2023 • 27 • 4

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/VJGB-23-35

Учредители

Сибирское отделение Российской академии наук

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

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

2023 • 27 • 4

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/VJGB-23-35

Founders

Siberian Branch of the Russian Academy of Sciences

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

The Vavilov Society of Geneticists and Breeders

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

СОДЕРЖАНИЕ • 2023 • 27 • 4

Генетика и селекция растений

297 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Генетический полиморфизм локусов высокомолекулярных субъединиц глютенина у сортообразцов мягкой пшеницы Предуральской степной зоны.
А.А. Галимова, А.Р. Кулуев, К.Р. Исмагилов, Б.Р. Кулуев

306 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Морфологические и филогенетические особенности крымской популяции *Juniperus deltoides* R.P. Adams.
А.О. Лантушенко, О.О. Коренькова, А.А. Сыровец, Я.В. Мегер, П.А. Кореньков, О.М. Шевчук

316 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Генетическая структура популяции дикорастущих форм винограда заповедника «Утриш».
Е.Т. Ильницкая, М.В. Макаркина, И.В. Горбунов, И.В. Степанов, Т.Д. Козина, Е.А. Кожевников, В.К. Котляр

323 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Создание и изучение гибридов полба (*Triticum dicoccum*) × тритикале.
О.Г. Силкова, Ю.Н. Иванова, П.И. Стёпочкин

Генетика и селекция животных

333 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Изучение половых различий ответов на введение фактора роста фибробластов 21 у мышей с ожирением, вызванным потреблением сладко-жирной диеты.
Н.М. Бажан, Т.В. Яковлева, А.Ю. Казанцева, Н.Е. Костина, П.Е. Орлов, Н.В. Балыбина, К.О. Баранов, Е.Н. Макарова (на англ. языке)

342 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Эффекты центрального введения Tau-белка человека на экспрессию генов *Bdnf*, *Trkb*, *p75*, *Mapt*, *Bax* и *Bcl-2* в мозге мышей. *А.С. Орешко, А.Я. Родный, Д.В. Базовкина, В.С. Науменко*

349 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Сравнение эволюционных паттернов ДНК-повторов у представителей древних и молодых букетов видов из озера Байкал.
Ван Юйсян, Т.Е. Перетолчина, Е.В. Романова, Д.Ю. Щербаков (на англ. языке)

357 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Композиция тела как индикатор метаболических изменений у мышей, полученных путем оплодотворения *in vitro*.
М.В. Анисимова, Л. Гон, Г.В. Концевая, А.В. Ромащенко, Н.В. Хоцкин, А.К. Станова, Л.А. Герлинская, М.П. Мошкин

366 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Генетические методы в селекции медоносной пчелы. *М.Д. Каскинова, А.М. Салихова, Л.Р. Гайфуллина, Е.С. Салтыкова*

Генетика человека

373 ОБЗОР
Роль Берингии в адаптации человека к условиям Арктики по результатам геномных исследований современного и древнего населения. *Б.А. Малярчук*

383 ОБЗОР
Клинико-генетические аспекты и молекулярный патогенез остеопетроза.
Д.Д. Надыршина, Р.И. Хусаинова

Медицинская генетика

393 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Генетические маркеры бронхиальной астмы у детей: предрасположенность к вариантам течения заболевания.
М.В. Смольникова, Э.В. Каспаров, М.А. Малинчик, К.В. Копылова

401 ОБЗОР
Эпигенетическая регуляция ремоделирования костной ткани и ее роль в патогенезе первичного остеопороза. *Б.И. Ялаев, Р.И. Хусаинова*

Популяционная генетика

411 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Таксономическая и экофизиологическая характеристика актинобактерий почв сухостепной зоны Селенгинского среднегорья (Западное Забайкалье).
Е.П. Никитина, Л.Б. Буянтуева, Е.Ю. Абидуева, Ч.-Х. Сун

Актуальные методы биотехнологии

421 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Влияние коэкспрессии тиоредоксина и прохимозина на рефолдинг рекомбинантного химозина альпака.
С.В. Беленькая, Д.Н. Щербаков, А.И. Шаповал, Т.И. Есина, В.В. Ельчанинов (на англ. языке)

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2023 • 27 • 4

Plant genetics and breeding

297 ORIGINAL ARTICLE
Genetic polymorphism of high-molecular-weight glutenin subunit loci in bread wheat varieties in the Pre-Ural steppe zone.
A.A. Galimova, A.R. Kuluev, K.R. Ismagilov, B.R. Kuluev

306 ORIGINAL ARTICLE
Morphological and phylogenetic features of the Crimean population of *Juniperus deltoides* R.P. Adams. *A.O. Lantushenko, O.O. Korenkova, A.A. Syrovets, Ya.V. Meger, P.A. Korenkov, O.M. Shevchuk*

316 ORIGINAL ARTICLE
Genetic structure of the population of wild-growing vines of the Utrish Nature Reserve. *E.T. Ilnitskaya, M.V. Makarkina, I.V. Gorbunov, I.V. Stepanov, T.D. Kozina, E.A. Kozhevnikov, V.K. Kotlyar*

323 ORIGINAL ARTICLE
Creation and study of emmer (*Triticum dicoccum*) × triticale hybrids.
O.G. Silkova, Y.N. Ivanova, P.I. Stepochkin

Animal genetics and breeding

333 ORIGINAL ARTICLE
Studying sex differences in responses to fibroblast growth factor 21 administration in obese mice consuming a sweet-fat diet.
N.M. Bazhan, T.V. Jakovleva, A.Yu. Kazantseva, N.E. Kostina, P.E. Orlov, N.Yu. Balybina, K.O. Baranov, E.N. Makarova

342 ORIGINAL ARTICLE
Effects of central administration of the human Tau protein on the *Bdnf*, *Trkb*, *p75*, *Mapt*, *Bax* and *Bcl-2* genes expression in the mouse brain. *A.S. Oreshko, A.Ya. Rodnyy, D.V. Bazovkina, V.S. Naumenko*

349 ORIGINAL ARTICLE
Comparison of the evolutionary patterns of DNA repeats in ancient and young invertebrate species flocks of Lake Baikal.
Wang Yuxiang, T.E. Peretolchina, E.V. Romanova, D.Y. Sherbakov

357 ORIGINAL ARTICLE
Body composition as an indicator of metabolic changes in mice obtained by *in vitro* fertilization. *M.V. Anisimova, Yanli Gon, G.V. Kontsevaya, A.V. Romashchenko, N.V. Khotskin, A.K. Stanova, L.A. Gerlinskaya, M.P. Moshkin*

366 ORIGINAL ARTICLE
Genetic methods in honey bee breeding.
M.D. Kaskinova, A.M. Salikhova, L.R. Gaifullina, E.S. Saltykova

Human genetics

373 REVIEW
The role of Beringia in human adaptation to Arctic conditions based on results of genomic studies of modern and ancient populations. *B.A. Malyarchuk*

383 REVIEW
Clinical, genetic aspects and molecular pathogenesis of osteopetrosis.
D.D. Nadyrshina, R.I. Khusainova

Medical genetics

393 ORIGINAL ARTICLE
Genetic markers of children asthma: predisposition to disease course variants.
M.V. Smolnikova, Ed.W. Kasparov, M.A. Malinchik, K.V. Kopylova

401 REVIEW
Epigenetic regulation of bone remodeling and its role in the pathogenesis of primary osteoporosis. *B.I. Yalaev, R.I. Khusainova*

Population genetics

411 ORIGINAL ARTICLE
Taxonomic and ecophysiological characteristics of actinobacteria in soils of the dry steppe zone of the Selenga Highlands (Western Transbaikalia).
E.P. Nikitina, L.B. Buyantueva, E.Yu. Abidueva, C.H. Sun

Current biotechnological methods

421 ORIGINAL ARTICLE
The effect of thioredoxin and prochymosin coexpression on the refolding of recombinant alpaca chymosin.
S.V. Belenkaya, D.N. Shcherbakov, A.I. Chapoval, T.I. Esina, V.V. Elchaninov

Вавиловский журнал генетики и селекции. 2023;27(4):297-305
DOI 10.18699/VJGB-23-36

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

Genetic polymorphism of high-molecular-weight glutenin subunit loci in bread wheat varieties in the Pre-Ural steppe zone

A.A. Galimova[1, 2] ✉, A.R. Kuluev[1], K.R. Ismagilov[3], B.R. Kuluev[1, 2]

[1] Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia
[2] Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
[3] Bashkir Research Institute of Agriculture of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia
✉ aiz.galimova@yandex.ru

Abstract. High-molecular-weight glutenins play an important role in providing high baking qualities of bread wheat grain. However, breeding bread wheat for this trait is very laborious and, therefore, the genotyping of variety samples according to the allelic composition of high-molecular-weight glutenin genes is of great interest. The aim of the study was to determine the composition of high-molecular-weight glutenin subunits based on the identification of the allelic composition of the *Glu-1* genes, as well as to identify the frequency of the *Glu-1* alleles in bread wheat cultivars that are in breeding work under the conditions of the Pre-Ural steppe zone (PSZ). We analyzed 26 winter and 22 spring bread wheat varieties from the PSZ and 27 winter and 20 spring varieties from the VIR collection. Genotyping at the *Glu-A1* locus showed that the Ax1 subunits are most common in winter varieties, while the predominance of the Ax2* subunits was typical of spring varieties and lines. In the *Glu-B1* locus, the predominance of alleles associated with the production of the Bx7 and By9 subunits was revealed for both winter and spring varieties. In the case of the *Glu-D1* gene, for all the wheat groups studied, the composition of the Dx5+Dy10 subunits was the most common: in 92.3 % of winter and 68.2 % of spring PSZ accessions and in 80 % of winter and 55 % of spring VIR accessions. The analysis of genotypes showed the presence of 13 different allelic combinations of the *Glu-A1*, *Glu-B1*, *Glu-D1* genes in the PSZ varieties, and 19 combinations in the VIR varieties. The *b b*/*al*/**c** *d* allelic combination (Ax2* Bx7+By8/8*/**9** Dx5+Dy10) turned out to be the most common for the PSZ spring varieties and lines, while for the PSZ winter accessions it was *a c d* (Ax1 Bx7+By9 Dx5+Dy10); the *b c a* and *b c d* genotypes (Ax2* Bx7+By9 Dx2+Dy12 and Ax2* Bx7+By9 Dx5+Dy10, respectively) occur with equal frequency among the VIR spring accessions; in the group of VIR winter varieties, the combination of the *a b*/*al d* alleles (Ax1 Bx7+By8/8* Dx5+Dy10) prevails. The most preferred combination of alleles for baking qualities was found in the spring variety 'Ekaterina' and winter varieties 'Tarasovskaya 97', 'Volzhskaya S3', as well as in lines k-58164, L43510, L43709, L-67, L-83, which are recommended for further breeding programs to improve and preserve baking qualities in the conditions of the Pre-Ural steppe zone.
Key words: *Triticum aestivum*; genotyping; baking qualities; high molecular weight glutenins; *Glu-1* genes.

For citation: Galimova A.A., Kuluev A.R., Ismagilov K.R., Kuluev B.R. Genetic polymorphism of high-molecular-weight glutenin subunit loci in bread wheat varieties in the Pre-Ural steppe zone. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):297-305. DOI 10.18699/VJGB-23-36

Генетический полиморфизм локусов высокомолекулярных субъединиц глютенина у сортообразцов мягкой пшеницы Предуральской степной зоны

А.А. Галимова[1, 2] ✉, А.Р. Кулуев[1], К.Р. Исмагилов[3], Б.Р. Кулуев[1, 2]

[1] Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
[2] Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
[3] Башкирский научно-исследовательский институт сельского хозяйства Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
✉ aiz.galimova@yandex.ru

Аннотация. Высокомолекулярные глютенины играют важную роль в обеспечении хлебопекарных качеств зерна мягкой пшеницы *Triticum aestivum* L. Однако селекция мягкой пшеницы по данному признаку весьма трудоемка, поэтому большой интерес представляет генотипирование сортообразцов по аллельному составу генов высокомолекулярных глютенинов. Цель исследования состояла в определении состава субъединиц высокомолекулярных глютенинов на основе выявления аллельного состава генов *Glu-1*, а также в выявлении частоты встречаемости аллелей генов *Glu-1* в сортообразцах мягкой пшеницы, находящихся в селекционной работе в Предуральской степной зоне (ПСЗ). Проанализированы 26 озимых и 22 яровых сортообразца мягкой пшеницы,

находящихся в селекционной работе в ПСЗ, и 27 озимых и 20 яровых сортообразцов из коллекции ВИР. Анализ генотипов показал наличие 13 различных аллельных сочетаний генов *Glu-A1*, *Glu-B1*, *Glu-D1* в сортообразцах ПСЗ и 19 сочетаний – в сортообразцах ВИР. Самым распространенным для яровых сортов и линий ПСЗ оказалось сочетание аллелей *b b/al/c d* (Ax2* Bx7+By8/8*/**9** Dx5+Dy10), для озимых – *a c d* (Ax1 Bx7+By9 Dx5+Dy10). Среди яровых сортообразцов коллекции ВИР с равной частотой встречались генотипы *b c a* (Ax2* Bx7+By9 Dx2+Dy12) и *b c d* (Ax2* Bx7+By9 Dx5+Dy10); в группе озимых сортов ВИР преобладал генотип *a b/al d* (Ax1 Bx7+By8/8* Dx5+Dy10). Наиболее предпочтительное в целях хлебопечения сочетание аллелей выявлено у ярового сорта Екатерина и озимых сортов Тарасовская 97, Волжская С3, а также у линий к-58164, Л43510, Л43709, Л-67, Л-83, которые рекомендуются для дальнейших селекционных программ по улучшению и сохранению хлебопекарных качеств в условиях Предуральской степной зоны.
Ключевые слова: *Triticum aestivum*; генотипирование; хлебопекарные качества; высокомолекулярные глютенины; гены *Glu-1*.

Introduction

The gluten complex of bread wheat grain (*Triticum aestivum* L.) plays an important role in determining its baking qualities (Gomez et al., 2011). Grain gluten is formed by different groups of proteins (Weiser et al., 2006; Koehle, Weiser, 2013). The most significant among them are high-molecular glutenins: it is between them that intermolecular disulfide bonds resistant to high temperatures are formed. Thus, the number and composition of high molecular weight glutenin subunits (HMW-GS) are important in the formation of bakery product crumb structure, which is formed during baking, and seem to be significant factors that determine the baking quality of wheat grain (Dhaka, Khatkar, 2015).

Each of the loci of high molecular weight glutenins *Glu-A1*, *Glu-B1*, *Glu-D1* contains two paralogous genes encoding 'x'- and 'y'-type proteins, differing in molecular weights and sequences of conservative N-terminal domains (Payne et al., 1982; Shewry et al., 2003). Despite the presence of six *Glu-1* genes (coding for HMW-GS Ax1, Bx1, Dx1, Ay1, By1 and Dy1), the number of expressed HMW genes varies from three to five in different varieties of bread wheat, since the genes of the Ax1 and By1 subunits may not be expressed, and expression of the Ay1 subunit is always blocked (Luo et al., 2018). Glutenin genes are characterized by alleles associated with high and low productivity, grain quality, and adaptive potential (Konarev et al., 2000). It was found that good quality of baking, in particular, the elasticity and strength of the dough and the volume of bread, is significantly affected by the subunits encoded by the D subgenome (Yang et al., 2014), followed by the influence of the B subgenome (Zhang L.J. et al., 2015).

Various combinations of the *Glu-1* genes alleles of the ABD subgenomes determine the diversity of HMW-GS combinations, which affects the quality of wheat dough and final products. For example, the *Glu-A1a* allele (encoding the Ax1 subunit) is associated with a high gluten index and a long test development time (Tabiki et al., 2006). The *Glu-A1b* (Ax2*) allele is usually associated with good dough strength (Vazquez et al., 2012). In turn, the *Glu-A1c* (null) allele has a negative effect on the quality of the dough (Anjum et al., 2007). The *Glu-B1* locus is characterized by the presence of many alleles, among which the *Glu-B1f* (Bx13+By16) and *Glu-B1i* (Bx17+By18) alleles have a positive effect on the rheological properties of the dough and baking quality (Guo et al., 2019). Among the alleles of the *Glu-B1* locus, a unique allele *Glu-B1al* was found, which is associated with overexpression of the Bx7 subunit ($Bx7^{OE}$) and increased dough strength (Zhao et al., 2020). It is known that the presence of the *Glu-D1d* (Dx5+Dy10) allele contributes to high rheological properties and quality of the dough (Wang G. et al., 1993). If we consider the effect of various HMW-GS on the rheological properties of the dough, they have the following rank order of contribution to the dough strength: Dx5+Dy10 > Dx2+Dy12 > Dx3+Dy12 > Dx4+Dy12 (Payne, Lawrence, 1983; Zhang Y. et al., 2018); at the *Glu-B1* and *Glu-A1* loci: Bx17+By18 > Bx13+By16 > Bx7+By9 > Bx7+By8 > Bx6+By8 and Ax2* > Ax1 > Null, respectively (Patil et al., 2015).

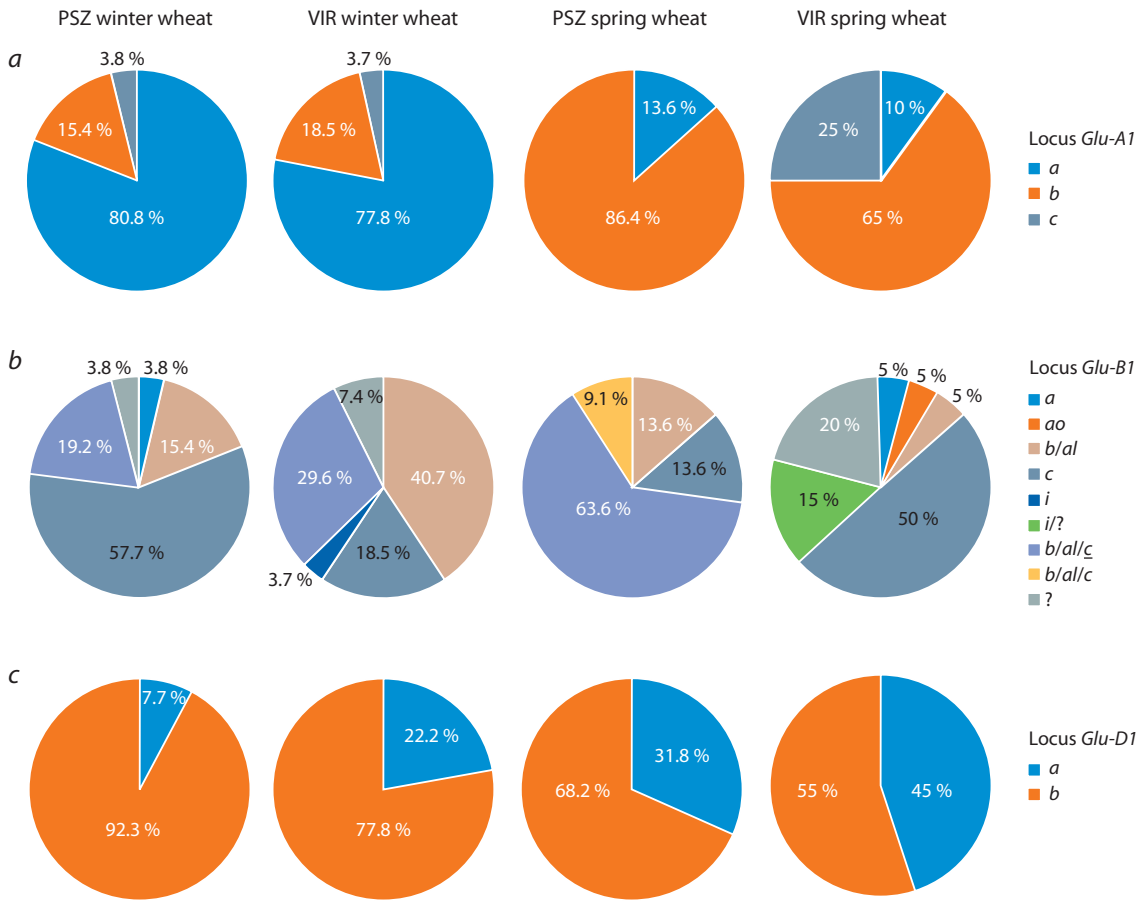
According to the estimates of the Federal State Budgetary Institution "Centre of Quality Assurance" for 2019, 46.2 % of bread wheat grain in the Republic of Bashkortostan belonged to Class 4, and 20.5 % was not food at all (http://fczerna.ru/). Thus, one of the problems in the cultivation of bread wheat in the Republic of Bashkortostan is the low quality of grain, which depends on the genotype, but the negative impact of the soil and climatic conditions of our region is also not excluded. Based on the importance of HMW-GS in the formation of high baking qualities of wheat grain, elucidating the composition of HMW-GS and their characteristics is an important and relevant task for breeding, aimed, among other things, at improving and preserving the baking qualities of wheat. Such studies have never been carried out before in the conditions of the Pre-Ural steppe zone (PSZ). Therefore, the aims of our work were to determine the allelic state of the *Glu-1* loci, to identify the composition of HMW-GS based on PCR, and to analyze the frequency of different genotypes occurrence in bread wheat varieties from the collections of the Bashkir Research Institute of Agriculture of the UFRC of the Russian Academy of Sciences and Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR).

Materials and methods

The objects of the study were 48 varieties and lines of bread wheat included in the breeding programs and zoned to the soil and climatic conditions of the PSZ, the baking qualities of which had not been studied, as well as 47 random varieties from the VIR collection with known baking qualities (Suppl. Material 1)[1].

