-124 and miR-9 mediate cell-cell communication between neurons and microglia. J. Neurosci. Res. 2019;97(2):162-184.

Volff J.N. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. BioEssays. 2006; 28:913-922.

Wang J., Li X., Wang L., Li J., Zhao Y., Bou G., Li Y., Jiao G., Shen X., Wei R., Liu S., Xie B., Lei L., Li W., Zhou Q., Liu Z. A novel long intergenic noncoding RNA indispensable for cleavage of mouse two-cell embryos. EMBO Rep. 2016;17:1452-1470.

Yuan Z., Sun X., Jianq D., Ding Y., Lu Z., Gong L., Liu H., Xie J. Origin and evolution of a placental-specific microRNA family in the human genome. BMC Evol. Biol. 2010;10:346-358.

Yuan Z., Sun X., Liu H., Xie J. MicroRNA genes derived from repetitive elements and expanded by segmental duplication events in mammalian genomes. PLoS One. 2011;6(3):e17666.

ORCID ID

R.N. Mustafin orcid.org/0000-0002-4091-382X
E.K. Khusnutdinova orcid.org/0000-0003-2987-3334

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

Received June 19, 2019. Revised September 22, 2019. Accepted October 2, 2019.

Вавиловский журнал генетики и селекции. 2020;24(2):219-227
DOI 10.18699/VJ20.614

Modern classification and molecular-genetic aspects of osteogenesis imperfecta

A.R. Zaripova[1] ✉, R.I. Khusainova[1, 2]

[1] Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia
[2] Republican Medical-Genetic Center, Ufa, Russia
✉ e-mail: a.ramilna@bk.ru

Abstract. Osteogenesis imperfecta (imperfect osteogenesis in the Russian literature) is the most common hereditary form of bone fragility, it is a genetically and clinically heterogeneous disease with a wide range of clinical severity, often leading to disability from early childhood. It is based on genetic disorders leading to a violation of the structure of bone tissue, which leads to frequent fractures, impaired growth and posture, with the development of characteristic disabling bone deformities and associated problems, including respiratory, neurological, cardiac, renal impairment, hearing loss. Osteogenesis imperfecta occurs in both men and women, the disease is inherited in both autosomal dominant and autosomal recessive types, there are sporadic cases of the disease due to *de novo* mutations, as well as X-linked forms. The term "osteogenesis imperfecta" was coined by W. Vrolick in the 1840s. The first classification of the disease was made in 1979 and has been repeatedly reviewed due to the identification of the molecular cause of the disease and the discovery of new mechanisms for the development of osteogenesis imperfecta. In the early 1980s, mutations in two genes of collagen type I (*COL1A1* and *COL1A2*) were first associated with an autosomal dominant inheritance type of osteogenesis imperfecta. Since then, 18 more genes have been identified whose products are involved in the formation and mineralization of bone tissue. The degree of genetic heterogeneity of the disease has not yet been determined, researchers continue to identify new genes involved in its pathogenesis, the number of which has reached 20. In the last decade, it has become known that autosomal recessive, autosomal dominant and X-linked mutations in a wide range of genes, encoding proteins that are involved in the synthesis of type I collagen, its processing, secretion and post-translational modification, as well as in proteins that regulate the differentiation and activity of bone-forming cells, cause imperfect osteogenesis. A large number of causative genes complicated the classical classification of the disease and, due to new advances in the molecular basis of the disease, the classification of the disease is constantly being improved. In this review, we systematized and summarized information on the results of studies in the field of clinical and genetic aspects of osteogenesis imperfecta and reflected the current state of the classification criteria for diagnosing the disease.
Key words: osteogenesis imperfecta; collagen; bone fragility; bisphosphonates; multiple fractures.

For citation: Zaripova A.R., Khusainova R.I. Modern classification and molecular-genetic aspects of osteogenesis imperfecta. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2020;24(2):219-227. DOI 10.18699/VJ20.614

Современная классификация и молекулярно-генетические аспекты незавершенного остеогенеза

А.Р. Зарипова[1] ✉, Р.И. Хусаинова[1, 2]

[1] Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
[2] Республиканский медико-генетический центр, Уфа, Россия
✉ e-mail: a.ramilna@bk.ru

Аннотация. Незавершенный остеогенез (несовершенный остеогенез в русскоязычной литературе) – наиболее распространенная наследственная форма ломкости костей, генетически и клинически гетерогенное заболевание с широким спектром клинической тяжести, основное клиническое проявление которого – множественные переломы начиная с натального периода жизни, зачастую приводящие к инвалидизации с детского возраста. К основным клиническим признакам незавершенного остеогенеза относятся голубые склеры, потеря слуха, аномалия дентина, повышенная ломкость костей, нарушение роста и осанки с развитием характерных инвалидизирующих деформаций костей и сопутствующих проблем, включающих дыхательные, неврологические, сердечные, почечные нарушения. Незавершенный остеогенез встречается и у мужчин, и у женщин, заболевание наследуется как по аутосомно-доминантному, так и аутосомно-ре-

цессивному типам, существуют спорадические случаи заболевания, обусловленные мутациями *de novo*, а также обнаружены Х-сцепленные формы. Термин «незавершенный остеогенез» был введен W. Vrolick в 1840-х гг. Первая классификация заболевания сделана в 1979 г. и неоднократно пересматривалась из-за идентификации молекулярной причины заболевания и открытия новых механизмов развития незавершенного остеогенеза. В начале 1980-х гг. мутации в двух генах коллагена типа I (*COL1A1* и *COL1A2*) впервые были ассоциированы с аутосомно-доминантным типом наследования незавершенного остеогенеза. С тех пор идентифицированы еще 18 генов, продукты которых участвуют в процессах формирования и минерализации костной ткани. До сих пор не определена степень генетической гетерогенности заболевания, исследователи продолжают идентифицировать новые гены, вовлеченные в его патогенез, число которых достигло 20. В последнее десятилетие стало известно, что аутосомно-рецессивные, аутосомно-доминантные и Х-связанные мутации в широком спектре генов, кодирующих белки, участвующие в синтезе коллагена типа I, его процессинге, секреции и посттрансляционной модификации, а также в белках, регулирующих дифференцировку и активность костеобразующих клеток, вызывают несовершенный остеогенез. Большое количество причинных генов усложнило классическую классификацию заболевания, и в связи с новыми достижениями в области молекулярных основ незавершенного остеогенеза постоянно совершенствуется и классификация. В этом обзоре мы систематизировали и обобщили информацию о результатах исследований в области изучения клинико-генетических аспектов незавершенного остеогенеза и отразили современное состояние классификационных критериев диагностики заболевания.
Ключевые слова: незавершенный (несовершенный) остеогенез; коллаген; хрупкость костей; бисфосфонаты; множественные переломы.

Introduction

Osteogenesis imperfecta (OI), also known as brittle bone disease, is a clinically and genetically heterogeneous hereditary disease of connective tissue, the main cause of which is a genetically determined violation of the quality of bone tissue, leading to frequent fractures with the development of disabling bone deformities and a complex of concomitant problems on the part of the respiratory, cardiovascular, neuromuscular systems.

Worldwide osteogenesis imperfecta occur with a frequency approximately of 1 in every 30,000 births. The desease affects both men and women. In Russia osteogenesis imperfecta is the most common genetic bone disease – one case per 10– 20 thousand newborns. According to the ministry of health in 2014 in Russia there are 556 adults and children with osteogenesis imperfecta (Kruchkova, Kruglov, 2014). In the past decade, (mostly) recessive, dominant and X-linked defects in a wide variety of genes encoding proteins involved in type I collagen synthesis, processing, secretion and post-translational modification, as well as in proteins that regulate the differentiation and activity of bone-forming cells have been shown to cause osteogenesis imperfecta (Marini et al., 2017). Also sporadic cases of osteogenesis imperfecta are affected by *de novo* mutations, which frequency is necessary to find out.

Nowadays, 20 genes are responsible for the development of different types of osteogenesis imperfecta and the search of new genes that take part in pathogenesis of the disease is still continuing. In past five years 6 new genes, which take part in pathogenesis of osteogenesis imperfecta, were identified. The last gene was identified in 2018 and it is not still known if the disease clinically and genetically heterogenious. Genetic defects, that lead to OI, are transformed into the defects of collagen synthesis, structures of its chains, post-translational modification of collagen, proper twisting into a triple helix and stitching (Nadyrshina et al., 2012). Also there are defects of bone tissue mineralization and osteoblasts differentiation. Due to the identification of new molecular causes of the disease, continuous improvement of diagnostic criteria and revision of classification of OI is carried out.

The aim of this article is the review of current state of clinical and genetic aspects of OI and the generalization of the results of molecular pathogenesis of the disease.

Evolution of classification criteria of osteogenesis imperfecta

The existance of clinical features, which are corresponded to osteogenesis imperfecta, had been known from ancient times. The earliest case of the disease was identified in 1000 BC in the study of a partially mummified skeleton of an infant from ancient Egypt (Lowenstein, 2009; Ramachandran, Jones, 2018). Also exist a story about Ivar The Boneless was a Viking leader who invaded Anglo-Saxon England. According to the Tale of Ragnar Lodbrok, Ivar's bonelessnes was the result of a curse. He was born with weak bones. While the sagas describe Ivar's physical disability, they also emphase his wisdom, cunning, and mastery of strategy and tactics in battle (Mahoney, 2017). Different publications of brittle bones and hearing loss studies have been appearing in medical literature since 1600. J.F. Lobstein и W. Vrolik were one of the first people, who could get the etiology of osteogenesis imperfecta. In 1825 J.F. Lobstein got some information about 3 sick children of different age. They had fractures of tubular bones without any reason. Author decided to name this disease as "ostepsathyrosys" and in his treatise on pathological anatomy devoted an entire chapter of it.

In 1849 W. Vrolik described "Osteogenesis Imperfecta" as syndrome of brittle bone with a lot of fractures which happened in prenatal period or immediately after birth. Searching the literature, we can see how gradually congenital bone fragility stood out from the concept of rickets. Since 1900 the authors began to point out the genetic nature of osteogenesis imperfecta.

J. Spurway in 1896 reported the first instance in which the condition of fragile bones (fragilitas ossium) was associated

Table 1. Expanded F.H. Glorieux classification of osteogenesis imperfecta

Type	Clinical severity	Typical features	Inheritance	Associated mutations
I	Mild nondeforming	Normal height or mild short stature; blue sclera	AD	*COL1A1, COL1A2*
II	Perinatal lethal	Multiple rib and long bone fractures at birth	AD Rarely AR	*COL1A1, COL1A2*
III	Severely deforming	Very short; triangular face; grayish sclera	AD	*COL1A1, COL1A2*
IV	Moderately deforming	Moderately short; grayish or white sclera	AD	*COL1A1, COL1A2*
V	Moderately deforming	Hyperplastic callus; white sclera	AD	Unknown
VI	Moderately to severely deforming	White sclera	AR	Unknown
VII	Moderately deforming	White sclera	AR	*CRTAP*
VIII	Severely deforming, perinatal lethal	White sclera	AR	*LEPRE1*

Note. AD – autosomal dominance; AR – autosomal recessivity; *COL1A1* and *COL1A2* – genes encoding type I collagen; *CRTAP* – cartilage-associated protein; *LEPRE1*, also known as *P3H1*, prolyl-3-hydroxylase 1.

with blue scleras. E. Bronson in 1917 and J. Hoeve and A. Kleyn in 1918 added to syndrome the third feature, deafness. J.A. Key in 1926 reffered to the syndrome as "hereditary hypoplasia of the mesenchyme" and called attention to the hypotonicity of the ligaments with hypermobile joints. The first classification of osteogenesis imperfecta was made by E. Looser, in 1906 who divided the condition on two forms, osteogenesis imperfecta congenita (also known as Vrolik disease) and osteogenesis imperfecta tarda (also known as Ekman-Lobstein disease) to distinguish the early and late forms of the disease.

In the 1970s, Dr. D. Sillence and his team of researchers in Australia developed the system of categorization using "Types" that is currently in use. His original four classifications (Type I, Type II, Type III and Type IV) combine clinical symptoms with genetic components. This listing is based on the number of people in the study who had similar symptoms. The types do not go from mildest to most severe. This classification system has been generally accepted world wide since 1979 OI continues to evolve as new information is discovered (Yakhyayeva et al., 2015b). Later this classification was supplemented by M. Ramachandran et al. (Pigarova et al., 2017), which also took into account the violation of dentinogenesis, OI the IV type of OI was subdivided into subtype B, which is accompanied by defects of dentinogenesis, and subtype A, which does not have these violations.