Genomic DNA was isolated from dried leaves by a standard method using CTAB (Doyle J.J., Doyle J.L., 1987). Genotyping of the samples was carried out by PCR analysis, visualization of the results was carried out in 1.6 % agarose and 8 % polyacrylamide gels with markers of DNA frag-

[1] Supplementary Materials 1–3 are available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Galimova_Engl_27_4.pdf



Distribution of the *Glu-1* genes alleles in cultivars and lines zoned to the conditions of the PSZ and from the VIR collection.
Alleles of ABD subgenomes and their frequency of occurrence: *a* – subgenome A; *b* – subgenome B; *c* – subgenome D.

ment with lengths of 50 bp and 100 bp (Eurogen, Russia). To identify the allelic states of the *Glu-1* genes, we used primer pairs UMN19F/19R (Liu et al., 2008), a/b (Lafiandra et al., 1997), and BxF/BxR (Ma et al., 2003), as well as ZSBy9aF1/ZSBy9aR3, ZSBy9aF2/ZSBy9aR2 (Lei et al., 2006) for subgenomes A and B, respectively. In the evaluation of varieties and lines at the *Glu-B1* locus, the BxF/BxR primer pair was used to detect alleles of the x-type subunits, and the ZSBy9aF1/ZSBy9aR3 and ZSBy9aF2/ZSBy9aR2 primer pairs were used to detect alleles of the y-type subunits. The allelic state of the *Glu-D1* gene was determined by duplex PCR using primer pairs UMN25F/25R and UMN26F/26R (Liu et al., 2008). Primer sequences and sizes of PCR products are given in Suppl. Material 2. Examples of the electrophoregrams of different alleles of the ABD subgenomes are shown in Suppl. Material 3.

Results

Glu-A1 locus genotyping

PCR analysis of PSZ winter wheat samples revealed the presence of *a* allele associated with subunit Ax1 in 21 out of 26 (80.8 %) samples, allele *b* (Ax2*) – in 4 (15.4 %) and allele *c* (Ax-null) – in 1 (3.8 %) winter wheat samples. Allele *a* was found in 3 out of 22 (13.6 %) varieties, allele *b* was found in 19 (86.4 %) (see Suppl. Material 3, *a*) varieties in the group of PSZ spring wheat cultivars and lines; cultivars with the *c* allele were not found in this group.

The following results were obtained for samples from the VIR collection: allele *a* was detected in 21 (77.8 %) winter and 2 (10 %) spring varieties; allele *b* – in 5 (18.5 %) winter and 13 (65 %) spring varieties; allele *c* – in 1 (3.7 %) winter and 5 (25 %) spring varieties and lines (see the Table and the Figure, *a*).

Glu-B1 locus genotyping

PCR analysis of 26 winter varieties of PSZ showed the presence of the Bx7 subunit in 25 (96 %) varieties (see the Figure, *b*). In the case of spring cultivars and lines of PSZ, it was revealed that all the studied varieties carry the Bx7 subunit. When analyzing winter varieties of VIR, it was found that all samples carry the allele associated with Bx7 subunits, except for cultivars Zarya, Kolkhoznitsa and Avesta. Amplicons, formed during PCR in the Zarya and Kolkhoznitsa cultivars, have lengths that are not characteristic of loci associated with subunits Bx6, Bx7/7* and Bx17; for cultivar Avesta the Bx17 subunit amplicon was detected (see Suppl. Material 3, *c*). In the course of genetic analysis of varieties and lines of spring wheat from VIR, the allele encoding the Bx7 subunit was established for 13 cultivars out of 20 (65 %), for 3 cultivars – the Bx17

Results of the allelic composition analysis of the *Glu-1* genes and the HMW-GS composition

No.	Cultivar/line	HMW-GS			*Glu-1* allele		
		Glu-A1	*Glu-B1*	*Glu-D1*	*Glu-A1*	*Glu-B1*	*Glu-D1*
			PSZ winter wheat				
1	Alabasskaya	1	7+null	5+10	*a*	*a*	*d*
2	Bashkirskaya 10	1	7+9	5+10	*a*	*c*	*d*
3	Bezenchukskaya 380	1	7+9	5+10	*a*	*c*	*d*
4	Volzhskaya kachestvennaya	2*	7+9	5+10	*b*	*c*	*d*
5	Lutescens 25520	1	7+9	5+10	*a*	*c*	*d*
6	Lutescens 47488	1	7+8/8*	5+10	*a*	*b/al*	*d*
7	Lutescens 60865	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
8	Lutescens 65532	1	7+9	5+10	*a*	*c*	*d*
9	Lutescens 65737	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
10	Lutescens 65752	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*
11	Lutescens 67750	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*
12	Lutescens 68347	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
13	Lutescens 68551	1	7+8/8*	5+10	*a*	*b/al*	*d*
14	Erythrospermum 37067	1	7+9	5+10	*a*	*c*	*d*
15	Erythrospermum 69577	1	7+9	5+10	*a*	*c*	*d*
16	Erythrospermum 70757	null	7+8/8*	5+10	*c*	*b/al*	*d*
17	Agidel	1	7+9	5+10	*a*	*c*	*d*
18	Agidel N	1	7+9	5+10	*a*	*c*	*d*
19	Agidel 2	1	7+9	5+10	*a*	*c*	*d*
20	Anastasia	1	?+ null/20	5+10	*a*	?	*d*
21	Aelita	1	7+9	5+10	*a*	*c*	*d*
22	Kalach 60	1	7+9	5+10	*a*	*c*	*d*
23	Lana	1	7+9	5+10	*a*	*c*	*d*
24	Moskovskaya 39	2*	7+9	5+10	*b*	*c*	*d*
25	Novoershovskaya	1	7+8/8*	5+10	*a*	*b/al*	*d*
26	Ufimka	1	7+9	5+10	*a*	*c*	*d*
			PSZ spring wheat				
1	Bashkirskaya 28	1	7+8/8*	5+10	*a*	*b/al*	*d*
2	Ekaterina	2*	7+8/8*	5+10	*b*	*b/al*	*d*
3	Zaural'skaya zhemchuzhina	2*	7+**8**/**8***/**9**	5+10	*b*	***b**/**al**/**c***	*d*
4	L43466	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
5	L43510	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
6	L43705	2*	7+9	5+10	*b*	*c*	*d*
7	L43706	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*
8	L43709	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
9	Omskaya 35	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
10	Tulaykovskaya 108	2*	7+9	2+12	*b*	*c*	*a*
11	Ekada 109	2*	7+9	5+10	*b*	*c*	*d*
12	Ekada 113	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*
13	Arhat	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
14	L-21	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*
15	L-63	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*
16	L-67	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
17	L-83	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
18	Omskaya 36	2*	7+9	5+10	*b*	*c*	*d*
19	Salavat Yulaev	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
20	Watan	1	7+**8**/**8***/**9**	2+12	*a*	***b**/**al**/**c***	*a*
21	Ekada 70	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
22	Saratovskaya 55	2*	7+8/8*/**9**	2+12	*b*	*b/al/**c***	*a*

Table (end)

No.	Cultivar/line	HMW-GS			*Glu-1* allele		
		Glu-A1	*Glu-B1*	*Glu-D1*	*Glu-A1*	*Glu-B1*	*Glu-D1*
			VIR winter wheat				
1	Zarya	1	?+9	2+12	*a*	?	*a*
2	Strelinskaya	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
3	Kolkhoznitsa	1	?	2+12	*a*	?	*a*
4	Gorkovchanka	1	7+8/8*	5+10	*a*	*b/al*	*d*
5	Bezostaya 1	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
6	Mironovskaya 808	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
7	Volna	1	7+8/8*/**9**	2+12	*a*	*b/al/**c***	*a*
8	Lutescens 61	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
9	Erythrospermum 60	2*	7+8/8*/**9**	5+10	*b*	*b/al/**c***	*d*
10	Lidia	1	7+8/8*/**9**	2+12	*a*	*b/al/**c***	*a*
11	Tarasovskaya 87	1	7+9	5+10	*a*	*c*	*d*
12	Zernogradka 9	1	7+8/8*/**9**	5+10	*a*	*b/al/**c***	*d*
13	Tarasovskaya 97	2*	7+8/8*	5+10	*b*	*b/al*	*d*
14	Prestizh	1	7+8/8*	5+10	*a*	*b/al*	*d*
15	Tarasovskaya ostistaya	1	7+9	5+10	*a*	*c*	*d*
16	Goryanka	1	7+8/8*	5+10	*a*	*b/al*	*d*
17	Avgusta	1	7+8/8*	5+10	*a*	*b/al*	*d*
18	Avesta	1	17+18	2+12	*a*	*i*	*a*
19	Agra	1	7+8/8*	2+12	*a*	*b/al*	*a*
20	Al'bina	1	7+8/8*	5+10	*a*	*b/al*	*d*
21	Doneko	1	7+8/8*	5+10	*a*	*b/al*	*d*
22	Dominanta	null	7+9	5+10	*c*	*c*	*d*
23	Volzhskaya S3	2*	7+8/8*	5+10	*b*	*b/al*	*d*
24	Zhemchuzhina Povolzh'ya	1	7+8/8*	5+10	*a*	*b/al*	*d*
25	Donna	1	7+9	5+10	*a*	*c*	*d*
26	Dmitriy	1	7+8/8*	5+10	*a*	*b/al*	*d*
27	Levoberezhnaya 3	1	7+9	5+10	*a*	*c*	*d*
			VIR spring wheat				
1	Tin-Ci-En 18	null	7+16	2+12	*c*	*ao*	*a*
2	Calanda	2*	7+9	5+10	*b*	*c*	*d*
3	Bola Picota	1	7+null	5+10	*a*	*a*	*d*
4	58006/382215	2*	7+9	5+10	*b*	*c*	*d*
5	58013/382239	1	7+9	5+10	*a*	*c*	*d*
6	Leningradka Krupnozernaya	null	?	2+12	*c*	?	*a*
7	C-75094	2*	17+18/15	5+10	*b*	*i*/?	*d*
8	58164/462069	2*	7+8/8*	5+10	*b*	*b/al*	*d*
9	Niab 545	2*	17+18/15	5+10	*b*	*i*/?	*d*
10	Bithoor	2*	17+18/15	5+10	*b*	*i*/?	*d*
11	Sheridan	null	?+8/8*/18/15	2+12	*c*	?	*a*
12	MS 4351	2*	7+9	2+12	*b*	*c*	*a*
13	MSK 1002	null	7+9	2+12	*c*	*c*	*a*
14	Egisar 29	2*	7+9	5+10	*b*	*c*	*d*
15	Nabat	2*	7+9	5+10	*b*	*c*	*d*
16	Ranger	2*	?+8/8*/18/15	2+12	*b*	?	*a*
17	Shortana	2*	7+9	2+12	*b*	*c*	*a*
18	Lutescens 275	2*	7+9	2+12	*b*	*c*	*a*
19	55870/356485	2*	7+9	2+12	*b*	*c*	*a*
20	56471/69970	null	?+8/8*/18/15	5+10	*c*	?	*d*

Note. HMW-GS – high molecular weight glutenin subunits. Bold and underline indicate subunits the expression of which is predominant; alleles and subunits that could not be accurately identified are marked with a question mark.

subunit (15 %), for 4 cultivars it was not possible to identify x-type subunits using the BxF/BxR primers pair (Leningradka Krupnozernaya, Sheridan, Ranger, 56471/69970).

The presence of monomorphic and polymorphic varieties and lines in the studied samples was revealed using pairs of primers ZSBy9aF1/ZSBy9aR3 and ZSBy9aF2/ZSBy9aR2. The share of monomorphic varieties was 80.8 and 70.4 % for winter wheat, and 27.3 and 100 % for spring wheat in the PSZ and VIR samples, respectively. Polymorphic cultivars and lines are the samples, during genotyping of which alleles of several glutenin-coding loci differed in electrophoregrams.

Polymorphism of the *Glu-B1* locus is represented by a combination of glutenin subunits Bx7+By8/8*/Bx7+By9. At the same time, this combination occurs in two forms: the same level of expression of the By-type subunits genes (7+**8/8*/9**) and the predominance of expression of the By9 subunit genes (7+8/8*/**9**). Thus, in the course of genetic analysis using primers ZSBy9aF1/ZSBy9aR3, the allele of the By9 subunit in the monomorphic form was detected in 15 out of 26 (57.7 %) varieties of winter wheat and in 3 out of 22 (13.6 %) varieties of spring wheat of the PSZ samples. The polymorphic form in the combinations of By (7+8/8*/**9**) and By (7+**8/8*/9**) was found in 19.2 and 0 % of winter wheat samples and in 63.6 and 9.1 % of spring wheat samples, respectively. The second pair of primers, ZSBy9aF2/ZSBy9aR2, made it possible to identify the By8/By8* allele in 4 winter wheat and 3 spring wheat varieties of PSZ. The Bynull/20 subunit allele was found in the cultivar Anastasia in the selection of winter wheat cultivars and lines of PSZ, while it was not found in the selection of spring wheat varieties of PSZ. The Alabasskaya cultivar was found to carry the *a* allele, which is characterized by the set of subunits Bx7+Bynull.

Analysis of varieties and lines of VIR using the ZSBy9aF1/ZSBy9aR3 primer pair showed that the monomorphic allele of the By9 subunit (see Suppl. Material 3, *d*) was present in 6 out of 27 (22.2 %) winter wheat and 10 out of 20 (50 %) spring wheat cultivars. Polymorphism in the VIR variety samples is found only in the form of By (7+8/8*/**9**) solely in winter cultivars (29.6 %). It was revealed that 11 out of 27 winter varieties and lines carry the Bx8/8* subunit allele using primers ZSBy9aF2/ZSBy9aR2; the y-type subunit was not established for the Kolkhoznitsa cultivar. It was found that 17 out of 20 spring varieties and lines form 2 reaction products, the three remaining varieties have 0, 1 and 3 amplification reaction products with primers ZSBy9aF2/ZSBy9aR2. Thus, according to the results of PCR of two primer pairs ZSBy9aF1/ZSBy9aR3 and ZSBy9aF2/ZSBy9aR2, it can be seen that 10 out of 20 (50 %) spring cultivars carry the allele associated with the production of the By9 subunit, 7 (35 %) cultivars carry the 8/8*/18/15 subunit allele, 1 cultivar (Bola Picota) – the null/20 subunit allele, 1 cultivar (Tin-Ci-En 18) – the By16 subunit allele (see the Table and the Figure, *b*).

Glu-D1 locus genotyping

Identification of the allelic composition of the *Glu-D1* gene was carried out using two pairs of primers to simultaneously detect alleles of genes encoding subunits Dx5 and Dy10 (5+10), Dx2 and Dy12 (2+12). When using these primers for samples with subunits 5+10 we observed the presence of 397 and 281 bp amplicons (allele *d*), and for samples with subunits 2+12 – 415 and 299 bp amplicons (allele *a*) (see Suppl. Material 3, *b*). Thus, in the sample of PSZ cultivars, the *d* allele was detected in 24 (92.31 %) winter and 15 (68.18 %) spring cultivars; in a sample of varieties and lines from the VIR collection, this allele was found in 24 (88.9 %) winter and 11 (55 %) spring cultivars of wheat (see the Table and the Figure, *c*).

Glu-1 gene allele combination (ABD)

The presence of 8 different combinations of the *Glu-1* (ABD) gene alleles was characteristic for winter wheat, and 9 different combinations of alleles – for spring wheat of PSZ varieties. The predominant alleles combination for winter wheat was the combination of alleles *a c d* – it was detected in 13 out of 26 varieties (50 %) of the samples of PSZ, while in the case of spring wheat, the combination of *b b/al/***c** *d* was predominant (36.4 %), so a significant number of spring varieties had a genotype with *b b/al/***c** *a* alleles combination (22.7 %).

The presence of 10 different combinations of the *Glu-1* (ABD) gene alleles was revealed in 27 winter wheat varieties and 11 combinations of alleles – in the spring wheat forms from 20 VIR cultivars and lines. The predominant combinations of alleles of winter varieties turned out to be the combination *a b/al d* (29.6 %; 8 varieties out of 27), in spring varieties, combinations of alleles *b c d* and *b c a* were the most common (20 % of each combination; 4 varieties out of 20). In addition, the combination of alleles *b i/? d* was observed in 15 % of varieties (3 varieties out of 20).

Discussion

To date, the existence of correlations between the presence of certain alleles of the *Glu-1* genes and indicators of the baking quality of bread wheat grain has been shown. HMW-GS genes loci are characterized by high polymorphism (Patil et al., 2015), which may be one of the reasons for the genetic variability of wheat cultivars in terms of the rheological and technological properties of the dough. On the other hand, these qualities are rather difficult to control in classical breeding. Therefore, genotyping of HMW-GS alleles is an important task aimed at the effective selection of parental forms with high baking qualities, and the results obtained can be applied in marker-assisted and genomic selection of bread wheat. However, first it is necessary to carry out work on the assessment of the genetic diversity of bread wheat cultivars according to the allelic composition of high molecular weight glutenin genes in certain regions, which was the subject of our study.

The Ax1 and Ax2* subunits have a positive effect on the quality of the dough, while the null subunit has a negative effect (Anjum et al., 2007). We have found that in winter varieties for subgenome A, allele *a* (Ax1) is predominant, while in spring varieties it is allele *b* (Ax2*) (see the Figure, *a*). This pattern is typical for both studied groups (PSZ and VIR).

Eight different *Glu-B1* alleles were identified for all the studied varieties and lines, the total number of which was 95 (see the Figure, *b*). At the same time, it was not possible to identify the exact allele for 7 varieties, since reaction products formed during PCR either did not correspond in size to the expected ones, or were absent. These samples require further research, including by sequencing, for the possible carriage of new alleles. Among the eight identified alleles, the highest

frequency of occurrence was determined for the *c* allele (Bx7+By9) (34.7 % of the total number of all PSZ and VIR samples), which correlates with the results of other studies that indicate a wide distribution of this allele (Payne, Lawrence, 1983; Gianibelli et al., 2001).

When analyzing a sample of varieties and lines of bread wheat PSZ, it was found that the most common alleles are *c* and *b/al/**c***. Thus, 57.7 % of winter varieties of PSZ carry the *c* allele (Bx7+By9), 63.6 % of spring varieties and lines of PSZ carry the *b/al/**c*** allele (Bx7+By8/8*/**9**). However, when analyzing varieties and lines from the VIR collection, it was revealed that most of the *c* alleles are carried by spring wheat varieties (50 %), and winter wheat varieties with a frequency of 40.7 % carry the *b/al* allele (Bx7+By8/8*). Thus, the winter wheat varieties of PSZ were more often monomorphic, in contrast to the spring wheat varieties, for which the production of PCR products characteristic of both the *b/al* allele (Bx7+By8/8*) and the *c* allele (Bx7+By9) was detected. At the same time, it should be noted that the production of amplicons associated with the *c* allele occurs much more often compared to the amplicons of the *b/al* allele (varieties have the *b/al/**c*** genotype (Bx7+By8/8*/**9**)). The distribution of alleles *b/al* (Bx7+By8/8*) and *c* is not surprising, since these alleles are associated with good baking qualities (Payne, Lawrence, 1983).

In the studied samples, it is worth paying attention to the winter wheat cultivar Avesta from the VIR collection, since this cultivar is likely to carry the *Glu-B1i* allele (Bx17+By18) associated with high grain quality, and can be selected for breeding programs aimed at improving grain quality. In addition to the Avesta cultivar, it is necessary to carry out additional studies in order to establish the exact allele, namely, to identify the allele of the y-type subunit in spring wheat varieties from VIR – C-75094, Niab 545, Bithoor. If the presence of the *Glu-B1i* allele (Bx17+By18) is confirmed, these wheat varieties can also be used in breeding programs. When planning hybridization work, it should be borne in mind that the Alabasskaya and Bola Picota cultivars carry the *a* allele (Bx7) of the B subgenome, which is associated with a rather low grain quality assessment. The Bola Picota cultivar, according to VIR data, really has low baking qualities (see Suppl. Material 1). So, the analysis of the *Glu-Bx* locus showed that the carriage of the allele associated with the production of Bx7 subunits is predominant for both winter and spring forms of bread wheat. Genetic analysis of the *Glu-By* locus revealed the greatest distribution of the allele associated with the production of the By9 subunit in wheats of both forms of vernalization.

During the genetic analysis of the *Glu-D1* gene, it was revealed that all the studied varieties and lines of PSZ and VIR carry one of two alleles – *a* or *d*. The results of studies of VIR samples also indicate a wide distribution of these alleles and the Dx2+Dy12 and Dx5+Dy10 glutenin subunits encoded by them (Ayala et al., 2016). In addition, it was found that for all the studied wheat groups, the composition of the subunits Dx5+Dy10 (allele *d*) is the most common (frequency of occurrence in total in cultivars and lines of PSZ is 83.3 %; in the total sample of VIR – 72.3 %). At the same time, the frequency of occurrence of this allele in winter wheat varieties is higher than in spring ones (see the Figure, *c*). It is known that the *d* allele (Dx5+Dy10) of the *Glu-D1* locus has a pronounced positive effect on flour quality (Payne, Lawrence, 1983), which is consistent with selection aimed at improving the baking qualities of grain. The second allele of the D subgenome identified by us is the allele *a* (Dx2+Dy12); theoretically, it can have a negative impact on the production of high-quality pan bread, but it is recommended for cultivars used for making hearth bread and noodles. However, it is worth mentioning that this allele is not always associated with poor grain quality. Relatively recently, genes have been discovered that cause the synthesis of Dy12.7 subunits (Peng et al., 2015) and Dy12** (Du et al., 2019), which have similar molecular weights to standard Dy12, but are associated with increased grain quality.

In addition to the influence of single alleles on the quality of the final product, the quality of dough and bread is affected by the total effect of alleles of all three subgenomes (Payne et al., 1981; Wang Z.J. et al., 2018; Zhao et al., 2020). In the sample of all analyzed wheat samples, 24 allelic combinations were identified at the *Glu-1* locus, among which the combination of alleles *b b/al d* (Ax2* Bx7+By8/8* Dx5+Dy10) is the most preferable for baking purposes (Pirozi et al., 2008). A similar combination of alleles was detected in the spring wheat cultivar Ekaterina. According to the data given in the Russian State Register for Selection Achievements, this cultivar has good baking qualities. Among the varieties of VIR, the combination of alleles *b b/al d* (Ax2* Bx7+By8/8* Dx5+Dy10) is found in winter wheat cultivars Tarasovskaya 97, Volzhskaya S3 and spring wheat line k-58164. Cultivars Tarasovskaya 97 and Volzhskaya S3, according to the Russian State Register for Selection Achievements, have good baking qualities. Thus, winter wheat cultivars Tarasovskaya 97, Volzhskaya S3 and spring wheat cultivar Ekaterina can be taken into account when selecting parental pairs for breeding bread wheat in order to improve baking qualities.

Combinations of alleles *a c d* (Ax1 Bx7+By9 Dx5+Dy10) (50 % – PSZ, 14.8 % – VIR) and *a b/al d* (Ax Bx7+By8/8* Dx5+Dy10) (11.5 % – PSZ, 29.6 % – VIR) turned out to be the most common in the total sample of winter wheat varieties and lines. It was reported that these combinations of subunits are rated 9 and 10 scores on the Payne scale. From the frequency of occurrence of alleles *a c d* (Ax1 Bx7+By9 Dx5+Dy10) (9 scores on the Payne scale) and *a b/al d* (Ax1 Bx7+By8/8* Dx5+Dy10) (10 scores) in the sample of winter wheat varieties of PSZ and the VIR collection, it can be seen that in regions with climatic conditions of PSZ, selection went towards the fixation of the *c* allele of the subgenome B in comparison with the VIR samples, which collected varieties and lines from different regions and countries (see the Figure, *b*). However, in the studied sample of spring wheat varieties, the opposite picture is observed – classical selection in PSZ goes towards the acquisition of the *b* allele (Bx7+By8) of the B subgenome (see the Figure, *b*). So, according to the frequency of occurrence of genotypes *b b/al/**c*** *a* (Ax2* Bx7+By8/8*/**9** Dx2+Dy12) (22.7 % – PSZ, 0 % – VIR), *b b/al/**c*** *d* (Ax2* Bx7+By8/8*/**9** Dx5+Dy10) (36.4 % – PSZ, 0 % – VIR), *b c d* (Ax2* Bx7+By9 Dx5+Dy10) (13.6 % – PSZ, 20 % – VIR) and *b c a* (Ax2* Bx7+By9 Dx2+Dy12) (0 % – PSZ, 20 % – VIR), it can be seen that heteromorphism of the *Glu-B1* locus appeared in the group of PSZ varieties and lines.

The lowest HMW-GS scores according to the Payne scale have genotypes *b c a* (Ax2* Bx7+By9 Dx2+Dy12) (Glu-1 score = 7). This combination of alleles was identified in the spring wheat cultivar Tulaykovskaya 108, but according to the Russian State Register for Selection Achievements, this cultivar belongs to strong wheat with good baking qualities. However, according to our data, in the conditions of the Republic of Bashkortostan, cultivar Tulaykovskaya 108 is characterized only by satisfactory baking qualities. These contradictory data, among other things, are probably related to the climatic conditions of the regions, and the genotype of the cultivar may become a limiting factor under adverse external conditions. In the group of cultivars and lines from the VIR collection, genotype *b c a* was found in 20 % of spring wheat samples. In addition to the *b c a* genotype, combinations of alleles *b b/al/**c** a* (Ax2* Bx7+By8/8*/**9** Dx2+Dy12) (Glu-1 score = 7 or 5) and *c c a* (Ax-null Bx7+By9 Dx2+Dy12) (Glu-1 score = 5) have low scores on the Payne scale. The last combination of alleles has been found in one spring wheat line of VIR – MSK 1002. Genotype *b b/al/**c** a* (Ax2* Bx7+By8/8*/**9** Dx2+Dy12), despite the low Glu-1 score, is widespread in the group of spring wheat varieties PSZ (22.7 %), two winter lines of PSZ also have this genotype – Lutescens 65752 and Lutescens 67750.

Conclusion

Polymorphism of bread wheat cultivars zoned to the conditions of PSZ was studied for the genes of high molecular glutenins *Glu-1* by PCR analysis. Among the 26 studied winter cultivars of PSZ, the predominant majority had genotypes with the *a c d* allele (Ax1 Bx7+By9 Dx5+Dy10). The *b b/ al/**c** d* (Ax2* Bx7+By8/8*/**9** Dx5+Dy10) alleles combination was found to be dominant for 22 spring cultivars of PSZ. Among the 24 identified ones, the combination of alleles *b b/al d* (Ax2* Bx7+By8/8* Dx5+Dy10) of the *Glu-1* genes is the most preferred for baking purposes. This genotype was detected in the spring wheat cultivar Ekaterina and winter wheat cultivars Tarasovskaya 97, Volzhskaya S3 and spring wheat line k-58164. Of the promising spring wheat lines of the PSZ, the lines L43510, L43709, L-67, L-83, which are recommended for further breeding programs to improve and preserve baking qualities in the conditions of the Pre-Ural steppe zone, may have the highest score according to Payne.

References

Anjum F.M., Khan M.R., Din A., Saeed M., Pasha I., Arshad M.U. Wheat gluten: high molecular weight glutenin subunits – structure, genetics, and relation to dough elasticity. *J. Food Sci*. 2007;72(3): R56-R63. DOI 10.1111/j.1750-3841.2007.00292.x.

Ayala M., Guzman C., Pena R.J., Alvarez J.B. Diversity of phenotypic (plant and grain morphological) and genotypic (glutenin alleles in *Glu-1* and *Glu-3* loci) traits of wheat landraces (*Triticum aestivum* L.) from Andalusia (Southern Spain). *Genet. Resour. Crop Evol.* 2016;63(3):465-475. DOI 10.1007/s10722-015-0264-0.

Dhaka V., Khatkar B.S. Effects of gliadin/glutenin and HMW-GS/LMW-GS ratio on dough rheological properties and bread-making potential of wheat varieties. *J. Food Qual*. 2015;38(2):71-82. DOI 10.1111/jfq.12122.

Doyle J.J., Doyle J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull*. 1987;19(1):11-15.

Du X., Hu J., Ma X., He J., Hou W., Guo J., Bo C., Wang H., Li A., Kong L. Molecular characterization and marker development for high molecular weight glutenin subunit 1Dy12** from Yunnan hulled wheat. *Mol. Breed.* 2019;39(1):4. DOI 10.1007/s11032-018-0910-2.

Gianibelli M.C., Larroque O.R., Macritchie F., Wrigley C.W. Biochemical, genetic, and molecular characterization of wheat glutenin and its component subunits. *Cereal Chem*. 2001;78(6):635-646. DOI 10.1094/CCHEM.2001.78.6.635.

Gomez A., Ferrero C., Calvelo A., Anon M.C., Puppo M.C. Effect of mixing time on structural and rheological properties of wheat flour dough for breadmaking. *Int. J. Food Prop*. 2011;14(3):583-598. DOI 10.1080/10942910903295939.

Guo H., Wu J., Lu Y., Yan Y. High-molecular-weight glutenin 1Bx17 and 1By18 subunits encoded by *Glu-B1i* enhance rheological properties and bread-making quality of wheat dough. *J. Food Qual.* 2019;2019:1958747. DOI 10.1155/2019/1958747.

Koehler P., Wieser H. Chemistry of cereal grains. In: Gobbetti M., Gänzle M. (Eds.) Handbook on Sourdough Biotechnology. MA, Boston: Springer, 2013;11-45. DOI 10.1007/978-1-4614-5425-0_2.

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

Lafiandra D., Tucci G.F., Pavoni A., Turchetta T., Margiotta B. PCR analysis of x- and y-type genes present at the complex *Glu-A1* locus in durum and bread wheat. *Theor. Appl. Genet.* 1997;94(2):235-240. DOI 10.1007/s001220050405.

Lei Z.S., Gale K.R., He Z.H., Gianibelli C., Larroque O., Xia X.C., Butow B.J., Ma W. Y-type gene specific markers for enhanced discrimination of high-molecular weight glutenin alleles at the *Glu-B1* locus in hexaploid wheat. *J. Cereal Sci.* 2006;43(1):94-101. DOI 10.1016/j.jcs.2005.08.003.

Liu S., Chao S., Anderson J.A. New DNA markers for high molecular weight glutenin subunits in wheat. *Theor. Appl. Genet.* 2008;118(1): 177-183. DOI 10.1007/s00122-008-0886-0.

Luo G., Song S., Zhao L., Shen L., Song Y., Wang X., Yu K., Liu Z., Li Y., Yang W., Li X., Zhan K., Zhang A., Liu D. Mechanisms, origin and heredity of *Glu-1Ay* silencing in wheat evolution and domestication. *Theor. Appl. Genet.* 2018;131(7):1561-1575. DOI 10.1007/s00122-018-3098-2.

Ma W., Zhang W., Gale K. Multiplex-PCR typing of high molecular weight glutenin alleles in wheat. *Euphytica.* 2003;134:51-60. DOI 10.1023/A:1026191918704.

Patil V.R., Talati J.G., Singh C., Parekh V.B., Jadeja G.C. Genetic variation in glutenin protein composition of aestivum and durum wheat cultivars and its relationship with dough quality. *Int. J. Food Properties*. 2015;18(9-12):2393-2408. DOI 10.1080/10942912.2014.980948.

Payne P.I., Corfield K.G., Holt L.M., Blackman J.A. Correlations between the inheritance of certain high-molecular weight subunits of glutenin and bread-making quality in progenies of six crosses of bread wheat. *J. Sci. Food Agric*. 1981;32(1):51-60. DOI 10.1002/jsfa.2740320109.

Payne P.I., Holt L.M., Worland A.J., Law C.N. Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin: Part 3. Telocentric mapping of the subunit genes on the long arms of the homoeologous group 1 chromosomes. *Theor. Appl. Genet.* 1982; 63(2):129-138. DOI 10.1007/BF00303695.

Payne P.I., Lawrence G.J. Catalogue of alleles for the complex gene loci, *Glu-A1*, *Glu-B1*, and *Glu-D1* which code for high-molecular-weight subunits of glutenin in hexaploid wheat. *Cereal Res. Commun.* 1983;11(1):29-35.

Peng Y., Yu K., Zhang Y., Islam S., Sun D., Ma W. Two novel y-type high molecular weight glutenin genes in Chinese wheat landraces of the Yangtze-River region. *PloS One*. 2015;10(11):e0142348. DOI 10.1371/journal.pone.0142348.

Pirozi M.R., Margiotta B., Lafiandra D., MacRitchie F. Composition of polymeric proteins and bread-making quality of wheat lines with allelic HMW-GS differing in number of cysteines. *J. Cereal Sci*. 2008;48(1):117-122. DOI 10.1016/j.jcs.2007.08.011.

Shewry P.R., Halford N.G., Tatham A.S., Popineau Y., Lafiandra D., Belton P.S. The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Adv. Food Nutr. Res*. 2003;45(3):219-302. DOI 10.1016/s1043-4526(03)45006-7.

Tabiki T., Ikeguchi S., Ikeda T.M. Effects of high-molecular-weight and low-molecular-weight glutenin subunit alleles on common wheat flour quality. *Breed. Sci.* 2006;56(2):131-136. DOI 10.1270/jsbbs.56.131.

Vazquez D., Berger A.G., Cuniberti M., Bainotti C., Miranda M.Z., Scheeren P.L., Jobet C., Zuniga J., Cabrera G., Verges R., Pena R.J. Influence of cultivar and environment on quality of Latin American wheats. *J. Cereal Sci*. 2012;56(2):196-203. DOI 10.1016/j.jcs.2012.03.004.

Wang G., Snape J., Hu H., Rogers W. The high-molecular-weight glutenin subunit compositions of Chinese bread wheat varieties and their relationship with bread-making quality. *Euphytica*. 1993;68:205-212. DOI 10.1007/BF00029874.

Wang Z.J., Wang X., Zhang Y., Zhang B., Florides Ch.G., Gao Z., Zhang X., Wei Y. Comparison of quality properties between high-molecular-weight glutenin subunits 5+10 and 2+12 near-isogenic lines under three common wheat genetic backgrounds. *Cereal Chem*. 2018;95(4):575-583. DOI 10.1002/cche.10061.

Wieser H., Bushuk W., MacRitchie F. The polymeric glutenins. In: Wrigley C.W., Bekes F., Bushuk W. (Eds.) Gliadin and Glutenin. The Unique Balance of Wheat Quality. USA, Minnesota, St. Paul: Am. Assoc. of Cereal Chemists, 2006;213-240. DOI 10.1094/9781891127519.010.

Yang Y.H., Li S.M., Zhang K.P., Dong Z., Li Y., An X., Chen J., Chen Q., Jiao Z., Liu X., Qin H., Wang D. Efficient isolation of ion beam-induced mutants for homoeologous loci in common wheat and comparison of the contributions of *Glu-1* loci to gluten functionality. *Theor. Appl. Genet.* 2014;127(2):359-372. DOI 10.1007/s00122-013-2224-4.

Zhang L.J., Chen Q.F., Su M.J., Yan B., Zhang X., Jiao Zh. High-molecular-weight glutenin subunit-deficient mutants induced by ion beam and the effects of *Glu-1* loci deletion on wheat quality properties. *J. Food Sci. Agric*. 2015;96(4):1289-1296. DOI 10.1002/jsfa.7221.

Zhang Y., Hu M., Liu Q., Sun L., Chen X., Lv L., Liu Y., Jia X., Li H. Deletion of high-molecular-weight glutenin subunits in wheat significantly reduced dough strength and bread-baking quality. *BMC Plant Biol.* 2018;18(1):319. DOI 10.1186/s12870-018-1530-z.

Zhao H., Gao C., Song W., Zhang Y.B., Gao D., Zhang X.M., Zhao L.J., Yang X.F., Liu D.J., Song Q., Zhang C.L., Xin W., Xiao Z.M. Quality differences between NILs of wheat variety Longmai 20 possessing HMW-GS 7^{OE} + 8* and 17 + 18. *Cereal Res. Commun*. 2020;48:493-498. DOI 10.1007/s42976-020-00052-x.

ORCID ID

A.A. Galimova orcid.org/0000-0002-7068-3359
A.R. Kuluev orcid.org/0000-0002-8563-1244
K.R. Ismagilov orcid.org/0000-0002-0212-116X
B.R. Kuluev orcid.org/0000-0002-1564-164X

Acknowledgements. Research by Galimova A.A. and Kuluev B.R. is supported by a grant from the Ministry of Science and Higher Education of the Russian Federation (agreement No. 075-15-2021-1066 dated September 28, 2021), works by Kuluev A.R. and Ismagilov K.R. are supported by the means of the federal budget, directed for performing the government assignment No. 122030200143-8 and FMRS-2022-0130.

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

Received October 25, 2022. Revised January 23, 2023. Accepted January 24, 2023.

Вавиловский журнал генетики и селекции. 2023;27(4):306-315
DOI 10.18699/VJGB-23-37

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

Morphological and phylogenetic features of the Crimean population of *Juniperus deltoides* R.P. Adams

A.O. Lantushenko[1], O.O. Korenkova[2] ✉, A.A. Syrovets[1], Ya.V. Meger[1], P.A. Korenkov[3], O.M. Shevchuk[2]

[1] Sevastopol State University, Sevastopol, Russia
[2] The Order of the Red Banner of Labour Nikitsky Botanical Gardens – National Scientific Center of the Russian Academy of Sciences, Yalta, Republik of the Crimea, Russia
[3] Moscow State University of Civil Engineering (National Research University), Moscow, Russia
✉ o.o.korenkova@mail.ru

Abstract. *Juniperus deltoides* is a relict species from the Tertiary Period. It is a typical representative of the Mediterranean group of the section *Juniperus*. It is included in the Red Books of the Republic of Crimea and the city of Sevastopol. Until recently, it was believed that a population of *J. oxycedrus* grew in Crimea. Currently, *J. deltoides* is described as a cryptic species, morphologically difficult to distinguish from *J. oxycedrus*. As a result, it became necessary to conduct a series of detailed studies to determine the morphological and phylogenetic features of the Crimean cryptic population in order to identify it as being one of the species of the cryptic pair. The studies were carried out in two stages: at the first stage, the morphological features of the vegetative and generative organs and their difference from *J. oxycedrus* were determined; the second stage included genetic research. The length of the needles of the Crimean population is 12.94 ± 0.19 mm, which corresponds to the Eastern Italian population of *J. deltoides*. At the same time, the width of the needles is 1.39 ± 0.02 mm, which is typical of the Portuguese population of *J. oxycedrus*. The dimensions of the cones are d_1 (conditional height) = 7.54 ± 0.14 mm, and d_2 (conditional width) = 9.11 ± 0.09 mm, which is more in line with *J. deltoides*. The shapes of the cones are very diverse. Some individuals have cones, the covering scales of which are visually indistinguishable, and their tops are completely fused. A similar phenomenon is characteristic of the Western Mediterranean populations of *J. oxycedrus*. Morphological analysis of the vegetative and generative organs of *J. deltoides* showed that when these two traits are combined, it is not possible to reliably distinguish between *J. deltoides* and *J. oxycedrus* individuals. Nuclear (ITS internal transcribed spacer) and chloroplast (petN-psbM, trnS-trnG) non-coding regions of the genome were used for genetic analysis. Studies have shown that the nuclear regions of genes have greater variability than chloroplast regions. The sequences obtained in this work formed a clade with *J. deltoides* samples 9430 and 9431 (BAYLU) growing in Turkey, which makes it possible to assign the samples studied to *J. deltoides.*

Key words: *Juniperus deltoids*; population; needles; cone berries; cryptic view; phylogenetic analysis; nuclear genes; chloroplast genes.

For citation: Lantushenko A.O., Korenkova O.O., Syrovets A.A., Meger Ya.V., Korenkov P.A., Shevchuk O.M. Morphological and phylogenetic features of the Crimean population of *Juniperus deltoides* R.P. Adams. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):306-315. DOI 10.18699/VJGB-23-37

Морфологические и филогенетические особенности крымской популяции *Juniperus deltoides* R.P. Adams

А.О. Лантушенко[1], О.О. Коренькова[2] ✉, А.А. Сыровец[1], Я.В. Мегер[1], П.А. Кореньков[3], О.М. Шевчук[2]

[1] Севастопольский государственный университет, Севастополь, Россия
[2] Ордена Трудового Красного Знамени Никитский ботанический сад – Национальный научный центр Росссийской академии наук, Ялта, Республика Крым, Россия
[3] Национальный исследовательский Московский государственный строительный университет, Москва, Россия
✉ o.o.korenkova@mail.ru

Аннотация. *Juniperus deltoides* – это реликтовый вид третичного периода. Является типичным представителем средиземноморской группы секции *Juniperus*. Включен в Красные книги Республики Крым и города Севастополя. До недавнего времени считалось, что на территории Крыма произрастает популяция *J. oxycedrus*. Сейчас *J. deltoides* описывают как криптический вид, морфологически сложно отличимый от *J. oxycedrus*. В результате возникла необходимость проведения ряда детальных исследований по определению морфологических и филогенетических особенностей криптической популяции Крыма с целью установления ее принадлежности к одному из видов криптической пары. Исследования проходили в два этапа: на первом этапе определяли морфологические особенности вегетативных и генеративных органов и их отличие от *J. oxycedrus*, на вто-

ром – выполняли генетические исследования. Длина хвои крымской популяции составляет 12.94 ± 0.19 мм, что соответствует восточноитальянской популяции *J. deltoides*. При этом ширина хвои равна 1.39 ± 0.02 мм, что свойственно португальской популяции *J. oxycedrus*. Размеры шишкоягод составляют: d_1 (условно высота) – 7.54 ± 0.14 мм, d_2 (условно ширина) – 9.11 ± 0.09 мм, что больше соответствует *J. deltoides*. Форма шишкоягод варьирует весьма сильно. Встречаются особи с шишкоягодами, кроющие чешуи которых визуально не отличимы, их верхушки полностью срастаются. Подобное явление характерно для западносредиземноморских популяций *J. oxycedrus*. Морфологический анализ вегетативных и генеративных органов *J. deltoides* показал, что при сочетании двух этих признаков достоверно различить особи *J. deltoides* и *J. oxycedrus* не представляется возможным. Для генетического анализа использовали ядерные (внутренний транскрибируемый спейсер ITS) и хлоропластные (petN-psbM, trnS-trnG) некодирующие участки генома. Исследования показали, что ядерные участки генов обладают большей вариабельностью, чем хлоропластные. Последовательности, полученные в данной работе, образовали кладу с образцами *J. deltoides* 9430 и 9431 (BAYLU), произрастающими в Турции, что позволяет отнести изученные образцы к виду *J. deltoides*.

Ключевые слова: *Juniperus deltoids*; популяция; хвоя; шишкоягоды; криптический вид; филогенетический анализ; ядерные гены; хлоропластные гены.

Introduction

The genus *Juniperus* L. is the largest in the cypress family (Cupressaceae Bartl.), assigned to the juniper subfamily (*Juniperoideae* Endl.), and includes 76 species (Pisarev, 2007; Adams, 2014b).

Junipers are distinguished by significant polymorphism, on the basis of which a number of subgenera, sections and series are distinguished in the genus. The taxonomy of the genus is based on two distinctive morphological features: the structure of generative organs (cones) and the structure of vegetative organs (needles) (Kolesnikov, 1974).

The genus was first described in 1700; since that time, the taxonomy of the genus has undergone significant changes (Novikov et al., 2014). Currently, the genus is divided into three sections, among which: the *Caryocedrus* section counting one species – *J. drupacea* Labill., the *Juniperus* section (synonymous with *Oxycedrus*), which includes 14 species, and the *Sabina* section, which consists of the remaining 61 species (Pisarev, 2007; Abaimov, 2009; Adams, 2014b). At the same time, the *Caryocedrus* section was often considered as a separate genus, but PCR studies conducted by R.P. Adams proved its common origin with the *Juniperus* section (Adams, 2014b).