In 2000, F.H. Glorieux presented a classification of osteogenesis imperfecta, in which, in addition to the already known types, four more types of OI (V, VI, VII, VIII) were identified that are not associated with the pathology of type I collagen. In this classification, modern advances in the field of molecular genetic studies of the disease were taken into account (Table 1).

In less than 5 % of patients diagnosed with OI, type V occurs, which is inherited in an autosomal dominant type. The clinical phenotype of OI type V differs from other types of OI and is characterized by calcification of the interosseous membrane of the forearm and the formation of hyperplastic callus. OI–V has a wide spectrum of disease severity.

Type VI OI is clinically similar to types II and IV but has different characteristic histological picture – forming osteoid due to a violation of mineralization (Glorieux et al., 2002).

Type VII is manifested by deformations of long bones, shortening of the proximal limbs, coxa vara (varus deformities of the femoral neck), accompanied by normal dentinogenesis and the usual color of the sclera. It is characterized by an autosomal recessive type of inheritance. VII type OI is caused by a gene mutation in the chromosome 3p22-24.1, which encodes a protein associated with cartilage (*CRTAP*). *CRTAP* is a co-factor for post-translational modification of type I collagen. The severity of the disease depends on the degree of *CRTAP* deficiency. In the complete absence of *CRTAP* protein, prenatal death occurs, or the baby is born with severe OI (Ward et al., 2002).

Type VIII– a severe type of the desease, clinically similar to type II of OI, characterized by an autosomal recessive type of inheritance, associated with the mutation in *LEPRE1*. Diagnosed at perinatal age. Severe bone deformities, white sclera, are characteristic, accompanied by normal dentinogenesis (Fratzl-Zelman et al., 2016).

Types I–V are predominantly autosomal dominant inheritance, VI–XVIII are autosomal recessive. When new genes were discovered, the classification expanded, and by 2015, the number of forms of the disease reached 18.

Osteogenesis imperfecta type I is characterized by the presence of a defect in the *COL1A1* gene, which leads to a decrease in the amount of type I collagen produced; in types II–IV, due to mutations in the *COL1A1* and *COL1A2* genes, type V is due to mutations in the *IFITM5* gene and dysregulation of bone mineralization, type VI occurs due to a mutation in the *SERPINF1* gene, which leads to a defect in bone mineralization; types VII (*CRTAP* gene), VIII (*LEPRE1*

Table 2. Modern classification of osteogenesis imperfecta

Type	Type name	Gene	Inheritance type
I	Non-deforming type with blue sclera	*COL1A1*, *COL1A2*, *SP7*, *BMP1*, *P3H1*, *PLS3*	AD, X-linked
II	Perinatally fatal, severe	*COL1A1*, *COL1A2*, *CRTAP*, *P3H1*, *CREB3L1*, *PPIB*, *BMP1*	AD, AR
III	Progressively deforming, moderately severe	*COL1A1*, *COL1A2*, *BMP1*, *CRTAP*, *FKBP10*, *P3H1*, *PLOD2*, *PPIB*, *SERPINF1*, *SERPINH1*, *TMEM38B*, *WNT1*, *CREB3L1*, *FAM46A*	AD, AR
IV	Variable OI with blue sclera, medium heavy	*COL1A1*, *COL1A2*, *WNT1*, *CRTAP*, *PPIB*, *SP7*, *PLS3*, *TMEM38B*, *FKBP10*, *SPARC*	AD, AR, X-linked
V	Moderate OI with ossification of the interosseous membrane of the forearm	*IFITM5*	AD

Note. AD is an autosomal dominant type of inheritance; AR is an autosomal recessive type of inheritance.

gene, also known as *P3H1*) and IX (*PPIB* gene) are the result of a defect in the collagen 3-hydroxylation process. The cause of osteogenesis imperfecta of the X and XI types is a violation of the processing and cross-linking of collagen due to mutations in the *SERPINH1* and *FKBP10* genes, respectively. Mutations in the *PLOD2* and *BMP1* genes lead to incomplete type XII osteogenesis. These genes are involved in post-translational modification, processing, folding, secretion, and crosslinking of type I procollagen. Types XIII–XVIII of osteogenesis imperfecta are characterized by a violation of the differentiation of osteoblasts: mutations in the *SP7* gene lead to the manifestation of the XIII type, in the *TMEM38B* gene – the XIV type, in *WNT1* – the XV type, in *CREB3L1* – the XVI type, in *SPARC* – the XVII type, in *MBTPS2* – XVIII type (Marini et al., 2017).

The classification of the disease, taking into account the molecular pathogenesis of the disease, complicated the work of clinical doctors and in 2016 the International committee of nomenclature of constitutional disorders of the skeleton, INCDS) reduced the classification to 5 forms, retaining 4 types, which were originally described by silence and adding a 5th type. In total, 5 groups of the disease were identified using the Arabic digital system, which indicates the unifying phenotypic characteristics, and individual (characteristic for a particular type) changes still retained their original Roman designation (Table 2) (Ignatovich et al., 2018). This characteristic leaves room for the inclusion of new genes found as the cause of osteogenesis imperfecta until the degree of heterogeneity of the disease is identified.

Thus, the classification of osteogenesis imperfecta has undergone a number of fundamental changes associated with advances in the study of the molecular pathogenesis of the disease. The degree of heterogeneity of the disease has not yet been determined, the incidence of *de novo* cases has not been estimated, and therefore, it will probably continue to improve the classification criteria for the diagnosis of osteogenesis imperfecta.

Modern views on the etiology and pathogenesis of osteogenesis imperfecta

Osteogenesis imperfecta is characterized by wide clinical and genetic heterogeneity; earlier, the disease was referred to as collagenopathies, because in most cases, the structure and function of the main protein of bone tissue – type I collagen, as well as its stability are disturbed. Later, in patients with osteogenesis imperfecta, mutations were revealed in genes that do not participate in the formation of collagen structure and folding (Tournis, Dede, 2017).

To date, 20 genes responsible for the development of OI have been identified. The autosomal dominant type of ND inheritance in most cases is caused by defects in the *COLIA1* or *COLIA2* genes of type I collagen chains encoding α1 (I) and α2 (I) peptide type I collagen chains, respectively (Ignatovich et al., 2018). Autosomal dominant disease inheritance options have also been described in several patients with mutations in the *IFITM5* (MIM: 614757) and *P4HB* (MIM: 176790) genes.