Five species of junipers grow in the Crimea (*J. communis* L., *J. deltoides* R.P. Adams, *J. excelsa* M.-Bieb., *J. foetidissima* Willd., *J. sabina* L.), which belong to two sections – *Juniperus* and *Sabina*. All of them are included in the Red Books of the Republic of Crimea and the city of Sevastopol (Yena, Fateryga, 2015; Red Book…, 2018).

Juniperus deltoides is a relict species from the Tertiary period. It is a typical representative of the Mediterranean group of the *Juniperus* section. *J. deltoides* is common in the Mediterranean and the Middle East. To a large extent, its range is limited to the Mediterranean climate, but in the Balkans it occurs in more continental conditions. The northern border of its range passes in the Crimea. The area of the cryptic population of the Crimea, according to 2006 data, is 4843 hectares (Adams, 2014b; Plugatar, 2015; Farjon, 2017; Rajčević et al., 2020; Sadykova, Neshataeva, 2020; Yousefi et al., 2021).

Until recently, it was believed that a population of *J. oxycedrus* grows on the territory of Crimea. This species was included in the Guide to Higher Plants of Crimea (Rubtsov, 1972). However, the scientist R.P. Adams, on the basis of DNA sequencing, found that in most of the Mediterranean – from Italy to the east through Turkey to the mountains of the Caucasus and Iran (including the Crimea), a juniper other than *J. oxycedrus* is common, which he described as a new species – *J. deltoides*. At the same time, no direct analysis of the genetic material from the territory of Crimea was carried out. Adams made a similar conclusion based on the geographic localization of populations. In his works, he described that junipers growing west of Italy belong to the species *J. oxycedrus*, and junipers found to the east are *J. deltoides* (Adams et al., 2005; Adams, 2014a; Roma-Marzio et al., 2017).

Currently, *J. deltoides* is described as a cryptic species that is morphologically difficult to distinguish from *J. oxycedrus* (Adams et al., 2005; Adams, 2014a; Roma-Marzio et al., 2017). As a result, it became necessary to conduct a series of detailed studies to determine the morphological and phylogenetic features of the Crimean population of *J. deltoides*, in order to establish its belonging to one of the species of the cryptic pair.

The study includes two main tasks: determining the correspondence between the morphological features of the vegetative and generative organs of the cryptic population of the Crimea to the species *J. deltoides*; conducting genetic studies using nuclear and chloroplast regions of marker sequences.

Materials and methods

In order to conduct morphological and phylogenetic studies within the population, test areas of 0.2 hectares were laid at an altitude of 40 to 620 m above sea level, in various edaphoorographic conditions from Inkerman to Sudak (Fig. 1).

According to generally accepted methods, 10 model trees were identified within the test areas (Yarmishko, Lyanguzova, 2002). For each model tree, 30 cones were measured in two mutually perpendicular planes (conventional width and height). In addition, according to the determinant key developed by Adams for *J. deltoides* (Adams, 2014a), the degree of accretion of cone scales was visually determined.

To determine the parameters of the vegetative organs, the length and width of the needles were measured and the average error was determined (30 needles for each model tree). Then, a cross section of the needles was carried out in order to establish the presence or absence of curvature of the adaxial surface of the needles. The shape of the base of the needles was determined (Adams, 2014a, b).

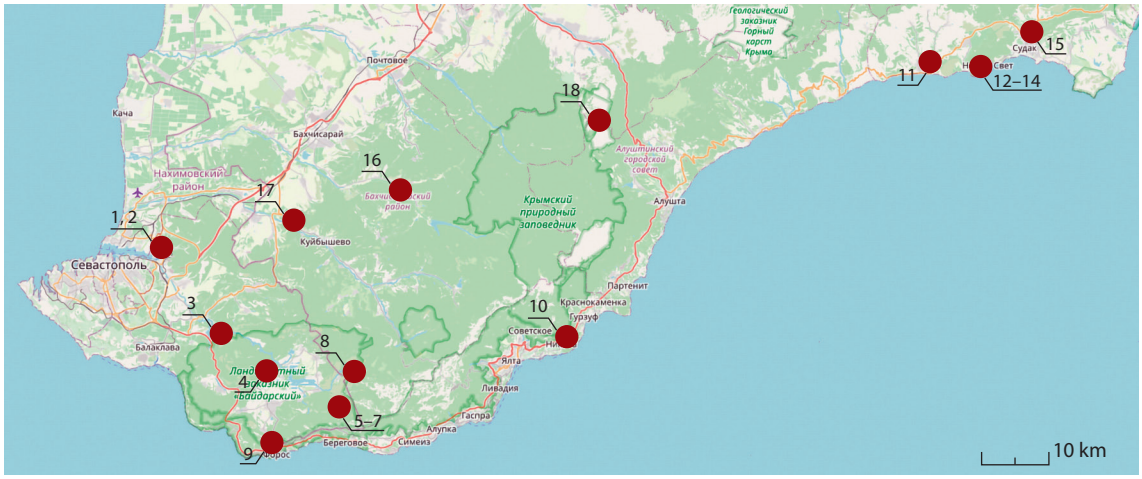


Fig. 1. Scheme of the location of sample plots in the cryptic populations of the Crimean Mountains.

1, 2 – the vicinity of the city of Inkerman; 3 – mt. Chirka-Kayasy; 4 – mt. Samnalykh; 5–7 – mt. Kara-Dag; 8 – mt. Tolaka- Bair; 9 – mt. Dragon; 10 – Cape Martyan; 11 – mt. Papaya-Kaya; 12, 13 – mt. Koba-Kaya; 14 – mt. Sokol; 15 – mt. Karshiters; 16 – Kullu-Kaya rocks; 17 – the vicinity of the village of Kudrino; 18 – mt. Chatyr-Dag.

Table 1. Nucleotide sequences of primers and PCR protocol used in this work (Hojjati et al., 2018)

Sequence name	Size, b.p.	Primer sequences	PCR Protocol
ITS	1100	F: GGAAGGAGAAGTCGTAACAAGG	1. 94 °C – 4 min 2. 34 cycles: 94 °C – 40 s 50 °C – 40 s 50 °C – 1.5 min 3. 72 °C – 7 min
		R: CTTTTCCTCCGCTTATTGATATG	
petN-psbM	764	F: AACGAAGCGAAAATCAATCA	
		R: AAAGAGAGGGATTCGTATGGA	
trnS-trnG	700	F: GCCGCTTTAGTCCACTCAGC	
		R: GAACGAATCACACTTTTACCAC	

18 samples of *J. deltoides* from different geographical locations of the Crimean peninsula were selected for genetic studies (see Fig. 1). DNA isolation from needles was carried out using the DNeasy Plant Mini Kit (Qiagen, Germany). The quantity and quality of the isolated DNA were analyzed using an Inplen nanophotometer (Germany). For PCR analysis, nuclear (internal transcribed spacer ITS) and chloroplast (petN-psbM, trnS-trnG) non-coding regions of the genome were used. Marker gene amplifications were performed using the universal primers and protocols described earlier (Table 1, Hojjati et al., 2018), using the ScreenMix reagent kit (Eurogen, Russia).

Sequencing of the obtained fragments was carried out on the genetic analyzer NANOPHOR-05 (Syntol, Russia) in the Resource Centre "Molecular Structure of Matter". Electrophoregrams unsuitable for analysis were obtained for two samples during sequencing of the ITS nuclear fragment, and the nucleotide sequences of 16 samples of the Crimean population were further studied (Table 2). The obtained sequences of ITS, petN-psbM and trnS-trnG were compared with those available in the database of the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/). The samples for comparison were taken from (Hojjati et al., 2018) and are indicated in Table 2. The alignment of nucleotide sequences for each marker site and their integration into the combined matrix was carried out in the MegaX program (Kumar et al., 2018). Phylogenetic reconstruction was performed using the Bayesian method implemented in MrBayes version 3.2.6 (Ronquist et al., 2012).

Haplotype diversity (HD), nucleotide diversity (π), and the number of nucleotide substitutions (M) were calculated for each species using DnaSP 6.0 (Rozas et al., 2017).

The relationships between haplotypes of sequences from three marker sites were reconstructed by the TCS method implemented in the PopArt program (Bandelt et al., 1999).

Results and discussion

In 2014, R.P. Adams (2014a) developed and published a determinant key for *J. deltoides*, which makes it possible to distinguish individuals of this species from *J. oxycedrus*. According to the determinant, the maximum needle length of *J. deltoides* is less than that of *J. oxycedrus* and equals 13.0 mm (for *J. oxycedrus* it is 15.0 mm). The length of the needles of the Crimean population of *J. deltoides* is 12.94 ± 0.19 mm, which corresponds to the dimensions declared by Adams. At the same time, a significant number of

Table 2. Codes, geographical location and Genbank codes of the analyzed samples

Sample code / species*	Location	Sequences	Genbank code
CRMD1	mt. Chirka-Kayasy	ITS	OP303319
		petN-psbM	OP321223
		trnS-trnG	OP321239
CRMD4	mt. Samnalykh	ITS	OP303320
		petN-psbM	OP321224
		trnS-trnG	OP321240
CRMD7	mt. Kara-Dag	ITS	OP303321
		petN-psbM	OP321225
		trnS-trnG	OP321241
CRMD10		ITS	OP303322
		petN-psbM	OP321226
		trnS-trnG	OP321242
CRMD14		ITS	OP303323
		petN-psbM	OP321227
		trnS-trnG	OP321243
CRMD16	mt. Tolaka-Bair	ITS	OP303324
		petN-psbM	OP321228
		trnS-trnG	OP321244
CRMD19	mt. Papaya-Kaya	ITS	OP303325
		petN-psbM	OP321229
		trnS-trnG	OP321245
CRMD22	mt. Sokol	ITS	OP303326
		petN-psbM	OP321230
		trnS-trnG	OP321246
CRMD25	mt. Koba-Kaya	ITS	OP303327
		petN-psbM	OP321231
		trnS-trnG	OP321247
CRMD27		ITS	OP303328
		petN-psbM	OP321232
		trnS-trnG	OP321248
CRMD28		ITS	OP303329
		petN-psbM	OP321233
		trnS-trnG	OP321249
CRMD29		ITS	OP303330
		petN-psbM	OP321234
		trnS-trnG	OP321250
CRMD33	mt. Karshiters	ITS	OP303331
		petN-psbM	OP321235
		trnS-trnG	OP321251
CRMD35	The vicinity of the city of Inkerman	ITS	OP303332
		petN-psbM	OP321236
		trnS-trnG	OP321252
CRMD38		ITS	OP303333
		petN-psbM	OP321237
		trnS-trnG	OP321253
CRMD40	mt. Chatyr-Dag	ITS	OP303334
		petN-psbM	OP321238
		trnS-trnG	OP321254
BAYLU 13730* *J. communis*	Bulgaria	ITS	LC420860
		petN-psbM	LC420856
		trnS-trnG	LC420859

Table 2 (end)

Sample code / species*	Location	Sequences	Genbank code
BAYLU 8765* *J. communis*	Armenia	ITS	LC420870
		petN-psbM	LC420866
		trnS-trnG	LC420869
BAYLU 9431* *J. deltoides*	Turkey	ITS	LC420895
		petN-psbM	LC420891
		trnS-trnG	LC420894
BAYLU 8795* *J. drupacea*	Greece	ITS	LC420930
		petN-psbM	LC420926
		trnS-trnG	LC420929
BAYLU 8796* *J. drupacea*		ITS	LC420935
		petN-psbM	LC420931
		trnS-trnG	LC420934
BAYLU 8785* *J. excelsa*		ITS	LC420800
		petN-psbM	LC420796
		trnS-trnG	LC420799
BAYLU 9433* *J. excelsa*	Turkey	ITS	LC420850
		petN-psbM	LC420846
		trnS-trnG	LC420849
BAYLU 5645* *J. foetidissima*	Greece	ITS	LC420900
		petN-psbM	LC420896
		trnS-trnG	LC420899
BAYLU 9039* *J. oxycedrus*	France	ITS	LC420880
		petN-psbM	LC420876
		trnS-trnG	LC420879
BAYLU 9040* *J. oxycedrus*		ITS	LC420885
		petN-psbM	LC420881
		trnS-trnG	LC420884
TARI IRN30492* *J. oxycedrus*	Iran	ITS	LC420974
		petN-psbM	LC420970
		trnS-trnG	LC420973
BAYLU 14171* *J. polycarpos* var. *polycarpos*	Azerbaijan	ITS	LC420790
		petN-psbM	LC420786
		trnS-trnG	LC420789
BAYLU 8757* *J. polycarpos* var. *turcomanica*	Turkmenistan	ITS	LC420690
		petN-psbM	LC420686
		trnS-trnG	LC420689
BAYLU 14316* *J. sabina*	Azerbaijan	ITS	LC420910
		petN-psbM	LC420906
		trnS-trnG	LC420909
BAYLU 9430* *J. deltoides*	Turkey	ITS	LC420890
		petN-psbM	LC420886
		trnS-trnG	LC420889
BAYLU 14317* *J. sabina*	Azerbaijan	ITS	LC420920
		petN-psbM	LC420916
		trnS-trnG	LC420919
TARI 17035* *J. foetidissima*	Azerbaijan, Iran	ITS	LC420989
		petN-psbM	LC420985
		trnS-trnG	LC420988

* The samples taken for comparison from (Hojjati et al., 2018), their species are indicated.

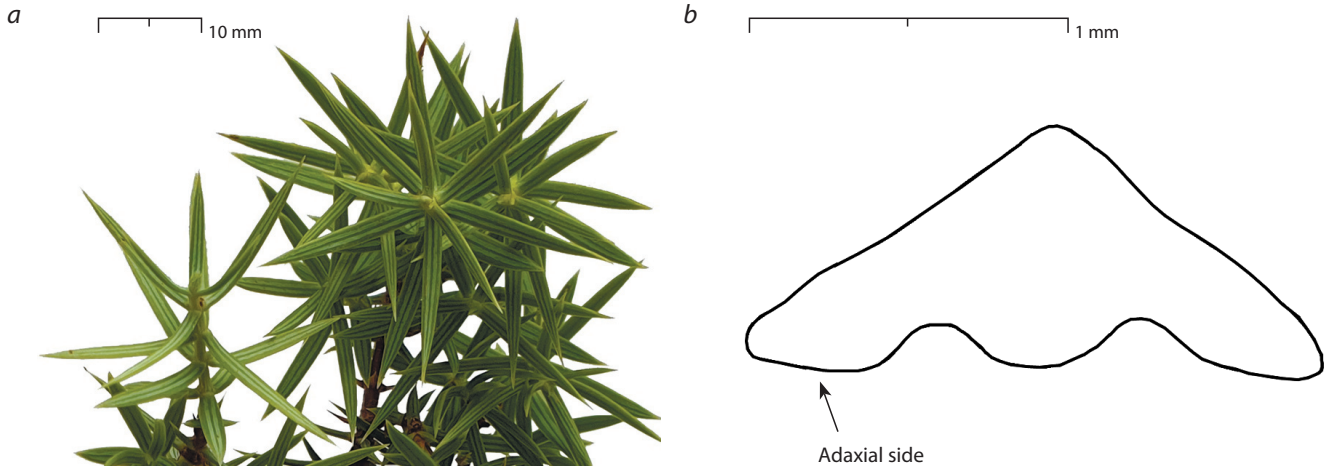


Fig. 2. Needles of the Crimean cryptic population (localization – mt. Tolaka-Bair):
a, general view; *b*, a schematic representation of the curvature of the adaxial surface of the needles.

Fig. 3. Morphological heterogeneity of cones of the cryptic population of Crimea.

individuals were found (in the vicinity of the city of Inkerman and the village of Kudrino; on the mountains of Kara-Dag, Koba-Kaya, Dragon, on the rocks of Kullu-Kaya and on Cape Martyan), the length of the needles of which is from 18 to 20 mm. These individuals aroused the greatest interest for further research. Needle width, according to Adams (2014a), is also a defining feature. Under the conditions of the Crimea, this indicator is 1.39±0.02 mm, which corresponds to the western group of junipers, namely *J. oxycedrus*.

The needles had no differences in color and shape of the base. All needles are light green with a deltoid base, which is typical for *J. deltoides* (Fig. 2, *a*). At the same time, the cross section of the needles showed that a significant part of the individuals (34 %) are characterized by the curvature of the adaxial surface of the needles (see Fig. 2, *b*). According to Adams (2014a), this is a distinctive feature of *J. oxycedrus*. The length of the needles with this type of stomatal bands is 11.87±0.24 mm. Thus, it was found that the morphological features of the needles of the Crimean population of *J. deltoides* simultaneously exhibit signs of both *J. deltoides* and *J. oxycedrus*. On the basis of the identified features of the vegetative organs, it is not possible to attribute individuals to one of the species.

The second distinctive morphological feature of junipers are cones. In the case of the cryptic pair, *J. deltoides*/*J. oxycedrus*, the main role in determining the species is played by the inosculation of cone scale tips, and to a lesser extent, by the size and color of the cone of cones, the presence or absence of plaque on them.

The cones of the Crimean cryptic population have a significant number of morphological variants (Fig. 3). At the same time, the coloration is almost the same for all of them, and the smoke-blue coating appears to varying degrees, regardless of the shape of the cones. The shape, in turn, differs very much (from spherical to triangular). There are individuals with cones, the covering cone scale tips of which are visually indistinguishable and their tops are completely fused. A similar phenomenon is characteristic of *J. oxycedrus*. At the same time, the needles of these individuals are defined as the needles of *J. deltoides*. The second type of cones is almost triangular due to clearly visible three covering cone scale tips (characteristic of *J. deltoides*). The needles of these individuals show signs of both species. To a greater extent, there are intermediate variants, in which the bases of the cone scale tips grow together, and their tops move away from each other to varying degrees.

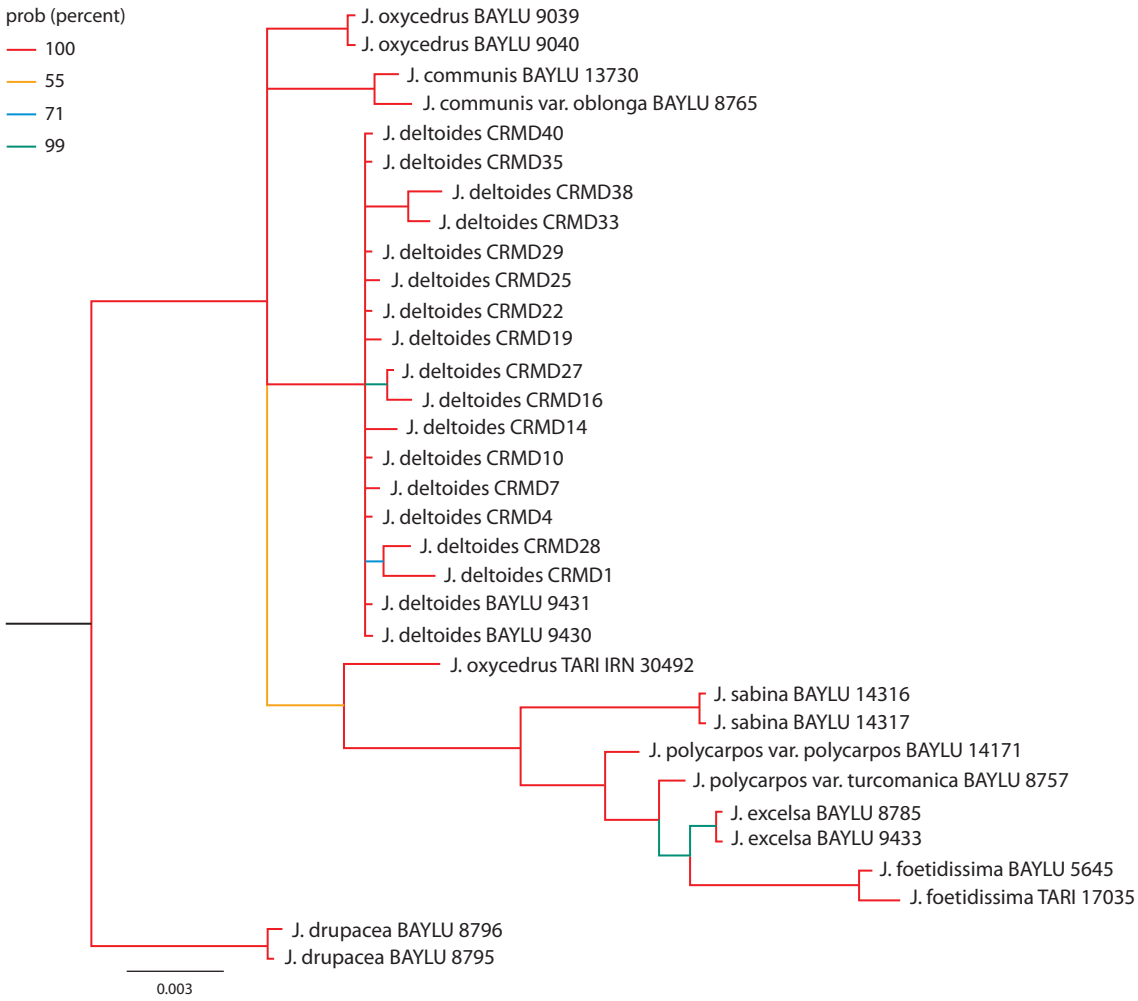


Fig. 4. A phylogenetic tree constructed by the Bayes method based on a combined sequence including nuclear (ITS) and chloroplast (petN-psbM, trnS-trnG) non-coding regions of the genome.

Node support values are shown in color.

The size of the cones turned out to be the most stable trait and differed within the error on all test areas. The average sizes of cones are: d_1 (conditional height) is 7.54±0.14 mm, and d_2 (conditional width) is 9.11±0.09 mm, which is more consistent with the Turkish population of *J. deltoides*.