P4HB encodes the beta subunit of prolyl 4-hydroxylase, which is involved in prolyl hydroxylation and folding of procollagen (Li et al., 2019), and *IFITM5* is a gene specific for osteoblasts associated with matrix mineralization (Glorieux et al., 2000). The *IFITM5* gene is located on chromosome 11 (p15.5) in a cluster of related genes (*IFITM1*, *2*, *3*, and *10*) and belongs to the family of genes encoding proteins containing two transmembrane domains that perform various significant cellular functions (Yakhyayeva et al., 2014).

Osteogenesis imperfecta is also transmitted in an autosomal recessive manner of inheritance, which is caused by mutations in the following genes: *BMP1* (MIM: 112264) (Asharani et al., 2012), *CRTAP* (MIM: 605497) (Morello et al., 2006), *FKBP10* (MIM: 607063) (Barnes et al., 2012), *P3H1* (MIM: 610339) (Cabral et al., 2007), *PLOD2* (MIM: 601865) (Puig-Hervás et al., 2012), *PPIB* (MIM: 123841) (VanDijk et al., 2009), *SEC24D* (MIM: 607186) (Zhang et al., 2017), *SERPINH1* (MIM: 600943) (Christiansen et al., 2010) and *TMEM38B* (MIM: 611236) (Rubinato et al., 2014), which are involved in post-translational modifications, processing, coagulation, secretion and cross-linking of procollagen (I). However, there is another group of OI loci with AR type of inheritance, which are not recognized as directly involved in the biosynthesis of type I collagen, but play a role in the mineralization or development of osteoblasts. This second group of genes includes *CREB3L1* (MIM: 616215) (Symoens et al., 2013), *SERPINF1* (MIM: 172860) (Becker et al., 2011), *SP7* (MIM: 606633)

(Lapunzina et al., 2010), *SPARC* (MIM: 182120) (Mendoza-Londono et al., 2015) and *WNT1* (MIM: 164820) (Laine et al., 2013; Pyott et al., 2013). Finally, mutations in the *PLS3* (MIM: 300131) genes (Costantini et al., 2018) and *MBTPS2* (MIM: 300294) were associated with two different forms of X-linked forms of OI.

It is known that there are two genes that encode proteins that are part of the metabolic chain that regulate intramembrane proteolysis (RIP) in osteoblasts, leading to the formation of the phenotype of osteogenesis imperfecta. During the intramembrane proteolysis, endopeptidases S1P (encoded by the *MBTPS1* gene) and S2P (encoded by the *MBTPS2* gene) in the Golgi membrane sequentially cleave regulatory proteins transported from the endoplasmic reticulum during stress endoplasmic reticulum or sterol metabolite deficiency. In patients with mutations in the *MBTPS2* gene, lysine hydroxylation of the α1 (I) chain and α2 (I) chain is reduced, collagen crosslinking is altered, and bone tissue strength is impaired. One of the transcription factors activated by RIP is a specific astrocyte-induced substance (OASIS; encoded by *CREB3L1*). A deficiency of this substance has been reported in association with a family with severe osteogenesis imperfecta. OASIS is a stress transducer of the endoplasmic reticulum, which regulates the transcription of genes involved in the development, differentiation and maturation of osteoblasts. In mice with the knocked out *CREB3L1* gene, severe osteopenia was observed with spontaneous fractures and a decrease in the production of type I collagen in the bone (Lindert et al., 2016).

In 2018, another gene was discovered – *FAM46*, which also leads to osteogenesis imperfecta with an autosomal recessive type of inheritance. FAM46A is a member of the superfamily of nucleotidyl transferase folded proteins, but its exact function is currently unknown. However, there is some evidence pointing to the corresponding role of *FAM46A* in bone development. Using RT-PCR analysis, specific *FAM46A* expression was detected in human osteoblasts and, interestingly, a nonsense mutation in *FAM46A* was recently discovered in a mouse model derived from ENU (N-ethyl-N-nitrosourea), characterized by a decrease in body length, limbs, deformation of the ribs, pelvis and skull and a decrease in the thickness of the cortex in long bones (Doyard et al., 2018) (Table 3).

About 90 % of the 3,000 people from the incomplete osteogenesis database (http://www.le.ac.uk/ge/collagen/) have changes in either the *COL1A1* gene or *COL1A2*, and the remaining 10 % show homozygous or heterozygous mutations in other genes involved in the pathogenesis of OI. However, major sequencing centers that offer a panel of causal mutations associated with incomplete osteogenesis identify a lower frequency of structural mutations in the *COL1A1* and *COL1A2* genes in patients with a moderate to severe clinical presentation of the disease. For example, heterozygous mutations in the *COL1A1* or *COL1A2* genes were identified in 77 % of 598 patients with ND from the Shreiners Clinic (Montreal, Canada), 9 % had one mutation in the *IFITM5* gene, and the rest had homozygous or heterozygous mutations in other genes causing incomplete osteogenesis. Lethal mutations in the collagen gene could be lost in this study. In populations with a high level of blood relationship, the frequency of incomplete osteogenesis is higher, for example, among African Americans in the United States of America, the frequency of the mutant variant in the *P3H1* gene (previously called *LEPRE1* encoding shedding 3-hydroxylase 1) is about 1 in 240 people. Homozygosity for this so-called West African allele accounts for 25 % of all cases of incomplete lethal osteogenesis in this population, which may be clinically erroneously classified as type II OI. Among the West Africans of Ghana and Nigeria, the frequency of occurrence of this allele is 1.5 %, which can lead to a frequency of lethal recessive incomplete osteogenesis equal to the frequency of *de novo* mutations in type I collagen.

Despite the large number of mutations recorded in the database on incomplete osteogenesis (https://oi.gene.le.ac.uk), each population has its own spectrum consisting of a small number of mutations, with each researcher finding previously undescribed in mutation literature.

As in the case of other recessive diseases, in some populations there are isolated cases of mutations in rare genes that are not found in other populations: an exon deletion in the *TMEM38B* gene was found in a family from Saudi Arabia; reading frame offset in the *FKBP10* gene was found in patients from Turkey; missense mutations in the *WNT1* gene in the Hmong ethnic group from Vietnam and China (Marini et al., 2017). Among the population of northern Ontario (Canada), the intron variant destabilizes the mRNA of the *CRTAP* gene (which encodes a protein associated with cartilage) and develops the phenotype of incomplete osteogenesis type VII.

The clinical picture of OI and the severity of the disease are diverse, they can be manifested by lethal variants, obvious abnormalities of the skeleton in children, or have an easy manifestation in people of mature age. The severity of the disease is determined by the frequency of fractures, progressive deformity, chronic bone pain and loss of mobility. Due to the clinical heterogeneity of the disease, there are difficulties in diagnosing and verifying the diagnosis; in children with OI, a delay in physical development, scoliosis, progressive deformations of long bones, hearing loss, pathology of teething are revealed, therefore only the identification of the molecular cause of the disease allows an accurate diagnosis to be established.

Thus, significant progress has been made in the study of the molecular pathogenesis of incomplete osteogenesis, but the degree of heterogeneity of the disease remains to be determined. With the development of genotyping technologies and the widespread adoption of deep sequencing and full-exomic sequencing methods, it has become possible to identify not only new mutations in known genes, but also to identify new genes involved in the development of the disease.

Prospects for the treatment of incomplete osteogenesis

Currently, active research is being conducted on the possibilities of targeted therapy for patients with hereditary diseases, taking into account the molecular defect. Encouraging results were obtained with the pathogenetic treatment of cystic fibrosis.

Bisphosphonates (BP) are the main drug for the treatment of both children and adult patients with OI. It is believed that

Table 3. Characterization of genes and their protein products responsible for the development of OI

Genes and their protein products	Localization	Type OI	Inheritance type	Function	Exons	Mutations
COL1A1 – collagen α1(I)	17q21.33	I, II, III, IV	Autosomal dominant	Part of collagen type I	52	1035
COL1A2 – collagen α2(I)	27q21.3	I, II, III, IV	»	Part of collagen type I	52	604
CRTAP – cartilage-associated protein	3p22.3	III, IV	Autosomal recessive	Participates in post-translational modification of collagen I	7	32
FKBP10 – 65kDa FK506-binding protein	17q21.2	III, IV	»	Serves as collagen chaperones	11	39
IFITM5 – bone-restricted interferon-induced transmembrane protein-like protein (BRIL; also known as IFM5)	11p15.5	V	Autosomal dominant	Expressed in skeletal tissue and is involved in bone formation	2	2
P3H1 – prolyl-3-hydroxylase 1	1p34.2	III	Autosomal recessive	Participates in post-translational modification of collagen I	16	69
SP7 – transcription factor *SP7* (also known as osterix)	12q13.13	III	»	Involved in the regulation of bone cell differentiation	5	2
TMEM38B – trimeric intracellular cation channel type B (TRIC B; also known as TM38B)	9q31.2	IV	»	Involved in the transfer of divalent Ca	6	6
WNT1 – proto-oncogene Wnt-1 (WNT1)	12q13.12	IV	»	Involved in the functioning of osteoblasts and bone development	4	36
BMP1 – bone morphogenetic protein 1	8p21.3	I, III, IV	»	Participates in the C-terminal processing of both procollagen protein chains	20	11
PPIB – peptidyl-prolyl cis–trans isomerase B (PPIase B)	15q22.31	III	»	Participates in post-translational modification of type I collagen	5	17
SERPINF1 – pigment epithelium-derived factor (PEDF)	17p13.3	III, IV	»	Involved in bone mineralization	9	38
SERPINH1 – serpin H1 (also known as HSP47)	11q13.5	III, IV	»	Is a collagen chaperone	7	9
PLS3 (plastin 3) – plastin 3	Xq23	I	X-linked type	The molecular function of plate-3 is not fully understood, may play a role in the differentiation of bone cells	21	11
CREB3L1 – old astrocyte specifically induced substance (OASIS; also known as CR3L1)	11p11.2	II	Autosomal recessive	Regulates the formation of type I collagen in the process of bone formation	13	4
P4HB – prolyl 4-hydroxylase, subunit beta	17q25.3	III	Autosomal dominant	Participates in the hydroxylation of lysine residues in collagen fibers; catalyzes the hydroxylation of proline residues in X-Pro-Gly repeats in the spiral domain of procollagen	11	2
PLOD2 – lysyl hydroxylase 2 (LH2)	3q24	III, IV	Autosomal recessive	Participates in the hydroxylation of lysine residues in collagen fibers	23	10
SEC24D – SEC24 homolog D, COPII coat complex component	4q26	III, IV	»	Function is not fully understood	25	7
SPARC – osteonectin	5q33.1	IV	»	Regulates the proliferation and interaction of cells and matrix by binding calcium ions with hydroxyapatite	10	2
FAM46A – family of similar sequences 46A	6q14.1	III	»	Function is not fully understood	3	3

BP can be less effective or even lead to adverse consequences in cases of insufficient calcium intake and/or vitamin D deficiency (Weaver et al., 2016).

There are also preclinical and a small number of clinical studies in adult patients with OI regarding denosumab, a monoclonal antibody targeted at RANKL (receptor activator of the nuclear factor kappa-B ligand). Regarding anabolic therapy, teriparatide, currently the only available anabolic agent, has shown promising results in adult patients with type I OI. Preclinical studies show that inhibition of TGF-beta signaling, as well as inhibition of sclerostin, can also play a role in treating bone fragility. In addition to pharmacological interventions, the multidisciplinary approach provided by experienced orthopedic surgeons, dental care specialists, physiotherapists and kinesitherapy specialists is of paramount importance for providing the best possible medical care.

Currently, bisphosphonates are widely used to treat children with OI. It has been shown that both oral (alendronate, risedronate) and intravenous administration of BP (pamidronate, zoledronate, neridronate) improve the level of BMD, especially in the spine. However, data from a randomized, placebo-controlled trial regarding fracture response, pain relief, and motor activity improvement are still missing. Recent studies have not found a consistent decrease in the frequency of fractures and an improvement in the clinical status of patients in the treatment of BP (Dwan et al., 2014).

Concerning the effect of BF in adults with OI, there is limited evidence that tested the effect of various BFs on BMD levels. Almost all studies reported a beneficial effect on the level of BMD of the lower spine (an increase of up to 13.9 %) with less pronounced effects on the total level of BMD of the thigh (an increase of up to 4.3 %) (Lindahl et al., 2014). Recently, a number of reports have been published about atypical hip fractures in adult patients with OI receiving treatment with BF.