Thus, the morphological analysis of the vegetative and generative organs of *J. deltoides* showed that when these two characters are combined; it is not possible to reliably distinguish between individuals of *J. deltoides* and individuals of *J. oxycedrus*. *J. oxycedrus* is a basal species, while *J. deltoides* is a cryptic one. As a result, the issue of conducting phylogenetic studies is especially acute for determining the systematic belonging of the Crimean population to one of the species.

For phylogenetic analysis, the nucleotide sequences of three marker sites (ITS, petN-psbM and trnS-trnG) of 16 Crimean samples and 17 samples from the work (Hojjati et al., 2018) were used (see Table 2).

Phylogenetic trees constructed from individual marker sequences are presented in Supplementary Materials 1–3[1]. The use of ITS and petN-psbM marker sequences allowed us to obtain topologies where each species forms a separate clade and their phylogenetic definition is unambiguous. The topology obtained by analyzing the trnS-trnG sequences does not allow separating the species of *J. communis* and *J. deltoids*.

A phylogenetic tree was also constructed taking into account all the nucleotide fragments studied in this work (Fig. 4). It can be seen that the samples of each species form clades with high (more than 75 %) support. The exception is the specimen *J. oxycedrus* TARI IRN 30492, for which the phylogenetic definition is not unambiguous: according to the sequences trnS-trnG and petN-psbM, it forms a clade with samples of the species *J. deltoids*, and according to the fragment of ITS and the analysis of the combined sequences, it forms a clade with species of the section *Sabina*. All the samples from the Crimea studied in this work formed a clade with sequences of the species *J. deltoides* 9430 and 9431 (BAYLU) growing in Turkey. In this clade, three pairs of samples (CRMD1 and CRMD28, CRMD16 and CRMD27, CRMD33 and CRMD38) form separate branches with high support, but there is no correlation between them by phenotypes and geographical location.

[1] Supplementary Materials 1–3 are available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Lantushenko_Engl_27_4.pdf

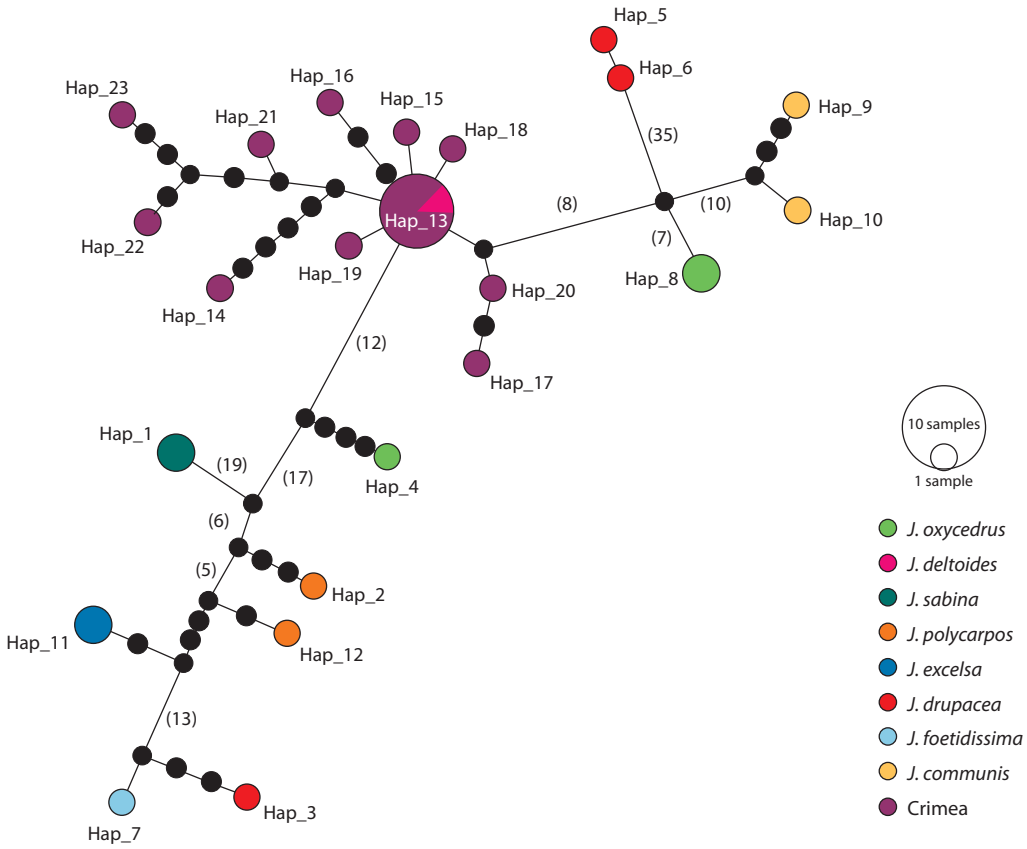


Fig. 5. Haplotypic network constructed by the TCS method.

The color indicates belonging to the species, the size of the circle is the number of samples in the haplotype, small distances are represented as dots (1 point – 1 replacement), for large ones the number of substitutions is given in brackets.

Table 3. Characteristics of the nucleotide sequences studied in this work

Parameter	Length	H	Hp	π	M
ITS	1295	10	0.867	0.00419	14
petN-psbM	896	4	0.442	0.00140	7
trnS-trnG	785	3	0.242	0.0005	3
Combined sequence	2976	11	0.875	0.002	24

Note. H is the number of haplotypes, Hp is the haplotypic diversity, π is the nucleotide diversity, M is the number of mutations.

The haplotypic network constructed from the nuclear and chloroplast regions of the genome for the samples listed in Table 2 is shown in Figure 5. It can be seen that the Crimean population of *J. deltoides* is characterized by a quite large number of haplotypes: 11.

In the previously studied populations of juniper trees of other species (*J. excelsa*, *J. polycarpos*, and *J. foetidissima*) on the northern border of the distribution area, in the Crimea, the Caucasus and Dagestan (Sadykova et al., 2021), much smaller haplotypic diversity was found: sequences of 17 *J. excelsa* samples form two haplotypes, 16 *J. foetidissima* samples form four haplotypes, 15 samples of *J. polycarpos* – one haplotype.

The genetic variability of nuclear and chloroplast regions of genes was analyzed. The analysis of the parameters presented in Table 3 allows us to conclude that the greatest variability is characteristic of the ITS nuclear fragment, and the least variability is characteristic of the trnS-trnG fragment. The greatest variability of the nuclear sites of marker nucleotide sequences is characteristic of other juniper species (Mao et al., 2010; Hojjati et al., 2018).

It follows from Figure 5 that the nucleotide sequences of CRMD4, CRMD10, CRMD22, CRMD29, CRMD35, CRMD40 samples formed a common haplotype with the sequences BAYLU:9430 and BAYLU:9431.

The sequences obtained in this work formed a clade with *J. deltoides* 9430 and 9431 (BAYLY) specimens growing in Turkey (Table 4). Thus, the analysis carried out in this work allows us to attribute the studied samples to the species *J. deltoides*.

Table 4. List of haplotypes of nucleotide sequences of samples of the Crimean population and samples taken for comparison from the work (Hojjati et al., 2018)

Haplotype	Species
Hap_1	*Juniperus sabina* BAYLU:14317, BAYLU:14316
Hap_2	*Juniperus polycarpos* var. *polycarpos* BAYLU:14168
Hap_3	*Juniperus foetidissima* TARI 17035
Hap_4	*Juniperus oxycedrus* TARI IRN: 30492
Hap_5	*Juniperus drupacea* BAYLU:8796
Hap_6	*Juniperus drupacea* BAYLU:8795
Hap_7	*Juniperus foetidissima* BAYLU:5645
Hap_8	*Juniperus oxycedrus* BAYLU:9040, BAYLU:9039
Hap_9	*Juniperus communis* var. *oblonga* BAYLU:8765
Hap_10	*Juniperus communis* BAYLU:13730
Hap_11	*Juniperus excelsa* BAYLU:9433, BAYLU:8785
Hap_12	*Juniperus polycarpos* var. *turcomanica* BAYLU:8757
Hap_13	*Juniperus deltoides* BAYLU:9430, BAYLU:9431,CRMD4, CRMD10, CRMD22, CRMD29, CRMD35, CRMD40
Hap_14	*Juniperus deltoides* CRMD1
Hap_15	*Juniperus deltoides* CRMD7
Hap_16	*Juniperus deltoides* CRMD14
Hap_17	*Juniperus deltoides* CRMD16
Hap_18	*Juniperus deltoides* CRMD19
Hap_19	*Juniperus deltoides* CRMD25
Hap_20	*Juniperus deltoides* CRMD27
Hap_21	*Juniperus deltoides* CRMD28
Hap_22	*Juniperus deltoides* CRMD33
Hap_23	*Juniperus deltoides* CRMD38

Conclusion

Based on the studies of vegetative organs, it was found that the length of the needles of individuals of the Crimean cryptic population is 12.94 ± 0.19 mm, which is typical for this species. At the same time, there are individuals with needles of much greater length (18–20 mm). The cross section of the needles, regardless of its length, in 34 % of cases shows signs of *J. oxycedrus*, expressed in the curvature of its adaxial surface.

Cones manifest significant morphological heterogeneity. Their shape varies from spherical to triangular, depending on the degree of fusion of the covering cone scale tips. At the same time, it was found that the same individuals can simultaneously show signs of *J. deltoides* (by vegetative organs) and signs of *J. oxycedrus* (by generative organs).

Thus, we can conclude that these morphological features are not reliable for determining the systematic affiliation of individuals. Thus, the only possible way to determine it is to conduct genetic research.

Phylogenetic analysis has shown that nuclear regions of genes have greater variability than chloroplast ones. The sequences obtained in this work for the Crimean population formed a clade with samples of *J. deltoides* 9430 and 9431 (BAYLU) growing in Turkey, which makes it possible to attribute the studied samples to the species *J. deltoides*.

References

Abaimov V.F. Dendrology. Moscow, 2009. (in Russian)

Adams R.P. Morphological comparison and key to *Juniperus deltoides* and *J. oxycedrus. Phytologia.* 2014a;96(2):58-62.

Adams R.P. Junipers of the World: The Genus *Juniperus*. Bloomington: Trafford Publ., 2014b.

Adams R.P., Morris A.J., Pandey R.N., Schwarzbach A.E. Cryptic speciation between *Juniperus deltoides* and *Juniperus oxycedrus* (*Cupressaceae*) in the Mediterranean. *Biochem. Syst. Ecol.* 2005;33(8): 771-787. DOI 10.1016/j.bse.2005.01.001.

Bandelt H.J., Forster P., Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 1999;16(1):37-48. DOI 10.1093/oxfordjournals.molbev.a026036.

Farjon A. A Handbook of the World's Conifers. Brill, 2017.

Hojjati F., Kazempour-Osaloo S.H., Adams R.P., Assadi M. Molecular phylogeny of *Juniperus* in Iran with special reference to the *J. excelsa* complex, focusing on *J. seravschanica*. *Phytotaxa.* 2018; 375(2):135-157. DOI 10.11646/phytotaxa.375.2.1.

Kolesnikov A.I. Ornamental Dendrology. Moscow, 1974. (in Russian)

Kumar S., Stecher G., Li M., Knyaz M., Tamura C. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol*. 2018;35(6):1547-1549. DOI 10.1093/molbev/msy096.

Mao K., Hao G., Liu J., Adams R.P., Milne R.I. Diversification and biogeography of *Juniperus* (*Cupressaceae*): variable diversification rates and multiple intercontinental dispersals. *New Phytol*. 2010; 188(1):254-272. DOI 10.1111/j.1469-8137.2010.03351.x.

Novikov O.O., Pisarev D.I., Zhilyakova E.T., Trifanov B.V. Juniper: Phytochemistry and pharmacology of the genus *Juniperus* L. Moscow, 2014. (in Russian)

Pisarev D.I. Current state of research in the taxonomy, chemistry, and pharmacology of the genus *Juniperus* L. In: Proceedings of the international scientific and practical conference "Current Issues in Phytodesign". Belgorod, 2007;296-304. (in Russian)

Plugatar Yu.V. Forests of Crimea. Simferopol, 2015. (in Russian)

Rajčević N., Dodoš T., Novaković J., Janaćković P., Marin P.D. Epicuticular wax variability of *Juniperus deltoids* R.P. Adams from the central Balkan – ecology and chemophenetics. *Biochem. Syst. Ecol*. 2020;89:104008. DOI 10.1016/j.bse.2020.104008.

Red Book of the City of Sevastopol. Sevastopol, 2018. (in Russian)

Roma-Marzio F., Najar B., Alessandri J., Pistelli L., Peruzzi L. Taxonomy of prickly juniper (*Juniperus oxycedrus* group): a phytochemical-morphometric combined approach at the contact zone of two cryptospecies. *Phytochemistry*. 2017;141:48-60. DOI 10.1016/j.phytochem.2017.05.008.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 2012;61(3):539-542. DOI 10.1093/sysbio/sys029.

Rozas J., Ferrer-Mata A., Sánchez-DelBarrio J.C., Guirao-Rico S., Librado P., Ramos-Onsins S.E., Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol*. 2017;34(12):3299-3302. DOI 10.1093/molbev/msx248.

Rubtsov N.I. Guide to Higher Plants of Crimea. Leningrad, 1972. (in Russian)

Sadykova G.A., Hantemirova E.V., Polezhaeva M.A., Aliev K.U. Genetic variability of tree junipers of section Sabina: data from Dagestan, Armenia, and Crimea. *Russ. J. Genet*. 2021;57:1223-1228. DOI 10.1134/S1022795421100100.

Sadykova G.A., Neshataeva V.Yu. Woodlands of *Juniperus excelsa* subsp. *polycarpos* in foothill Dagestan. *Botanicheskij Zhurnal = Botanical Journal*. 2020;105(2):179-195. DOI 10.31857/S0006813619110164. (in Russian)

Yarmishko V.T., Lyanguzova I.V. Methods for Studying Forest Communities. St. Petersburg, 2002. (in Russian)

Yena A.V., Fateryga A.V. Red Book of the Republic of Crimea: Plants, algae, and fungi. Simferopol, 2015. (in Russian)

Yousefi S., Avand M., Yariyan P. Identification of the most suitable afforestation sites by *Juniperus excels* specie using machine learning models: Firuzkuh semi-arid region, Iran. *Ecol. Inform.* 2021;65:101427. DOI 10.1016/j.ecoinf.2021.101427.

ORCID ID

Lantushenko A.O. orcid.org/0000-0003-3614-9387
Korenkova O.O. orcid.org/0000-0001-6482-7312
Korenkov P.A. orcid.org/0000-0003-1847-4303
Shevchuk O.M. orcid.org/0000-0002-6565-7930

Acknowledgements. Genetic studies were carried out at the Center for Collective Use "Molecular Structure of Matter" with the support of the Priority-2030 program of the Sevastopol State University (strategic project No. 3, No. 121121700318-1).

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

Received September 12, 2022. Revised February 3, 2023. Accepted February 28, 2023.

Вавиловский журнал генетики и селекции. 2023;27(4):316-322
DOI 10.18699/VJGB-23-38

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

Genetic structure of the population of wild-growing vines of the Utrish Nature Reserve

E.T. Ilnitskaya[1] ✉, M.V. Makarkina[1], I.V. Gorbunov[2], I.V. Stepanov[1], T.D. Kozina[1], E.A. Kozhevnikov[1], V.K. Kotlyar[1]

[1] North Caucasian Federal Scientific Center of Horticulture, Viticulture, Wine-making, Krasnodar, Russia
[2] Anapa Zonal Experimental Station of Viticulture and Wine-making – Branch of North Caucasian Federal Scientific Center of Horticulture, Viticulture, Wine-making, Anapa, Russia
✉ ilnitskaya79@mail.ru

Abstract. Grapes are one of the most common agricultural crops in the world. Currently, the analysis of genotypes directly at the DNA level is considered to be the most accurate method for studying the plant gene pool. The study of wild vines and ancient varieties in various regions of viticulture is an important direction of research in this field. The purpose of this work was to study the population of wild grapes growing on the territory of the Utrish Nature Reserve on the Black Sea coast of Krasnodar Region. The territory of the reserve is of interest as it is a site of ancient settlements, and the environmental conditions are suitable for the growth of wild grapes. During the survey of the territory, 24 samples of wild grapes were found, which were described according to the main morphological characteristics and analyzed by the molecular genetic method. The found vines were genotyped using 15 DNA markers, including nine commonly used for DNA fingerprinting (VVS2, VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VrZAG62, VrZAG79) and VVIb23, which allows determining hermaphrodite and dioecious vines. Statistical processing of microsatellite loci polymorphism data was carried out using the GenAlEx 6.5 program. The genetic relationships of the studied vines were evaluated using the PAST 2.17c program. The samples were found to be morphologically and genetically polymorphic. The number of alleles identified in the sample varied from 5 to 18 and averaged 8 alleles per locus. Statistical processing of DNA analysis data made it possible to identify two genetically different populations among the wild discovered vines. An assessment of genetic similarity of the found vines with some local varieties of geographically close viticulture regions, rootstocks and representatives of *Vitis sylvestris* from other territories was made. One of the populations found in the Utrish Nature Reserve is close to a number of *V. sylvestris* genotypes, the DNA profiles of which are presented in the *Vitis* International Variety Catalogue.
Key words: wild-growing vines; *Vitis sylvestris*; DNA profiling; genetic polymorphism; SSR loci.

For citation: Ilnitskaya E.T., Makarkina M.V., Gorbunov I.V., Stepanov I.V., Kozina T.D., Kozhevnikov E.A., Kotlyar V.K. Genetic structure of the population of wild-growing vines of the Utrish Nature Reserve. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):316-322. DOI 10.18699/VJGB-23-38

Генетическая структура популяции дикорастущих форм винограда заповедника «Утриш»

Е.Т. Ильницкая[1] ✉, М.В. Макаркина[1], И.В. Горбунов[2], И.В. Степанов[1], Т.Д. Козина[1], Е.А. Кожевников[1], В.К. Котляр[1]

[1] Северо-Кавказский федеральный научный центр садоводства, виноградарства, виноделия, Краснодар, Россия
[2] Анапская зональная опытная станция виноградарства и виноделия – филиал Северо-Кавказского федерального научного центра садоводства, виноградарства, виноделия, Анапа, Россия
✉ ilnitskaya79@mail.ru

Аннотация. Виноград – одна из самых распространенных сельскохозяйственных культур в мире. В настоящее время наиболее точным методом для изучения генофонда растений считается анализ генотипов на уровне ДНК. Дикорастущие формы и древние сорта различных регионов виноградарства – актуальные объекты исследований в данной области. Целью работы было изучение популяции дикорастущего винограда, произрастающей на территории заповедника «Утриш» на Черноморском побережье Краснодарского края. Территория заповедника представляет интерес для подобного рода исследований, так как является местом древних поселений, а экологические условия пригодны для произрастания диких форм винограда. В процессе обследования территории обнаружено 24 образца дикорастущего винограда, которые были описаны по основным морфологическим характеристикам и проанализированы молекулярно-генетическим методом. Найденные формы генотипированы с помощью 15 ДНК-маркеров, в том числе девяти общепринятых для ДНК-паспортизации сортов винограда (VVS2, VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VrZAG62, VrZAG79) и маркера VVIb23, позволяющего определять обоеполые и двудомные формы. Статистическая обработка данных полиморфизма микросателлитных локусов выполнена в программе GenAlEx 6.5. Генетические взаимосвязи исследуемых форм винограда оценивали в программе PAST 2.17c. В результате обнаружено, что образцы характеризуются мор-

фологическим и генетическим полиморфизмом. Количество идентифицированных аллелей в выборке варьировало от 5 до 18 и составило в среднем 8 аллелей на локус. Проведение статистической обработки данных ДНК-анализа позволило выявить две генетически различные популяции среди обнаруженных дикорастущих форм. Оценено генетическое сходство найденных форм с некоторыми аборигенными сортами географически близких регионов виноградарства, подвойными сортами и представителями *Vitis sylvestris* с других территорий. Одна из обнаруженных в заповеднике «Утриш» популяций близка к ряду генотипов *V. sylvestris*, ДНК-профили которых представлены в международной базе данных *Vitis* International Variety Catalogue.
Ключевые слова: дикорастущие формы винограда; *Vitis sylvestris*; ДНК-профилирование; генетический полиморфизм; SSR-локусы.

Introduction

Grapes are one of the most widespread agricultural crops in the world. The most significant, both economically and sociohistorically, is the species *Vitis vinifera* L. This species of the genus *Vitis* L. (family Vitaceae), originating from Eurasia, supposedly appeared about 65 million years ago (This et al., 2006). Currently, two subspecies are distinguished within the species, *V. vinifera* L. subsp. *sylvestris* (Gmel.) Hegi, which includes wild populations, and *V. vinifera* subsp. *sativa* (DC.) Hegi (or subsp. *vinifera*), which includes cultivars.

Wild and cultivated vines differ in a number of features, including their reproductive biology: wild grapevines are dioecious and cross-pollinate, while the cultivated vines are mostly hermaphrodite and self-pollinating. The domestication of grapes, which occurred about 8 thousand years ago, is closely associated with the emergence of winemaking, although it is still not known for certain which process preceded the other. The Middle East and the Caucasus are considered to be the initial centers of domestication of *V. vinifera*. The earliest evidence of wine production from 7,400–7,000 BC is found in Iran (McGovern, 2004). Seeds of domesticated grapes, about 8,000 years old, have also been found in Georgia and Turkey. However, Neolithic seed remains found in Western Europe also suggest grape exploitation during this time, and wild form seed remains have been found at Bronze Age sites in France (This et al., 2006).

Wild grape populations are currently represented by wild vines of *V. vinifera* cultivars and a wild subspecies. The study of wild grapevines in the ancient regions of grape cultivation has been actively conducted in recent years at the molecular genetic level (Doulati-Baneh et al., 2015; Gorislavec et al., 2017; De Michele et al., 2019; Margaryan et al., 2019; Cunha et al., 2020; Zdunić et al., 2020; Kupe et al., 2021; Lukšić et al., 2022).