A number of studies have evaluated the effect of denosumab in patients with OI caused by a mutation in SERPINF1, characterized by a weak response to BP, as well as in patients with OI I/IV (n = 8) and OI III (n = 2) types (Hoyer-Kuhn et al., 2016). The dose used was 1 mg/kg subcutaneously every 3 months. All studies reported a significant increase in BMD and the absence of significant side effects of treatment over a two-year period.

Sclerostin inhibition may be another treatment option for bone fragility in OI. Recently published studies have shown that administering romososumab (a sclerostin-binding monoclonal antibody) within one year reduces the incidence of spinal fractures and osteoporosis in postmenopausal women with osteoporosis (Sinder et al., 2015; Grafe et al., 2016).

In a mouse model with OI, it was shown that increased TGF-β signaling is involved in the OI phenotype, while inhibition of TGF-β improves bone mass and strength. Phase 1 of the study verifies the safety of fresolumumab, a high-affinity neutralizing antibody that targets all 3 TGF-β isoforms, in adults with a mild clinical presentation of OI. Combination therapy with antiresorptive and anabolic agents is another potential treatment option for bone fragility in patients with OI. Other treatments, such as bone marrow transplantation and gene therapy, are in the process of evaluating the effectiveness of treating severe forms of OI (Marini et al., 2017).

Thus, despite the progress made in understanding the pathophysiology of OI, additional research is still needed to determine the best therapeutic approach to this heterogeneous disease.

Conclusion

Summarizing the foregoing, we can conclude that there has been a breakthrough in the identification of the molecular pathogenesis of incomplete osteogenesis, which is due to the introduction of modern next-generation sequencing technologies (NGS). However, questions about the prevalence of the disease as a whole and its individual clinical forms in various world populations are still far from over. Also, a final determination of the degree of molecular heterogeneity of OI has not been achieved; the identification of new pathogenetic mechanisms of the formation of the phenotype of the disease continues on the basis of identifying new genes involved in the pathogenesis of OI. Currently, attempts are being made to develop targeted therapy for the disease, taking into account new knowledge about the clinical and genetic aspects of OI, but there are still many conflicting results and the solution to the problem of treating the disease is far from complete.

References

Asharani P.V., Keupp K., Semler O., Wang W., Li Y., Thiele H., Yigit G., Pohl E., Becker J., Frommolt P., Sonntag C., Altmüller J., Zimmermann K., Greenspan D.S., Akarsu N.A., Netzer C., Schönau E., Wirth R., Hammerschmidt M., Nürnberg P., Wollnik B., Carney T.J. Attenuated *BMP1* function compromises osteogenesis, leading to bone fragility in humans and zebrafish. Am. J. Hum. Genet. 2012;90(4):661-674. DOI 10.1016/j.ajhg.2012.02.026.

Barnes A.M., Cabral W.A., Weis M., Makareeva E., Merta E.L., Leikin S., Eyre D., Trujillo C., Marini J.C. Absence of *FKBP10* in recessive type XI OI leads to diminished collagen cross-linking and reduced collagen deposition in extracellular matrix. Hum. Mutat. 2012;33(11):1589-1598. DOI 10.1002/humu.22139.

Becker J., Semler O., Gilissen C., Li Y., Bolz H.J., Giunta C., Bergmann C., Rohrbach M., Koerber F., Zimmermann K. Exome sequencing identifies truncating mutations in human *SERPINF1* in autosomal-recessive osteogenesis imperfecta. Am. J. Hum. Genet. 2011;88(3):362-371. DOI 10.1016/j.ajhg.2011.01.015. Belgian Bone Club. 2019. Available at: http://www.bbcbonehealth.org/osteogenesis-imperfecta.

Cabral W.A., Chang W., Barnes A.M., Weis M.A., Scott M.A., Leikin S., Makareeva E., Kuznetsova N.V., Rosenbaum K.N., Tifft C.J., Bulas D.I., Kozma C., Smith P.A., Eyre D.R., Marini J.C. Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat. Genet. 2007;39(3):359-365. DOI 10.1038/ng1968.

Christiansen H.E., Schwarze U., Pyott S.M., AlSwaid A., Al Balwi M., Alrasheed S., Pepin M.G., Weis M.A., Eyre D.R., Byers P.H. Homozygosity for a missense mutation in *SERPINH1*, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. Am. J. Hum. Genet. 2010;86(3):389-398. DOI 10.1016/j.ajhg.2010.01.034.

Costantini A., Krallis P.N., Kämpe A., Karavitakis E.M., Taylan F., Mäkitie O., Doulgeraki A. A novel frameshift deletion in *PLS3* causing severe primary osteoporosis. J. Hum. Genet. 2018;63(8):923-926. DOI 10.1038/s10038-018-0472-5.

Doyard M., Bacrot S., Huber C., Di Rocco M. *FAM46A* mutations are responsible for autosomal recessive osteogenesis imperfecta. J. Med. Genet. 2018;55(4):278-284. DOI 10.1136/jmedgenet-2017-104999. Epub 2018. Jan 22.

Dwan K., Phillipi C.A., Steiner R.D. Bisphosphonate therapy for osteogenesis imperfecta. Cochrane Database Syst. Rev. 2014;7:Cd005088.

Fratzl-Zelman N., Barnes A.M., Weis M., Carter E., Hefferan T.E., Perino G., Chang W., Smith P.A., Roschger P., Klaushofer K., Glorieux F.H., Eyre D.R., Raggio C., Rauch F., Marini J.C. Non-lethal type VIII osteogenesis imperfecta has elevated bone matrix mineralization. J. Clin. Endocrinol. Metab. 2016;101(9):3516-3525. DOI 10.1210/jc.2016-1334.

Glorieux F.H., Rauch F., Plotkin H., Ward L., Travers R., Roughley P., Lalic L., Glorieux D.F., Fassier F., Bishop N.J. Type V osteogenesis imperfecta: a new form of brittle bone disease. J. Bone Miner. Res. 2000;15(9):1650-1658. DOI 10.1359/jbmr.2000.15.9.1650.

Glorieux F.H., Ward L.M., Rauch F., Lalic L., Roughley P.J., Travers R. Osteogenesis imperfecta type VI: a form of brittle bone disease with a mineralization defect. J. Bone Miner. Res. 2002;17(1):30-38. DOI 10.1359/jbmr.2002.17.1.30.