The study of the local gene pools of various viticulture regions (including native varieties and wild specimens) at the DNA level makes it possible to more fully assess the genetic diversity of varieties and vines, to identify closer and more distant genotypes. Wild exemplars of agricultural crops are also significant for breeding as unique sources of genetic variability (Ellstrand et al., 2010).

The territory of the Utrish Nature Reserve on the Black Sea coast of the Krasnodar Territory is of interest for this kind of research, since it is a site of ancient settlements, and the ecological conditions of the territory are suitable for the growth of wild vines (Chernodubov, Rudenok, 2015).

Materials and methods

Expeditionary research to search for wild grapevines, study the ecological conditions of their habitats and morpho-biological features was carried out for three years (2019–2021) on the territory of the Utrish Nature Reserve (Krasnodar Territory). The reserve is located in the northwestern part of the Black Sea coast of the Western Caucasus, on the Abrau Peninsula. The climate is sub-Mediterranean, moderately warm.

For molecular genetic analysis, 24 samples of wild-growing grapes were selected. DNA samples were isolated from shoot apical parts of vine plants by a method based on the use of CTAB (cytyltrimethylammonium bromide) (Rogers, Bendich, 1985).

Genotyping was performed at 15 microsatellite loci, 9 of which are standard for DNA fingerprinting of grape varieties (VVS2, VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VrZAG62, VrZAG79) (This et al., 2004; This, 2007). DNA markers linked to grape pathogens resistance loci were also included in the study – downy mildew resistance: UDV305, UDV737 (*Rpv3*) and GF09-46 (*Rpv10*) and powdery mildew resistance: ScORGF15-02 (*Ren3*), CenGen6 (*Ren9*) (Di Gaspero et al., 2012; Schwander et al., 2012; van Heerden et al., 2014; Zendler et al., 2017). All found specimens were also analyzed with the VVIb23 marker, which makes it possible to detect hermaphrodite and dioecious samples (Merdinoglu et al., 2005; Riaz et al., 2013).

Polymerase chain reaction (PCR) was carried out using an Eppendorf MasterCycler nexus GX2 device (Germany) according to the following scheme: 5 minutes at +95 °C (initial denaturation); 35 cycles: 10 seconds at +95 °C (denaturation), primer annealing for 30 seconds at +55 °C for VVS2, VVMD5, VVMD7, VVMD27, UDV305, UDV737, CenGen6, VVIb23, at +58 °C for VrZAG62, VrZAG79, ScORGF15-02, at +60 °C for VVMD25, VVMD28, VVMD32, GF09-46, 30 seconds at +72 °C (elongation); the final cycle was 15 minutes at +72 °C. PCR mixture with a total volume of 20 µl contained: 50 ng of genomic DNA, 1.5 units of Taq polymerase, 1x buffer for Taq polymerase with ammonium sulfate and magnesium, 2 mM $MgCl_2$, 0.2 mM of each dNTP (deoxynucleotide triphosphates) (SibEnzyme-M, Moscow) and 200 µM of each of the primers (Sintol, Moscow). The results were visualized by capillary electrophoresis using a Nanofor 05 genetic analyzer (Institute of Analytical Instrumentation, Russian Academy of Sciences, St. Petersburg, Russia) and a special built-in software package.

The following varieties were used as reference genotypes: Pinot noir (for markers VVS2, VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VrZAG62, VrZAG79), Saperavi Severniy (GF09-46), Villard blanc (UDV305, UDV737), Regent (ScORGF15-02, CenGen6), Kishmish Vatkana (VVIb23), since the allelic compositions in the DNA of these varieties for the analyzed loci are known.

Statistical processing of microsatellite loci polymorphism data was carried out using the GenAlEx 6.5 program (Pea-

kall, Smouse, 2012). The genetic relationships of the studied grapevines were assessed using the PAST 2.17c program using the method of principal coordinates (PCoA) (Hammer et al., 2001).

In order to study the genetic similarity of the studied wild-growing samples with the indigenous gene pool of grapes, the work included the sampling of DNA profiles of cultivars according to nine standard SSR loci from the international database *Vitis* International Variety Catalog (*V*IVC), which belong to the local forms of Dagestan (Republic of Dagestan, Russian Federation), Don (Rostov region, Russian Federation), Georgia, Crimea (Republic of Crimea, Russian Federation), as well as to rootstock grape cultivars (the greatest contribution to the genotype of which was made by different North American species) and *V. sylvestris* genotypes from different geographical zones. Bayesian analysis was carried out in the Structure 2.3.4 program using 65 genotypes (41 varieties from the *V*IVC database and 24 genotypes of the studied forms) with the following parameters: 500000 Burn-in period, 500000 Reps, K=7 (Pritchard et al., 2000).

The study was carried out using the instrument park of the Center for Collective Use of Technological Equipment in the direction of "Genomic and postgenomic technologies" of the North Caucasian Federal Scientific Center for Horticulture, Viticulture, Wine-making.

Results and discussion

In the process of surveying the territory of the Utrish Nature Reserve, 24 samples of wild-growing grapes were found, the description of which according to the main morphological features is given in Table 1.

Polymorphism of morphological traits was noted both between populations from different sampling sites, and between plants within nominal populations, designated by us according

Table 1. Morphological characteristics of wild-growing vines, Utrish Nature Reserve

Sample	Tip shape of young shoot	Prostrate hairs on tip of young shoot	Anthocyanin coloration of tip of young shoot	Shape of leaf blade	Depth of upper lateral sinuses of leaf	Degree of openness of leaf petiole sinus	Leaf teeth shape	Prostrate hairs between main veins on the lower side of leaf blade	Erect hairs on main veins on the lower side of leaf blade	Flower type	Colour of berry skin (without bloom)	Anthocyanin coloration of flesh berry
A1	3	3	5	3	9	2	3	5	3	1	–	–
A2	2	5	7	2	9	3	3	3	5	1	–	–
A3	3	5	3	3	9	2	3	5	3	–	–	–
Sh1	2	7	3	3	7	5	5	5	3	–	–	–
Sh2	3	5	3	4	9	5	5	5	5	–	–	–
Sh3	2	7	3	4	9	5	5	5	3	4	6	5
Sh4	3	5	5	2	5	3	3	7	5	4	6	3
Sh5	2	7	3	3	7	3	3	7	3	1	–	–
L1	3	5	3	3	5	4	4	7	1	–	–	–
L2	4	5	3	4	5	2	5	3	3	–	–	–
L3	2	7	5	2	7	2	4	7	5	–	–	–
L4	3	5	5	3	9	1	3	5	5	–	–	–
L5	3	5	3	1	9	2	4	5	3	–	–	–
L6	5	3	3	3	3	1	4	3	1	–	–	–
V1	3	5	3	3	5	4	4	7	1	–	–	–
V2	4	5	3	4	5	2	5	3	3	–	–	–
V3	2	7	3	2	7	2	4	7	5	–	–	–
V4	3	5	3	3	9	1	3	5	5	–	–	–
V5	3	5	3	1	9	2	4	5	3	–	–	–
V6	5	3	3	3	3	1	4	3	1	–	5	3
V7	2	5	3	3	5	2	5	3	3	–	5	3
V8	2	5	3	2	3	2	3	3	3	–	5	3
V9	4	7	3	3	5	4	4	3	1	–	5	3
V10	3	3	3	3	5	2	3	5	1	–	5	3

Note. Tip shape of young shoot: 2 – slightly open; 3 – half open; 4 – wide open; 5 – fully open. Prostrate hairs on tip of young shoot: 3 – sparse; 5 – medium; 7 – dense. Anthocyanin coloration of tip of young shoot: 3 – weak; 5 – medium; 7 – strong. Shape of leaf blade: 1 – cordate; 2 – wedge-shaped; 3 – pentagonal; 4 – circular. Depth of upper lateral sinuses of leaf: 3 – shallow; 5 – medium; 7 – deep; 9 – very deep. Degree of openness of leaf petiole sinus: 1 – very wide open; 2 – wide open; 3 – open, 4 – slightly open; 5 – closed. Leaf teeth shape: 3 – both sides convex; 4 – one side concave, one side convex; 5 – mixture of both sides straight and both sides convex. Prostrate hairs between main veins on the lower side of leaf blade: 3 – sparse; 5 – medium; 7 – dense. Erect hairs on main veins on the lower side of leaf blade: 1 – absent or very sparse; 3 – sparse; 5 – medium. Flower type: 1 – male; 4 – female. Colour of berry skin (without bloom): 5 – dark red-violet; 6 – blue black. Anthocyanin coloration of flesh berry: 3 – weak; 5 – medium. Dash – no data available.

Table 2. Ecological and geographical characteristics of the habitats of the analyzed wild-growing vines, Utrish Nature Reserve

Grape samples		Relief	Slope steepness	Slope exposition	Vegetation type	Soil type
No.	Code					
Atmacheva Shchel						
1	A1	Low-mountain	3	SW	Oak and ash forest	Brown carbonate rocky
2	A2					
3	A3					
Shirokaya Shchel						
4	Sh1	Low-mountain	5	SE	Oak and ash forest	Brown carbonate rocky
5	Sh2					
6	Sh3					
7	Sh4					
8	Sh5					
Lobanova Shchel						
9	L1	Low-mountain	10	SE	Linden-rocky-oak forest	Brown carbonate rocky
10	L2					
11	L3					
12	L4					
13	L5					
14	L6					
Vodopadnaya Shchel						
15	V1	Low-mountain	20	NE	Fluffy-oak-ash forest	Brown leached
16	V2					
17	V3					
18	V4					
19	V5					
20	V6		35	SW	Pistachio-juniper woodlands	Brown carbonate
21	V7					
22	V8					
23	V9					
24	V10					

to their places of growth on the territory of the reserve (plants were found in the locations of Atmacheva Shchel (A1–A3), Shirokaya Shchel (Sh1–Sh5), Lobanova Shchel (L1–L6), Vodopadnaya Shchel (V1–V10) (Gorbunov, Lukyanov, 2020; Gorbunov et al., 2020). The ecological and geographical characteristics of the places where grape plants were found are presented in Table 2.

DNA analysis of the found grape samples revealed different levels of polymorphism in the studied 15 loci – the number of identified alleles in the sample varied from 5 (VVMD25, VVMD27 and GF09-46) to 18 (UDV305) and amounted to an average of 8 alleles per locus (Table 3). The mean observed heterozygosity (Ho = 0.664) was lower than expected (He = 0.712). DNA marker analysis using VVIb23 showed that all wild vines were dioecious. Among the nine microsatellite loci, the data on polymorphism of which are used for DNA fingerprinting of grape genotypes, the most polymorphic was VVS2 (10 alleles were determined), the least – VVMD25, VVMD27 (5 types of alleles each) (see Table 3).

A similar situation was noted in the study of the diversity of wild vines in Armenia – the studied sample (77 samples) was also the most polymorphic in the locus of VVS2 – 13 types of alleles, and the least in VVMD25 and VVMD27 (5 and 8 alleles were identified, respectively) (Margaryan et al., 2019). At the same time, in the study of Gorislavec S.M. and co-authors (2017), in which Crimean wild-growing vines were studied, the smallest polymorphism (5 types of alleles) was revealed at the VVS2 locus (Gorislavec et al., 2017).

When distributing genotypes in the space of the main coordinates, a group of samples from territory of the Vodopadnaya Shchel can be distinguished (Fig. 1). In general, all samples from this place are localized separately from others in the space of the main coordinates, while samples V1, V2, V3, V4, V9, V10 form a separate subgroup. Samples V6 and

Table 3. Characteristics of microsatellite loci in the studied sample of 24 wild-growing vines

Locus	Na	Ne	Ho	He	Locus	Na	Ne	Ho	He
VVS2	10.000	7.067	0.917	0.859	VrZAG79	6.000	1.717	0.500	0.418
VVMD5	7.000	2.730	0.583	0.634	UDV305	18.000	8.662	0.542	0.885
VVMD7	8.000	6.508	0.958	0.846	UDV737	7.000	3.740	0.750	0.733
VVMD25	5.000	4.482	0.625	0.777	GF09-46	5.000	3.200	0.708	0.688
VVMD27	5.000	1.775	0.500	0.437	ScORGF15-02	8.000	3.080	0.292	0.675
VVMD28	9.000	3.008	0.583	0.668	CenGen6	10.000	6.436	0.792	0.845
VVMD32	9.000	6.160	0.542	0.838	VVIb23	6.000	2.654	0.750	0.623
VrZAG62	7.000	4.000	0.917	0.750	Mean	8.000	4.348	0.664	0.712

Note. Na – number of different alleles, Ne – number of effective alleles, Ho – observed heterozygosity, He – expected heterozygosity.

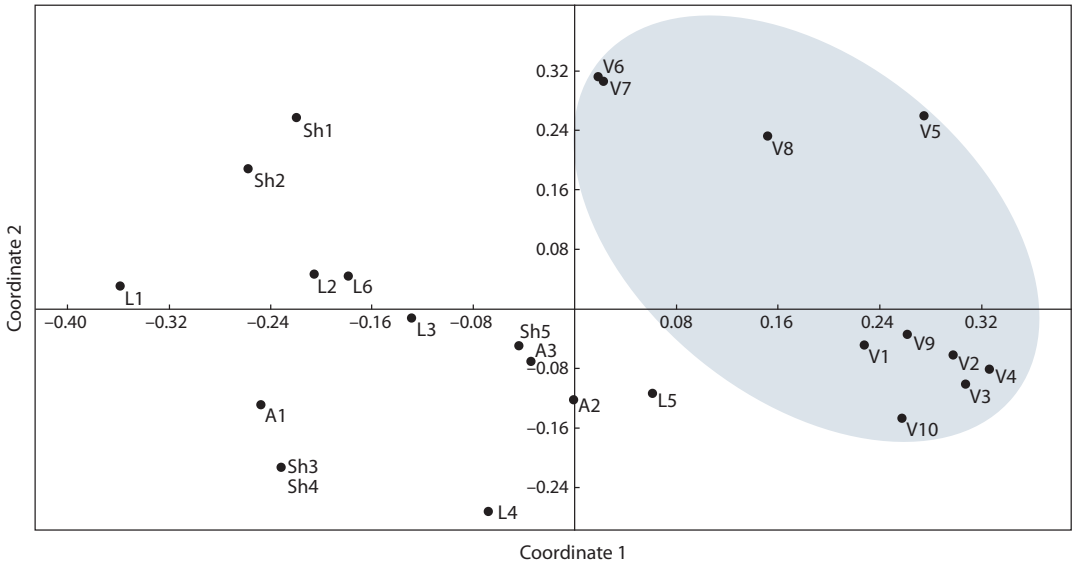


Fig. 1. Distribution of the studied grape genotypes in the space of the main coordinates.

V7 have very similar genotypes. Among other found samples, the complete coincidence of genotypes at the studied loci was determined in wild vines Sh3 and Sh4.

Analysis of samples using DNA markers linked to the resistance genes to downy mildew *Rpv3*, *Rp10* and powdery mildew *Ren3*, *Ren9* did not reveal resistance loci in the genotypes of wild vines. The inclusion of these DNA markers in the study was carried out in order to study the polymorphism of wild grapes at the analyzed loci, as well as a tool for the possible identification of wild-growing vines of hybrid origin. The resistance determined by *Ren3*, *Ren9* and *Rpv3* is inherited from North American grape species (*V. riparia*, *V. rupestris*, *V. labruska*, *V. lincecumii*), *Rpv10* – from *V. amurensis*.

DNA fingerprints of the studied wild-growing vines by 9 SSR loci standard for the identification of grape genotypes were checked for a coincidence in the catalog of grape varieties' DNA profiles of the international database *V*IVC (*V*IVC, 2022). No coincidences were found. To analyze the genetic similarity, DNA profiles of varieties belonging to geographically close regions of viticulture, where there are local varieties, some of which may have originated from wild grapevines growing earlier in these territories, were included into the research. For comparison, the identified DNA profiles of *V. sylvestris* genotypes from different geographical locations (Israel, Tunisia, France, Armenia, Turkey) presented in *V*IVC were also used. The study also included a group of rootstock varieties of complex interspecific origin with the largest share of the genetic contribution of North American grape species in order to exclude the presence of rootstock varieties among the found wild vines, which are characterized by high adaptability to various abiotic and biotic stress factors. The number of clusters equal to 7 (K=7) was used for Bayesian analysis. The results of the analysis are shown in Figure 2.

The data obtained during the analysis allowed us to identify certain patterns of clustering of samples, relative to their origin. The fourth and fifth clusters mainly include samples selected from wild grapevines of the Utrish Nature Reserve. The exceptions are three genotypes of the grape subspecies *Vitis*

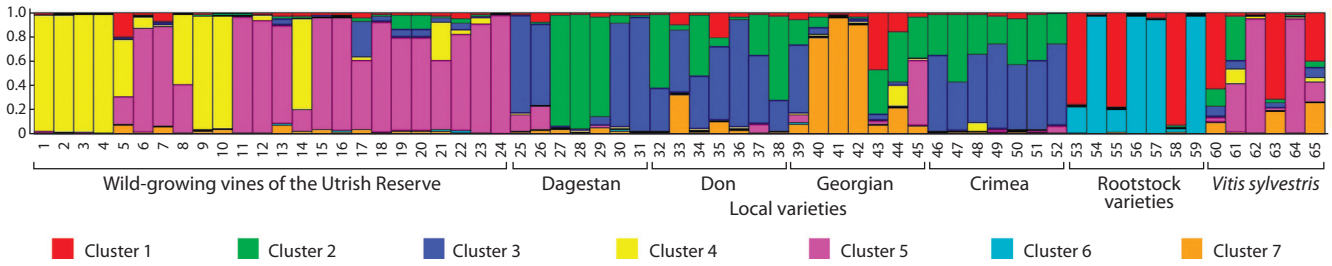


Fig. 2. Clustering of 65 grape genotypes by origin.

The vertical axis denotes the probability of assigning each genotype to putative clusters, indicated by different colors. Genotypes: 1 – V1, 2 – V2, 3 – V3, 4 – V4, 5 – V5, 6 – V6, 7 – V7, 8 – V8, 9 – V9, 10 – V10, 11 – L1, 12 – L2, 13 – L3, 14 – L4, 15 – L5, 16 – L6, 17 – Sh1, 18 – Sh2, 19 – Sh3, 20 – Sh4, 21 – Sh5, 22 – A1, 23 – A2, 24 – A3, 25 – Agadai, 26 – Rish baba, 27 – Tavlinskiy pozdniy, 28 – Hatal baar, 29 – Khop khalat, 30 – Sarakh, 31 – Khatmi, 32 – Varyushkin, 33 – Mushketnyi, 34 – Sibirkovyi, 35 – Efremovskiy, 36 – Shilokhvostyi, 37 – Tsimlyanskiy chernyi, 38 – Shampanchik bessergenevskiy, 39 – Tsitska, 40 – Aleksandrouli, 41 – Mtsvane Kakhuri, 42 – Rkatsiteli, 43 – Saperavi, 44 – Tsolikouri, 45 – Chkhaveri, 46 – Sary kokur, 47 – Kharko, 48 – Kefesiya, 49 – Sary pandas, 50 – Shabash, 51 – Dzhevat kara, 52 – Kokur belyi, 53 – Couderc 1616, 54 – Kober 5 BB, 55 – Millardet et Grasset 101-14, 56 – Teleki 8 B, 57 – Rupestris du lot, 58 – Fercal, 59 – Paulsen 1103, 60 – Gesher Hardof (Israel), 61 – Khedhayria (Tunisia), 62 – Lambrusque Abbadia H (France), 63 – Sveni (Armenia), 64 – Sylvestris Dirmstein 2 (unknown), 65 – Sylvestris Guemuelduer 104-64 (Turkey).

vinifera L. subsp. *sylvestris* (Gmelin) (*V. sylvestris*) – Khedhayria (Tunisia) (61), Lambrusque Abbadia H (France) (62) and Sylvestris Dirmstein 2 (64), and one local Georgian variety Chkhaveri (45) belonging to cluster 5 with a probability not exceeding 50 %. It is worth noting that the three studied wild grape vines (5, 8 and 21) were not unambiguously assigned to cluster 4 or 5. In turn, clusters 2, 3 and 7 include all the local varieties presented in the study: of Dagestan, Don, Georgian and Crimean origin. At the same time, cluster 7 is typical for three varieties of Georgian origin – Alexandrouli (40), Mtsvane Kakhuri (41) and Rkatsiteli (42); in other cases, the probability of its contribution to the genotypes is not significant. With the exception of the three varieties mentioned above, representatives of the local gene pool are distributed between clusters 2 and 3 with varying degrees of confidence. The sixth cluster was formed by a number of genotypes of rootstock varieties, the minor contribution of this cluster was also revealed only in varieties from this sample. Cluster 1 includes some rootstock varieties, part of the *V. sylvestris* genotypes and one Georgian variety Saperavi (43).

Conclusion

Based on the above, we can conclude that the samples of wild grapes of the Utrish Nature Reserve selected during the expedition are represented by two hypothetical populations (expressed as clusters 4 and 5). There are transitional forms between the two populations. And if the first nominal population (cluster 4) is localized on the Vodopadnaya Shchel territory, then representatives of the second population (cluster 5) are found at all expeditionary points of sampling of plant material. A genetic relationship was also established between the second nominal population and some genotypes of *V. sylvestris* and the Georgian variety Chkhaveri (45), with varying degrees of probability included in cluster 5.

Thus, it can be assumed that at least part of the genotypes found in the Utrish Nature Reserve are close to the genotypes of the subspecies *V. sylvestris* presented in the *V*IVC international database. Samples from the hypothetical first population (primarily localized – Vodopadnaya Shchel) are genetically different from other vines and are predominantly allocated to cluster 4. An insignificant contribution of cluster 4 was noted in the genotypes from the gene pool of Georgian and Crimean local varieties (44 –Tsolikouri, and 48 – Kefesia), and also in two accessions of *V. sylvestris* (61 – Khedhayria (Tunisia), 65 – Sylvestris Guemuelduer 104-64 (Turkey)). The genetic contribution of American rootstock varieties to the wild grape population of the Utrish Nature Reserve has not been identified.

References

Chernodubov A.I., Rudenok Y. Phytocenoses biodiversity in “Utrish” Reserve. *Lesotehnicheskiy Zhurnal = Forestry Engineering Journal*. 2015;5(1):120-127. DOI 10.12737/11269. (in Russian)

Cunha J., Ibáñez J., Teixeira-Santos M., Brazão J., Fevereiro P., Martínez-Zapater J.M., Eiras-Dias J.E. Genetic relationships among portuguese cultivated and wild *Vitis vinifera* L. germplasm. *Front. Plant Sci*. 2020;11:127. DOI 10.3389/fpls.2020.00127.

De Michele R., La Bella F., Gristina A.S., Fontana I., Pacifico D., Garfi G., Motisi A., Crucitti D., Abbate L., Carimi F. Phylogenetic relationship among wild and cultivated grapevine in Sicily: a hotspot in the middle of the Mediterranean basin. *Front. Plant Sci*. 2019;10:1506. DOI 10.3389/fpls.2019.01506.

Di Gaspero G., Copetti D., Coleman C., Castellarin S.D., Eibach R., Kozma P., Lacombe T., Gambetta G., Zvyagin A., Cindrić P., Kovács L., Morgante M., Testolin R. Selective sweep at the *Rpv3* locus during grapevine breeding for downy mildew resistance. *Theor. Appl. Genet*. 2012;124(2):227-286. DOI 10.1007/s00122-011-1703-8.

Doulati-Baneh H., Mohammadi S., Labra M., De Mattia F., Bruni I., Mezzasalma V., Abdollahi R. Genetic characterization of some wild grape populations (*Vitis vinifera* subsp. *sylvestris*) of Zagros mountains (Iran) to indentify a conservation strategy. *Plant Genet. Resour.* 2015;13(1):27-35. DOI 10.1017/S1479262114000598.

Ellstrand N.C., Heredia S.M., Leak-Garcia J.A., Heraty J.M., Burger J.C., Li Y., Nohzadeh-Malakshah S., Ridley C.E. Crops gone wild: evolution of weeds and invasive from domesticated ancestors. *Evol. Appl*. 2010;3(5-6):494-504. DOI 10.1111/j.1752-4571.2010.00140.x.

Gorbunov I.V., Lukyanov A.A. Kuban wild-growing grapes and their morphological features. *Vinogradarstvo i Vinodelie = Viticulture and Winemaking*. 2020;49:30-33. (in Russian)

Gorbunov I.V., Lukyanov A.A., Bykhalova O.N. Morphological peculiarities of the Kuban wild-growing forms of grapes. *Plodovodstvo i Vinogradarstvo Yuga Rossii = Fruit Growing and Viticulture of South Russia*. 2020;65(5):70-82. DOI 10.30679/2219-5335-2020-5-65-70-82. (in Russian)

Gorislavec S.M., Risovannaya V.I., Volkov Y.A., Kolosova A.A., Volodin V.A. Identification and evaluation of wild growing vines on the territory of Yalta mountain-forest nature reserve using molecular markers. *"Magarach". Vinogradarstvo i Vinodelie = "Magarach". Viticulture and Winemaking*. 2017;1:19-21. (in Russian)

Hammer Ø., Harper D.A.T., Ryan P.D. PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electron*. 2001;4(1):1-9.

Kupe M., Ercisli S., Jovanović-Cvetković T., Eyduran S.P., Ayed R.B. Molecular characterization of wild grapes from northeastern part of Turkey. *Genetika*. 2021;53(1):93-102. DOI 10.2298/GENSR2101093K.

Lukšić K., Zdunić G., Hančević K., Mihaljević M.Ž., Mucalo A., Maul E., Riaz S., Pejić I. Identification of powdery mildew resistance in wild grapevine (*Vitis vinifera* subsp. *sylvestris* Gmel Hegi) from Croatia and Bosnia and Herzegovina. *Sci. Rep*. 2022;12(1):2128. DOI 10.1038/s41598-022-06037-6.

Margaryan K., Maul E., Muradyan Z., Hovhannisyan A., Melyan G., Aroutiounian R. Evaluation of breeding potential of wild grape originating from Armenia. *BIO Web Conf*. 2019;15:01006. DOI 10.1051/bioconf/20191501006.

McGovern P.E. Ancient Wine: The Search for the Origins of Viniculture. Princeton: Princeton Univ. Press, 2004. DOI 10.2307/j.ctvfjd0bk.

Merdinoglu D., Butterlin G., Bevilacqua L., Chiquet V., Adam-Blondon A.F., Decroocq S. Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. *Mol. Breed*. 2005;15:349-366. DOI 10.1007/s11032-004-7651-0.

Peakall R., Smouse P.E. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*. 2012;28(19):2537-2539. DOI 10.1093/bioinformatics/bts460.

Pritchard J.K., Stephens M., Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155(2):945-959. DOI 10.1093/genetics/155.2.945.

Riaz S., Boursiquot J.M., Dangl G.S., Lacombe T., Laucou V., Tenscher A.C., Walker M.A. Identification of mildew resistance in wild and cultivated Central Asian grape germplasm. *BMC Plant Biol*. 2013;13(1):149. DOI 10.1186/1471-2229-13-149.

Rogers S.O., Bendich A.J. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol*. 1985;5:69-76. DOI 10.1007/BF00020088.

Schwander F., Eibach R., Fechter I., Hausmann L., Zyprian E., Töpfer R. *Rpv10*: a new locus from the Asian *Vitis* gene pool for pyramiding downy mildew resistance loci in grapevine. *Theor. Appl. Genet*. 2012;124(1):163-176. DOI 10.1007/s00122-011-1695-4.

This P. Microsatellite markers analysis. In: Minutes of the First Grape Gen06 Work-shop, March 22nd and 23rd. INRA, Versailles (France). Paris: INRA, 2007;3-42.

This P., Jung A., Boccacci P., Borrego J., Botta R., Costantini L., Crespan M., Dangl G.S., Eisenheld C., Ferreira-Monteiro F., Grando S., Ibáñez J., Lacombe T., Laucou V., Magalhães R., Meredith C.P., Milani N., Peterlunger E., Regner F., Zulini L., Maul E. Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theor. Appl. Genet*. 2004;109(7):1448-1458. DOI 10.1007/s00122-004-1760-3.

This P., Lacombe T., Thomas M.R. Historical origins and genetic diversity of wine grapes. *Trends Genet.* 2006;22(9):511-519. DOI 10.1016/j.tig.2006.07.008.

van Heerden C.J., Burger P., Vermeulen A., Prins R. Detection of downy and powdery mildew resistance QTL in a 'Regent' × 'RedGlobe' population. *Euphytica*. 2014;200:281-295. DOI 10.1007/s10681-014-1167-4.

VIVC. *Vitis* International Variety Catalogue. Microsatellites by profile. Last modified: 2022-08-10. Julius Kühn-Institut, 2022. Available at: https://www.vivc.de/index.php?r=eva-analysis-mikrosatelliten-vivc%2Findex

Zdunić G., Lukšić K., Nagy Z.A., Mucalo A., Hančević K., Radić T., Butorac L., Jahnke G.G., Kiss E., Ledesma-Krist G., Regvar M., Likar M., Piltaver A., Žulj Mihaljević M., Maletić E., Pejić I., Werling M., Maul E. Genetic structure and relationships among wild and cultivated grapevines from Central Europe and part of the Western Balkan Peninsula. *Genes.* 2020;11(9):962. DOI 10.3390/genes11090962.

Zendler D., Schneide P., Töpfe R., Zyprian E. Fine mapping of *Ren3* reveals two loci mediating hypersensitive response against *Erysiphe necator* in grapevine. *Euphytica*. 2017;213(68):1029. DOI 10.1007/s10681-017-1857-9.

ORCID ID

E.T. Ilnitskaya orcid.org/0000-0002-2446-0971
M.V. Makarkina orcid.org/0000-0002-3397-0666
I.V. Gorbunov orcid.org/0000-0002-4702-9148
I.V. Stepanov orcid.org/0000-0002-6251-300X
T.D. Kozina orcid.org/0000-0003-2908-6461
E.A. Kozhevnikov orcid.org/0000-0003-1305-3614
V.K. Kotlyar orcid.org/0000-0002-4865-0323

Acknowledgements. This study was supported by the Russian Foundation for Basic Research and the Administration of the Krasnodar Territory (grant No. 19-416-230025 r_a).

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

Received August 15, 2022. Revised October 11, 2022. Accepted October 31, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):323-332
DOI 10.18699/VJGB-23-39

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

Creation and study of emmer (*Triticum dicoccum*) × triticale hybrids

O.G. Silkova[1] ✉, Y.N. Ivanova[1], P.I. Stepochkin[2]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Siberian Research Institute of Plant Production and Breeding – Branch of the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ silkova@bionet.nsc.ru

Abstract. Triticale (× *Triticosecale* Wittmack) is of great interest as an insurance crop that can ensure the stability of the gross harvest of feed and food grains at a lower cost. In Western Siberia, only winter triticale varieties are cultivated, however, spring triticales are important for cultivation in regions not suitable for winter crops. To create spring varieties with high yields and good grain quality, it is necessary to study and enrich the gene pool, identify donors of economically valuable traits. One of the possible ways to solve this problem can be through the production of secondary hexaploid triticales with the involvement of the tetraploid wild-growing species of emmer wheat *Triticum dicoccum* (Schrank) Schuebl. The aim of this work was to create and study hybrids of emmer *T. dicoccum* (Schrank) Schuebl. with hexaploid triticale using genomic *in situ* hybridization for staining of meiotic chromosomes and analysis of plant productivity elements in F_4–F_8. DT4, DT5, DT6 plants and the prebreeding F_6 forms obtained from them – DT 4/168, DT 5/176 and DT 6/186 – were selected according to the characteristics of the productivity and the nature of the grain in the F_4 hybrid population. The offspring of hybrids DT4 and DT5 and prebreeding forms DT 4/168 and DT 5/176 had an increased grain nature (over 750 g/l), but low productivity. The hybrid DT6 and the breeding form DT 6/186 obtained from it had high grain productivity (785 ± 41 and 822 ± 74 g/m^2, respectively), but, like the paternal form of triticale UK 30/33, had a reduced nature of the grain. In F_8 DT 6/186 plants, 7 homologous pairs of rye chromosomes and from 27 to 30 wheat chromosomes were found in meiosis, which indicates the presence of a complete rye genome and two wheat AABB genomes. Rye chromosomes showed stable formation of bivalents in contrast to wheat chromosomes, which caused the presence of aneuploids in plant populations. Thus, hexaploid forms DT 4/168 and DT 5/176 with well-made smooth grain and high grain size were obtained, which can be used as a source of this trait for selection of food-grade triticale. DT 6/186 is a promising form for further breeding in order to obtain high-yielding forms of triticale.

Key words: triticale; *Triticum dicoccum*; wide hybrids; genomic *in situ* hybridization; productivity traits; meiosis; prebreeding forms.

For citation: Silkova O.G., Ivanova Y.N., Stepochkin P.I. Creation and study of emmer (*Triticum dicoccum*) × triticale hybrids. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):323-332. DOI 10.18699/VJGB-23-39

Создание и изучение гибридов полба (*Triticum dicoccum*) × тритикале

О.Г. Силкова[1] ✉, Ю.Н. Иванова[1], П.И. Стёпочкин[2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Сибирский научно-исследовательский институт растениеводства и селекции – филиал Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ silkova@bionet.nsc.ru

Аннотация. Тритикале (× *Triticosecale* Wittmack) представляет большой интерес как страховая культура, способная обеспечить стабильность валового сбора фуражного и продовольственного зерна с более низкими затратами. В Западной Сибири возделываются сорта только озимых тритикале, однако яровые тритикале являются значимыми для выращивания в регионах, не пригодных для озимых культур. Для создания яровых сортов с высокой урожайностью и хорошим качеством зерна необходимо изучение и обогащение генофонда, выделение доноров хозяйственно ценных признаков. Одним из возможных путей решения этой задачи может быть получение вторичных гексаплоидных тритикале с привлечением тетраплоидного дикорастущего вида пшеницы полба *Triticum dicoccum* (Schrank) Schuebl. Целью данной работы было создание и изучение гибридов полбы *T. dicoccum* (Schrank) Schuebl. с гексаплоидной тритикале с использованием геномной *in situ* гибридизации при окрашивании мейотических хромосом и анализ элементов продуктивности растений в F_4–F_8. По признакам продуктивности и натуры зерна в гибридной популяции F_4 были отобраны растения ДТ4, ДТ5, ДТ6

и выявленные в их потомстве пребридинговые формы F_6 – ДТ 4/168, ДТ 5/176 и ДТ 6/186. Потомства гибридов ДТ4 и ДТ5 и форм ДТ 4/168 и ДТ 5/176 имели повышенную натуру зерна (свыше 750 г/л), но невысокую продуктивность. Гибрид ДТ6 и полученная от него форма ДТ 6/186 отличались высокими показателями продуктивности зерна (785 ± 41 и 822 ± 74 г/м^2 соответственно), но, как и отцовская форма тритикале УК 30/33, имели пониженную натуру зерна. У растений потомства F_8 ДТ 6/186 в мейозе обнаружено 7 гомологичных пар хромосом ржи и от 27 до 30 хромосом пшеницы, что свидетельствует о наличии полного генома ржи и двух геномов пшеницы AABB. Хромосомы ржи демонстрировали стабильное формирование бивалентов в отличие от хромосом пшеницы, что вызвало анеуплоидию в популяциях растений. Таким образом, получены гексаплоидные формы ДТ 4/168 и ДТ 5/176 с хорошо выполненным гладким зерном и высокой натурой зерна, которые можно использовать в качестве источника этого признака для селекции тритикале пищевого направления. Форма ДТ 6/186 перспективна для дальнейшего селекционного процесса с целью получения высокоурожайных форм тритикале.
Ключевые слова: тритикале; *Triticum dicoccum*; отдаленные гибриды; геномная *in situ* гибридизация; признаки продуктивности; мейоз; пребридинговые формы.

Introduction

Triticale (×*Triticosecale* Wittmack) is a wheat and rye hybrid of a relatively short evolutionary history as an allopolyploid species. The first naturally created fertile wheat×rye hybrids were discovered in the late 1920s at the South-Eastern agricultural experimental station in Saratov (Meister, 1921). The plants had intermediate traits and were described as a new botanical species *Triticum Secalotriticum saratoviense* Meister by G.K. Meister (Levitsky, 1978). Meister immediately predicted the practical value of these intergeneric crossings. The first man-made wheat×rye hybrids (*Triticosecale* Wittmack) were obtained in 1888 by the German plant breeder W. Rimpau (Müntzing, 1974), who described 12 plants descending from a wheat×rye hybrid generally recognized as the first triticales (×*Triticosecale* Wittmack). The cytological analysis of the first triticales developed in Russia and Germany showed the somatic chromosome number of 56 (8*x*) (Müntzing, 1961; Levitsky, Benetzkaja, 1978), which demonstrated the combination of four genomes as follows: AABBDDRR, AABBDD, soft wheat, and RR rye.

Octoploid triticales were of great interest for plant breeders due to high seed set per spike, increased plant pathogen resistance, and resistance to environmental stresses. However, the primary triticale lines suffered from meiotic errors (Müntzing, 1961; Shkutina, Khvostova, 1971; Lukaszewski, Gustafson, 1987) and high frequency of aneuploidy in earlier generations (Krolow, 1962; Stepochkin, Vladimirov, 1978; Silkova et al., 2021), which caused reduced fertility. Grain shriveling and lateness of maturity in octoploid forms were also considered as limitations for their introduction as cultivars. As a result, there have been global efforts to develop the technology for obtaining more promising lines with various valuable breeding traits, which eventually produced hexaploid triticales (Shulyndin, 1975).

Primary hexaploid triticales of the AABBRR genome type were obtained as crossings of tetraploid wheats (*T. turgidum*, *T. durum*) and *S. cereale* rye (Müntzing, 1974). However, undesirable traits were not completely eliminated. Agronomically valuable traits were improved by enriching the triticale gene pool with crossing experiments involving octoploid lines and commercial soft wheat varieties, as well as hexaploid triticales with full rye genome and two wheat genomes identified in the progeny of octoploid triticales as a result of D-genome chromosome elimination (Stepochkin, 1978; Dou et al., 2006; Zhou et al., 2012; Hao et al., 2013; Li et al., 2015; Evtushenko et al., 2019). The crossing of two primary octoploid triticales produced hexaploid offspring, and breeding programs were focused on hybridization of these octoploids with the hexaploids identified in the progeny (Pisarev, 1964; Jenkins, 1969; Ammar et al., 2004; Oettler, 2005). Hexaploid triticales were more stable in terms of productivity (Müntzing, 1974; Oettler, 2005). As a result, various recombinant forms of secondary triticales were derived and karyotyped (Merker, 1975; Gustafson, Bennett, 1976; Lukaszewski, Gustafson, 1983; Badaev et al., 1985; Cheng, Murata, 2002; Mergoum et al., 2009; Shishkina et al., 2009; Fu et al., 2014), and some of them having combined wheat-rye genomes showed commercial value (Merker, 1975; Oettler, 2005; Zhou et al., 2012).

Agronomic traits of triticales were further improved by intergeneric and interspecific crossings. *Aegilops crassa* (2*n* = 4*x* = 28; DDMcrMcr), *Ae. juvenalis* (2*n* = 6*x* = 42; DDCuCu MjMj), *Ae. squarrosa* (syn. *Ae. tauschii*; 2*n* = 2*x* = 14; DD) and *Ae. triaristata* (2*n* = 6*x* = 42; CuCuMtMtMt2Mt2) (Gruszecka et., 1996), *Agropyron intermedium* ssp. *trichophorum* (2*n* = 42) (Gupta, Fedak, 1986a), *Hordeum parodii* Covas (Gupta, Fedak, 1986b), *H. vulgare* L. (Balyan, Fedak, 1989) and *T. monococcum* L. (Neumann, Kison, 1992) plants were used in hybridization experiments with hexaploid triticales. Intergeneric polyploid triticales were also bred through hybridization with intermediate forms sharing at least one set of chromosomes (genome) with the triticale genome. These hybridization efforts produced plants with resistance genes against diseases. Lines isolated in the progeny of triticale hybrids (AABBRR) and amphidiploids (wheat × *Psathyrostachys huashanica*, 2*n* = 8*x* = 56, AABBDDNsNs) were resistant to yellow rust (Kang et al., 2016). Hybridization of a hexaploid triticale and an amphiploid intermediate form (*Ae. variabilis* × *S. cereale*, 2*n* = 6*x* = 42; UUSvSvRR) produced addition and substitution lines with the 3Sv *Ae. variabilis* chromosome carrying the powdery mildew resistance gene *Pm13* (Kwiatek et al., 2016). Addition and substitution lines for chromosome 2D with the leaf-rust resistance gene *Lr39* and semidwarf gene *Rht8*, as well as chromosome 3D (or 3D/3B) with the *Lr32*

gene were isolated in the progeny of triticale × amphiploid hybrids (*Ae. tauschii* × *S. cereale*, $2n = 4x = 28$; DDRR) (Kwiatek et al., 2015).

Thanks to breeding achievements, triticale has become a new economically significant cereal species characterized by high grain and vegetative mass productivity, that could be used as forage and green feed (McGoverin et al., 2011; Ayalew et al., 2018). In the last three decades, its products have become increasingly significant, which is demonstrated by increasing crop acreage across the world, from 1,453,269 ha in 1994 to 4,157,018 ha in 2016. Triticale grains are used to produce bioethanol and food wrap, as well as various food products (bread, biscuits, pastas, flatbreads, and malt) (Zhu, 2018), and bran is used as the source of prebiotics and antioxidants for yoghurts. Food-grade triticale grains are comparable to wheat in macro- and micronutrient contents (Zhu, 2018). Protein content in triticale grains is 1–1.5 % higher than in wheat and 3–4 % higher than in rye, gluten content matches that in wheat or is 2–4 % higher (Meleshkina et al., 2015). However, triticale underperforms in test weight. This parameter is closely related to the plumpness and hardness of grains, as well as to their size and shape. Average test weight for wheat is 700–810 g/l. At test weights below 740 g/l, flour yield tends to decrease rapidly as test weight goes down. Most spring triticales have shriveling grains and low flour yields, which limits their use in bread making (Rakha et al., 2011).

Development of domestic high-productivity triticale varieties with high grain quality requires further study and enrichment of the gene pool, as well as identification of donors for economically valuable traits. One way to solve this problem is to obtain secondary hexaploid triticales using emmer wheat *T. dicoccum* (Schrank) Schuebl., a tetraploid wild-growing wheat with long, large, and plump grains.

The goal of the present study was to breed new forms of hexaploid triticales (AABBRR genome type) with improved test weights by crossing emmer (*T. dicoccum* Schrank, AABB genome) with triticale and study their productivity and meiotic stability with chromosome staining using genomic *in situ* hybridization.

Material and methods

Plant material. New forms of hexaploid triticale were obtained by hybridization of emmer (*T. dicoccum* (Schrank) Schuebl.) and triticale (×*Triticosecale* Wittmack). Intergeneric F_1 hybrid of emmer lines (L133 × PKK) × k-25516 (AABB genome) was used as maternal plants. Awned semi-hulless emmer (L133 × PKK) created by VIR researchers is characterized by brittle spikes and low productivity, and awnless emmer k-25516 was obtained at Siberian Research Institute of Plant Production and Breeding (SRI PPB), ICG SB RAS from the population of awned emmer wheat from the VIR collection. The paternal plants were represented by a selection form of hexaploid triticale UK 30/33 selected from the population of cytogenetically unstable octoploid triticale UK30 (AABBDDRR genome) developed at the SRI PPB by crossing the Ulyanovka soft wheat variety with the Korotkostebelnaya 69 short stem rye with subsequent doubling of chromosome number induced by colchicine water solution.