Grafe I., Alexander S., Yang T., Lietman C., Homan E.P., Munivez E., Chen Y., Jiang M.M., Bertin T., Dawson B., Asuncion F., Ke H.Z., Ominsky M.S., Lee B. Sclerostin antibody treatment improves the bone phenotype of *Crtap*$^{(-/-)}$ mice, a model of recessive Osteogenesis Imperfecta. J. Bone Miner. Res. 2016;31(5):1030-1040.

Hoyer-Kuhn H., Franklin J., Allo G., Kron M., Netzer C., Eysel P., Hero B., Schoenau E., Semler O. Safety and efficacy of denosumab in children with osteogenesis imperfect – a first prospective trial. J. Musculoskelet Neuronal. Interact. 2016;16(1):24-32.

Ignatovich O.N., Namazova-Baranova L.S., Margieva T.V., Yakhyaeva G.T., Zhurkova N.V., Savostyanov K.V., Pushkov A.A., Krotov I.A. Osteogenesis imperfecta: diagnostic feature. Pediatricheskaya Pharmacologiya = Pediatric Pharmacology. 2018;15(3): 224-232. DOI 10.15690/pf.v15i3.1902. (in Russian)

Kruchkova O.A., Kruglov S.V. Treatment of osteogenesis imperfecta. Symptoms of osteogenesis imperfecta. 2014. Available at: https://ymkababy.ru/pregnancy/nesovershennyi-osteogenez-lechenie-nesovershennyi-osteogenez-simptomy.html. (in Russian)

Laine C.M., Joeng K.S., Campeau P.M., Kiviranta R., Tarkkonen K., Grover M., Lu J.T., Pekkinen M., Wessman M., Heino T.J., Nieminen-Pihala V., Aronen M., Laine T., Kröger H., Cole W.G., Lehesjoki A.E., Nevarez L., Krakow D., Curry C.J., Cohn D.H., Gibbs R.A., Lee B.H., Mäkitie O. *WNT1* mutations in early-onset osteoporosis and osteogenesis imperfecta. N. Engl. J. Med. 2013; 368:1809-1816. DOI 10.1056/NEJMoa1215458.

Lapunzina P., Aglan M., Temtamy S., Caparrós-Martin J.A., Valencia M., Letón R., Martinez-Glez V., Elhossini R., Arm K., Vilaboa N., Ruiz-Perez V.L. Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. Am. J. Hum. Genet. 2010;87(1):110-114. DOI 10.1016/j.ajhg.2010.05.016.

Li L., Zhao D., Zheng W., Wang O., Jiang Y., Xia W., Xing X., Li M. A novel missense mutation in P4HB causes mild osteogenesis imperfecta. Biosci. Rep. 2019;39(4). DOI 10.1042/BSR20182118.

Lindahl K., Langdahl B., Ljunggren O., Kindmark A. Treatment of osteogenesis imperfecta in adults. Eur. J. Endocrinol. 2014;171(2): R79-R90. DOI 10.1530/EJE-14-0017.

Lindert U., Cabral W.A., Ausavarat S., Tongkobpetch S. *MBTPS2* mutations cause defective regulated intramembrane proteolysis in X-linked osteogenesis imperfecta. Nat. Commun. 2016;7:11920. DOI 10.1038/ncomms11920.

Lowenstein E.J. Osteogenesis imperfecta in a 3,000-year-old mummy. Childs Nerv. Syst. 2009;25(5):515-516. DOI 10.1007/s00381-009-0817-7.

Mahoney M. Ivar the Boneless. 2017. Available at: www.englishmonarchs.co.uk/vikings_10.html.

Marini J.C., Forlino A., Bächinger H.P., Bishop N.J., Byers P.H., De Paepe A., Fassier F., Fratzl-Zelman N., Kozloff K.M., Krakow D., Montpetit K., Semler O. Osteogenesis imperfecta. Nat. Rev. Dis. Primers. 2017;3:1-19. DOI 10.1038/nrdp.2017.52.

Mendoza-Londono R., Fahiminiya S., Majewski J. Care4Rare Canada Consortium; Tétreault M., Nadaf J., Kannu P., Sochett E., Howard A., Stimec J., Dupuis L., Roschger P., Klaushofer K., Palomo T., Ouellet J., Al-Jallad H., Mort J.S., Moffatt P., Boudko S., Bächinger H.P., Rauch F. Recessive osteogenesis imperfecta caused by missense mutations in *SPARC*. Am. J. Hum. Genet. 2015;96(6): 979-985. DOI 10.1016/j.ajhg.2015.04.021.

Morello R., Bertin T.K., Chen Y., Hicks J., Tonachini L., Monticone M., Castagnola P., Rauch F., Glorieux F.H., Vranka J., Bachinger H.P., Pace J.M., Schwarze U., Byers P.H., Weis M.A., Fernandes R.J., Eyre D.R., Yao Z., Boyce B.F., Lee B. *CRTAP* is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. Cell. 2006;127(2):291-304. DOI 10.1016/j.cell.2006.08.039.

Nadyrshina D.D., Khusainova R.I., Khusnutdinova E.K. Studies of type I collagen (COL1A1) α1 chain in patients with osteogenesis imperfecta. Russ. J. Genet. 2012;48(3):321-328.

Pigarova E.A., Sheremeta M.S., Kulikova K.S., Belovalova I.M., Tulpakov A.N., Rumiantsev P.O. Osteogenesis imperfecta in combination with Graves disease. Ozhirenie i Metabolism = Obesity and Metabolism. 2017;14(4):77-82. DOI 10.14341/OMET2017477-82. (in Russian)

Puig-Hervás M.T., Temtamy S., Aglan M., Valencia M., Martínez-Glez V., Ballesta-Martínez M.J., López-González V., Ashour A.M., Amr K., Pulido V., Guillén-Navarro E., Lapunzina P., Caparrós-Martín J.A., Ruiz-Perez V.L. Mutations in *PLOD2* cause autosomal-recessive connective tissue disorders within the Bruck syndrome-osteogenesis imperfecta phenotypic spectrum. Hum. Mutat. 2012;33(10):1444-1449. DOI 10.1002/humu.22133.

Pyott S.M., Tran T.T., Leistritz D.F., Pepin M.G., Mendelsohn N.J., Temme R.T., Fernandez B.A., Elsayed S.M., Elsobky E., Verma I., Nair S., Turner E.H., Smith J.D., Jarvik G.P., Byers P.H. *WNT1* mutations in families affected by moderately severe and progressive recessive osteogenesis imperfecta. Am. J. Hum. Genet. 2013;92(4): 590-597. DOI 10.1016/j.ajhg.2013.02.009.

Ramachandran M., Jones D. Osteogenesis imperfecta. 2018. Available at: https://emedicine.medscape.com/article/1256726-overview.

Rubinato E., Morgan A., D'Eustacchio A., Pecile V., Gortani G., Gasparini P. A novel deletion mutation involving *TMEM38B* in a patient with autosomal recessive osteogenesis imperfecta. Gene. 2014; 545(2):290-292. DOI 10.1016/j.gene.2014.05.028.

Sinder B.P., Salemi J.D., Ominsky M.S., Caird M.S., Marini J.C., Kozloff K.M. Rapidly growing Brtl/+ mouse model of osteogenesis imperfecta improves bone mass and strength with sclerostin antibody treatment. Bone. 2015;71:115-123.

Symoens S., Malfait F., D'hondt S., Callewaert B., Dheedene A., Steyaert W. Deficiency for the ER-stress transducer OASIS causes severe recessive osteogenesis imperfecta in humans. Orphanet J. Rare Dis. 2013;8:154. DOI 10.1186/1750-1172-8-154.

Tournis S., Dede A.D. Osteogenesis imperfecta – a clinical update. Metabolism. 2018;80:27-37. DOI 10.1016/j.metabol.2017.06.001.

VanDijk F.S., Nesbitt I.M., Zwikstra E.H., Nikkels P.G.J., Piersma S.R., Fratantoni S.A., Jimenez C.R., Huizer M., Morsman A.C., Cobben J.M., van Roij M.H.H., Elting M.W., Verbeke M.I.J.L., Wijnaendts L.C.D., Shaw N.J., Högler W., McKeown C., Sistermans E.A., Dalton A., Meijers-Jeijboer H., Pals G. *PPIB* mutations cause severe osteogenesis imperfecta. Am. J. Hum. Genet. 2009; 85(4):521-527. DOI 10.1016/j.ajhg.2009.09.001.

Ward L.M., Rauch F., Travers R., Chabot G., Azouz E.M., Lalic L., Roughley P.J., Glorieux F.H. Osteogenesis imperfecta type VII: an autosomal recessive form of brittle bone disease. Bone. 2002;31(1): 12-18.

Weaver C.M., Alexander D.D., Boushey C.J., Dawson-Hughes B., Lappe J.M., LeBoff M.S., Liu S., Looker A.C., Wallace T.C., Wang D.D. Calcium plus vitamin D supplementation and risk of fractures: an updated meta-analysis from the National Osteoporosis Foundation. Osteoporos. Int. 2016;27:367-376.

Yakhyayeva G.T., Margieva T.V., Namazova-Baranova L.S., Savostyanov K.V., Pushkov A.A., Zhurkova N.V., Zherdev K.V., Vashakmadze N.D., Gevorkyan A.K. Clinical case of rare type V osteogenesis imperfecta. Pediatricheskaya Pharmakologiya = Pediatric Pharmacology. 2015a;12(1):79-84. (in Russian)

Yakhyayeva G.T., Namazova-Baranova L.S., Margieva T.V. New aspects of the genetic basis, classification and treatment of osteogenesis imperfecta: a literature review. Pediatricheskaya Pharmakologiya = Pediatric Pharmacology. 2015b;12(5):579-588. DOI 10.15690/pf.v12i5.1461. (in Russian)

Zhang H., Yue H., Wang C., Gu J., He J., Fu W., Hu W., Zhang Z. Novel mutations in the *SEC24D* gene in Chinese families with autosomal recessive osteogenesis imperfecta. Osteoporos. Int. 2017; 28(4):1473-1480. DOI 10.1007/s00198-0163866-2.

ORCID ID
R.I. Khusainova orcid.org/0000-0002-8643-850X

Acknowledgements. This research was supported by the Russian Foundation for Basic Research grant No. 19-015-00489_a.
Conflict of interest. The authors declare no conflict of interest.
Received June 11, 2019. Revised August 6, 2019. Accepted October 17, 2019.

Прием статей через электронную редакцию на сайте http://vavilov.elpub.ru/index.php/jour
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции»/"Vavilov Journal of Genetics and Breeding"
до 2011 г. выходил под названием «Информационный вестник ВОГиС»/
"The Herald of Vavilov Society for Geneticists and Breeding Scientists".

Регистрационное свидетельство ПИ № ФС77-45870 выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 20 июля 2011 г.

«Вавиловский журнал генетики и селекции» включен ВАК Минобрнауки России в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, Российский индекс научного цитирования, ВИНИТИ, базы данных Emerging Sources Citation Index (Web of Science), Zoological Record (Web of Science), Scopus, Ebsco, DOAJ, Ulrich's Periodicals Directory, Google Scholar, Russian Science Citation Index на платформе Web of Science, каталог научных ресурсов открытого доступа ROAD.

Открытый доступ к полным текстам:
на сайте ИЦиГ СО РАН – bionet.nsc.ru/vogis/
платформе Elpub – vavilov.elpub.ru/index.php/jour
платформе Научной электронной библиотеки – elibrary.ru/title_about.asp?id=32440

Подписку на «Вавиловский журнал генетики и селекции» можно оформить в любом почтовом отделении России. Индекс издания 42153 по каталогу «Пресса России».

✉ e-mail: vavilov_journal@bionet.nsc.ru

Издатель: Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»,
проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Секретарь по организационным вопросам С.В. Зубова. Тел.: (383)3634977.
Издание подготовлено информационно-издательским отделом ИЦиГ СО РАН. Тел.: (383)3634963*5218.
Начальник отдела: Т.Ф. Чалкова. Редакторы: В.Д. Ахметова, И.Ю. Ануфриева. Дизайн: А.В. Харкевич.
Компьютерная графика и верстка: Т.Б. Коняхина, О.Н. Савватеева.

Подписано в печать 10.04.2020. Выход в свет 30.04.2020. Формат 60 × 84 $^{1}/_{8}$. Усл. печ. л. 13.72.
Уч.-изд. л. 16.2. Тираж 150 экз. (1-й завод 1–50 экз.). Заказ № 310. Цена свободная.

Отпечатано в типографии ФГУП «Издательство СО РАН», Морской проспект, 2, Новосибирск, 630090.