The progeny of three F_4 hybrids DT4, DT5, and DT6, and three selection forms DT 4/168, DT 5/176, and DT 6/186 isolated in F_5 from hybrid populations DT4, DT5, and DT6 respectively was selected for the study. In 2020, their F_6 progeny was seeded for research purposes along with F_4 hybrids at the nursery for distant wheat hybrids in the field of SRI PPB.

Fluorescence *in situ* hybridization of meiotic chromosomes. To evaluate the meiotic stability in the selection forms, two most productive plants DT 6/186/156 and DT 6/186/165 were selected in the F_7 progeny of plant DT 6/186, and their F_8 seeds were seeded in the hydroponic greenhouse of the ICG SB RAS in spring 2021. The chromosomal behavior of the progeny was studied using routine acetocarmine staining protocol and FISH staining (fluorescence *in situ* hybridization) following the technique described earlier (Ivanova et al., 2019).

The meiocytes were analyzed at the diakinesis stage, metaphase I (MI), anaphase I (AI), and telophase II (TII). The probes used in the analysis were as follows: *Aegilops tauschii* pAet6-09 specific for centromere repeats in rice, wheat, rye, and barley chromosomes (Zhang et al., 2004); pAWRc specific for centromere repeats in rye chromosomes (Francki, 2001), and rye genomic DNA. DNA repeat samples of pAet6-09 and pAWRc were the courtesy of Dr. A. Lukaszewcki (University of California, Riverside, the United States). Centromere-specific probes were labeled with biotin 16-dUTP or digoxigenin 11-dUTP by means of polymerase chain reaction (PCR). The total rye DNA was labeled by Nick translation with digoxigenin 11-dUTP. The probes were combined in different proportions and mixed with blocking wheat DNA. The preparations were mounted in Vectashield antifade solution (Vector Laboratories) slowing down fluorescence fading and including 1 μg/ml DAPI (4′,6-diamidino-2-phenylindol, Sigma-Aldrich, the United States) for chromatin staining. All preparations were analyzed using an Axio Imager M1 microscope (Karl Zeiss, Germany), the images were recorded using a ProgRes MF camera (Meta Systems, Jenoptic) at the Center of Microscopic Analysis of Biological Objects, SB RAS and processed using Adobe Photoshop CS2 software.

Analysis of economically valuable traits. Structural analysis of the plants was performed at the facility equipped for metric measurements, threshing, and seed weighting. As a result, the following data on productivity elements were obtained: spike length; spike density; grain weight per spike; thousand kernel weight; grain number per spikelet; test weight measured using a microchondrometer (Stepochkina, Stepochkin, 2015); and grain productivity per 1 m^2. Statistical processing of the results was carried out following the standard technique (Dospekhov, 1985). Significance of differences between mean values of two samples was estimated using Student's *t*-test.

Results

The parental-line plants had different spike and grain morphology. The emmer plants (L133 × PKK) had short, awned, brittle spikes and smooth, long grains (Fig. 1, *a*). The k-25516 emmer sample was an awnless plant with thin, long grains (see Fig. 1, *b*). Triticale UK 30/33 had a dense, awned spike and

Fig. 1. Spikes and seeds: *a*, emmer wheat (L133 × PKK); *b*, emmer wheat k-25516; *c*, triticale UK 30/33; F_4 hybrids: *d*, DT4; *e*, DT6; *f*, DT5.

Table 1. Results of structural analysis of triticale plants and emmer × triticale hybrids

Name of triticale line or emmer × triticale hybrid	Spike length, cm	Spike density	Grain weight per spike, g	Thousand kernel weight, g	Grain number per spikelet	Test weight, g/l	Productivity, g/m^2
UK 30/33	8.3 ± 0.3	2.89 ± 0.08	2.6 ± 0.2	47 ± 2	2.48 ± 0.10	706 ± 8	584 ± 33
DT4	11.0 ± 0.4**	2.10 ± 0.06	2.1 ± 0.2	50 ± 2	1.78 ± 0.10**	806 ± 14***	435 ± 34*
DT6	10.2 ± 0.2**	3.14 ± 0.07*	3.7 ± 0.2**	48 ± 1	2.43 ± 0.09	749 ± 16**	785 ± 41**
DT5	8.3 ± 0.1	2.64 ± 0.12*	1.3 ± 0.2*	39 ± 3*	1.56 ± 0.11**	789 ± 7***	280 ± 44**
DT 4/168	10.8 ± 0.8**	2.31 ± 0.08**	2.5 ± 0.4	47 ± 2	2.15 ± 0.26	760 ± 20*	531 ± 92
DT 6/186	10.8 ± 0.5**	2.38 ± 0.14**	3.9 ± 0.3**	48 ± 5	3.20 ± 0.18**	706 ± 25	822 ± 74**
DT 5/176	9.6 ± 0.7*	2.28 ± 0.11**	2.0 ± 0.3	47 ± 3	1.95 ± 0.13**	772 ± 48*	423 ± 67*

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ – significant differences between the hybrid and 6*x* triticale UK 30/33.

grains similar to those of soft wheat in shape but shriveling (see Fig. 1, *c*).

Plants of three emmer×triticale F_4 hybrids had different spike morphology (Table 1, see Fig. 1, *d–f*). The DT6 hybrid had a dense, awned spike. The plants of the remaining two hybrids had looser spikes. The spikes were sterile at the ends and therefore often susceptible to ergot.

All hybrids had a hair neck similar to the paternal triticale UK 30/33. This means that these genotypes have a gene responsible for manifestation of this trait localized in the long arm of chromosome 5R.

Triticale UK 30/33 had a short spike, low test weight, and medium grain productivity (see Table 1). Selection forms and hybrids, except for DT5, had a longer spike than the paternal form UK 30/33. Hybrid DT5 had low productivity due to flower sterility in the upper part of the spike and low thousand kernel weight. Hybrid DT4 had a high thousand kernel weight of 50 ± 2 g, and the highest test weight of 806 ± 14 g/l.

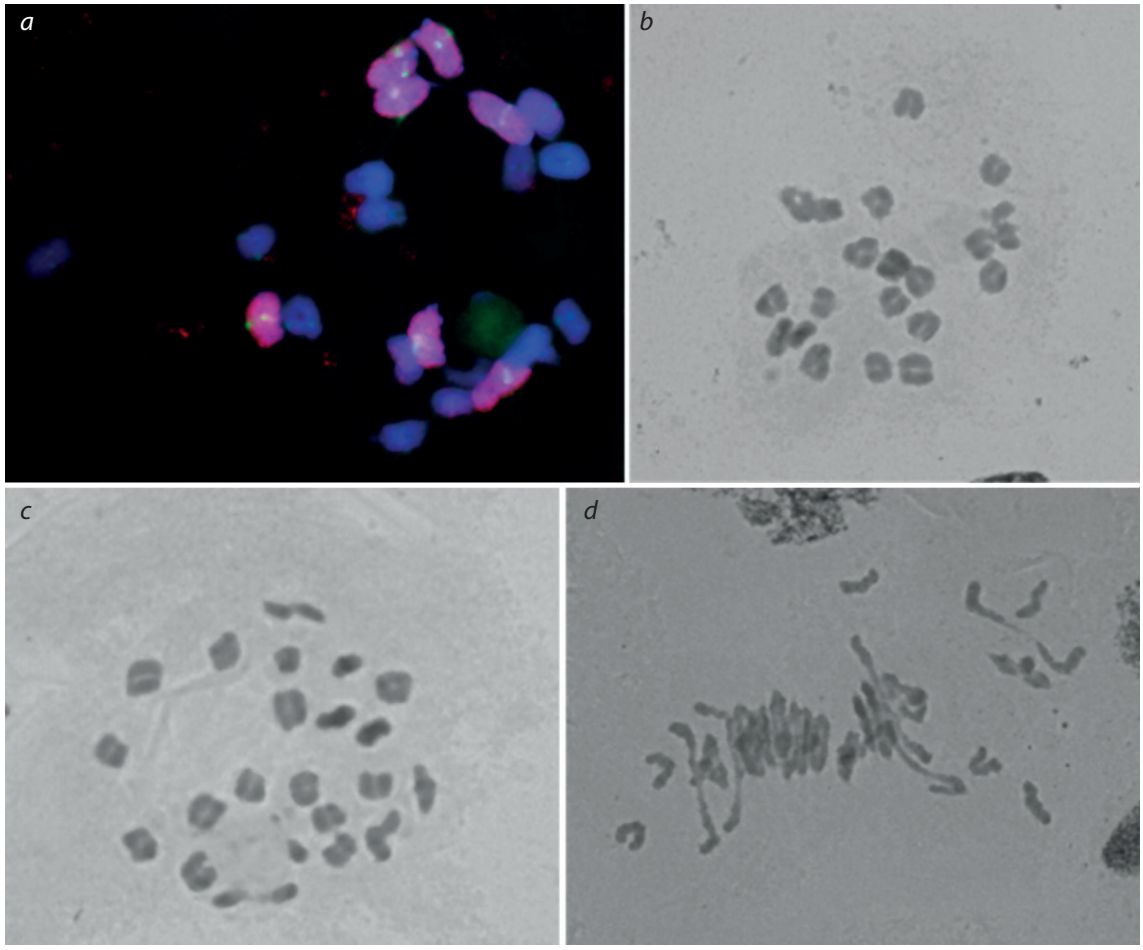


Fig. 2. Diakinesis (*a*–*c*) and metaphase I (*d*) meiotic stages in the progeny of DT 6/186/156 plant (*a*, *b*) and DT 6/186/165 plant (*c*, *d*).
a, *b*, 21 bivalents, *a*, 7 rye bivalents; *c*, 21 bivalents and one univalent; *d*, univalents in metaphase I. *a*, genomic *in situ* hybridization, rye chromosomes stained red; *b*–*d*, acetocarmine staining.

However, its grain productivity, although higher than that of DT5, was still not too high due to low seed set of spikes and spikelets. Flowers at the top of the spike were often sterile as well. Hybrid DT6 was characterized by dense spikes and high grain productivity (785 ± 41 g/m^2), as well as seed set of spikes and spikelets. The spikes were fertile along their full length.

Selection forms DT 4/168, DT 5/176, and DT 6/186 had different grain weights per spike, grain numbers per spike, test weights, and productivities. DT 4/168 had denser spikes and slightly higher seed sets of spikes and spikelets compared to the DT4 hybrid it was obtained from. Breeding sample DT 5/176 was more productive than the initial DT5 hybrid, but it had the lowest productivity among the three breeding samples studied. Small spike size and sterility of 3–7 spikelets in the upper part of the spike resulted in low productivity of hybrids DT5 and DT 5/176. The DT 6/186 line had the highest seed set and productivity among the studied lines. Despite the lowest test weight values similar to triticale UK 30/33 (706 ± 25 g/l), grain productivity per plot reached 822 ± 74 g/m^2. DT 6/186 had a smoother overall morphology than the initial hybrid DT6, but higher breeding value due to its higher grain productivity.

Cytological analysis of the progeny of two plants (DT 6/186/156 and DT 6/186/165) from the highly productive DT 6/186 line revealed instability in chromosome number and errors in chromosomal behavior in the first and second meiotic divisions. Chromosome staining using genomic *in situ* hybridization in these plants showed 14 rye chromosomes forming bivalents, which implied the presence of seven homologous pairs (Fig. 2, *a*). However, the bivalent chromosomes demonstrated premature separation in some meiocytes (desynapsis), with rye chromosomes becoming univalent and distributing anomalously between the poles (Fig. 3, *a*).

Chromosome number in the discovered aneuploid plants varied from $2n = 41$ to $2n = 44$. Among the ten plants from the DT 6/186/156 progeny, only one plant had chromosome number $2n = 42$ (see Fig. 2, *b*), while there were no plants with euploid chromosome numbers in the progeny of DT 6/186/165.

Univalents were discovered in the metaphase I (see Fig. 2, *d*), which were lagging at the equator during chromosome separation in anaphase I in 86.75 ± 4.56 and 61.32 ± 2.81 % of cells in DT 6/186/156 and DT 6/186/165 respectively (Table 2, see Fig. 3, *c*, *d*). The lagging chromosomes were divided into sister chromatids (see Fig. 3, *a*, *c*, *d*)

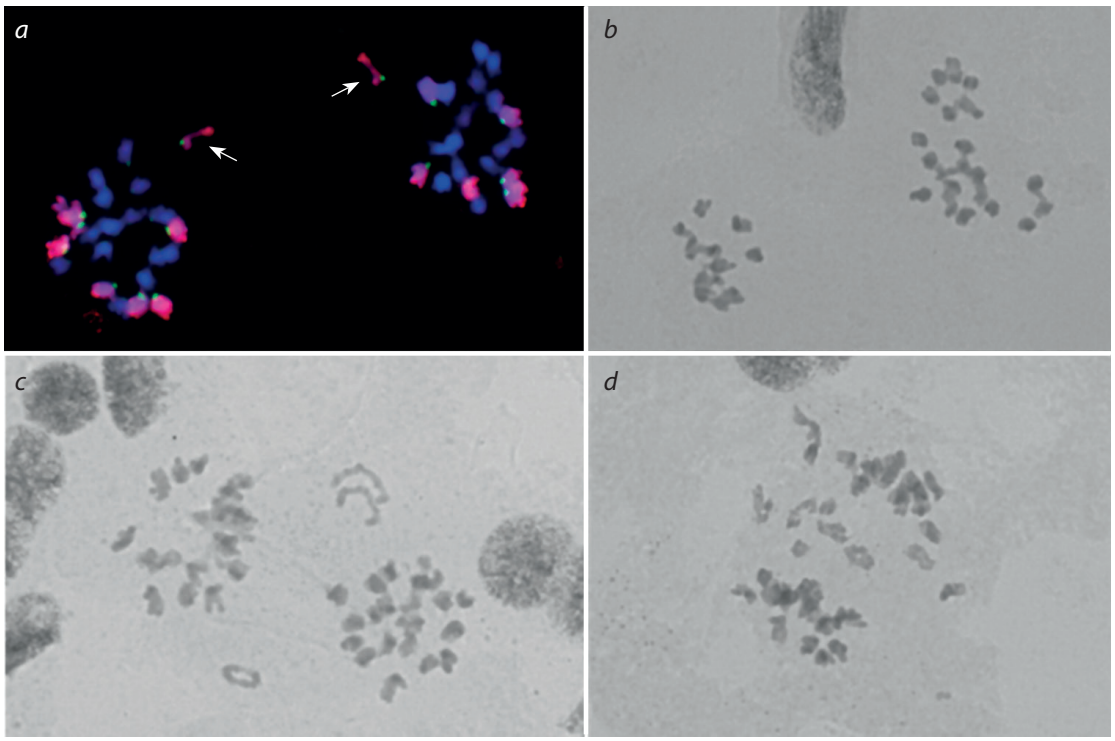


Fig. 3. Chromosome separation anomalies in meiotic anaphase I in the progeny of DT 6/186/165: *a*, unequal separation of wheat and rye chromosomes, the rye chromosome is divided into sister chromatids (arrows); *b*, unequal chromosome separation; *c*, *d*, lagging chromosomes at the equator and univalent division into sister chromatids.

a, genomic *in situ* hybridization, rye chromosomes stained red; *b*–*d*, acetocarmine staining.

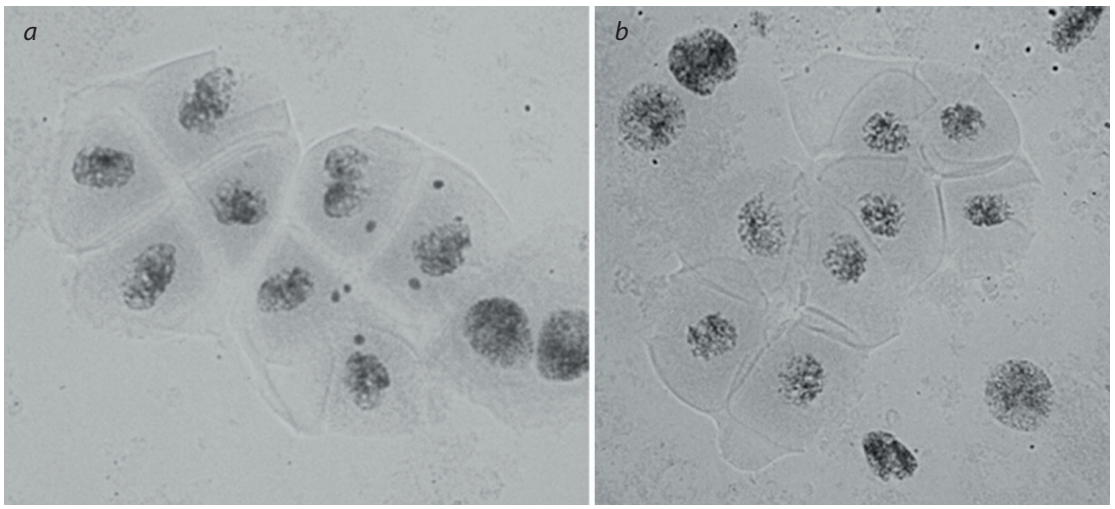


Fig. 4. Tetrads with (*a*) and without (*b*) micronuclei.

or broken at the centromere. Instability of chromosome separation during the division resulted in micronuclei formation at the tetrad stage (Fig. 4, *a*).

Micronuclei were observed in 60.29 ± 3.14 and 72.16 ± ± 2.29 % tetrads in DT 6/186/156 and DT 6/186/165, respectively (see Table 2). Even the euploid plant with $2n = 42$ (DT 6/186/156) had micronuclei in 51.48 % tetrads, which prevented us from considering this plant fully cytogenetically stable.

Plants with the minimum number of anomalous meiocytes in anaphase I and telophase II were selected from populations DT 6/186/156 and DT 6/186/165 based on the results of the analysis.

Discussion

Triticale (× *Triticosecale* Wittmack) as an agricultural crop combines wheat's high yield potential with rye's resistance to biotic and abiotic stresses, which increases its adaptability

Table 2. Analysis of meiotic chromosomal behavior in F_8 plants

Selection form	Total analyzed in anaphase I		Number of meiocytes with anomalous chromosome separation in anaphase I, %	Total analyzed by tetrads		Number of tetrads with micronuclei, %
	plants	cells		plants	cells	
DT 6/186/156	10	534	86.75 ± 4.56	8	1830	60.29 ± 3.14
DT 6/186/165	9	698	61.32 ± 2.81	8	2866	72.16 ± 2.29

to cultivation conditions in salty or highly acidic soils and in presence of toxic heavy metals. Thanks to these traits, triticale is of great interest as an emergency crop ensuring stable gross harvest of forage and food grains at lower costs (McGoverin et al., 2011). Despite today's applications of triticale grains being mostly restricted to forage in animal husbandry and production of feed and bioethanol, there is a rising interest in the use of triticale grains in human food products. Triticale grains have proven nutritional and dietary value (Meleshkina et al., 2015; Zhu, 2018), since they include not only proteins, carbohydrates, and fats, but also vitamins, minerals, and dietary fibres (14–18 %) (Rakha et al., 2011; Zhu, 2018). Compared to that of wheat, the protein from its grains has a richer amino acid profile, particularly in indispensable amino acids, such as lysine, threonine, and leucine (Meleshkina et al., 2015; Torikov et al., 2019). Triticale starch amounting to 3/4 of the kernel weight has a significantly lower amylose content compared to rye and wheat (Zhu, 2018), which ensures its better digestion by humans (Meleshkina et al., 2015).

To increase the share of triticales in production of breads and pastries, in recent decades breeding efforts have been aimed at increasing the quality of grains and finished products, which has resulted in development of triticale bread-making varieties (Grabovets et al., 2013), so State standards for triticale flour have been developed (State Standard 34142-2017). The characteristics of winter triticale varieties also include applicability for bread making. Triticale breeding efforts in Russia and other countries are primarily aimed at developing winter varieties (State Register…, 2022). However, grain quality assessment of the samples from the spring triticale collection showed that spring triticales have good potential for creating bread-making varieties (Krokhmal, Grabovets, 2015; Bocharnikova et al., 2017; Abdelkawy et al., 2020; Yerzhebayeva et al., 2020). The samples with such improved parameters as protein content, test weight, falling number, vitreousness, gluten quantity and quality, etc., are used in breeding for improved grain quality (Krokhmal, Grabovets, 2015; Bocharnikova et al., 2017; Turbayev et al., 2019; Abdelkawy et al., 2020; Yerzhebayeva et al., 2020).

A trait such as high test weight can be transferred by distant hybridization. This parameter is closely related to genetically determined parameters such as grain plumpness, hardness, and shape. In the present paper, tetraploid emmer wheat *T. dicoccum* (Schrank) Schuebl. with long, large, and plump grains was used as the maternal line for hybridization with hexaploid triticale (× *Triticosecale* Wittmack). As a result, new hexaploid forms of triticale with AABBRR genome types were created, which is confirmed by the analysis of meiotic chromosomal behavior using genomic *in situ* hybridization. Seven pairs of rye chromosomes and 27 to 30 wheat chromosomes were observed in the plants, which implies the presence of a complete rye genome and two wheat genomes. Three plants, DT4, DT5, and DT6, were isolated in the F_4 population, the progeny of which demonstrated test weight values significantly exceeding those of the initial triticale line, and test weights in the progeny of DT 4/168, DT 5/176, and DT 6/186 forms from F_6 were higher or on par with those of the initial triticale line. These plants had different productivity values, with the highest ones in F_4 recorded in the DT6 line (785 ± 41 g/m^2). The respective value for DT 6/186, i. e., the progeny of the latter in F_6, reached 822 ± 74 g/m^2. The DT 4/168 and DT 5/176 lines are of moderate interest for further research into improvement of bread-making properties of triticale grains.

The study of the meiotic behavior of rye and wheat chromosomes in the progeny of F_8 plants DT 6/186/156 and DT 6/186/165 showed that they had not yet achieved cytogenetic stability, which was indicated by the discovered chromosome separation errors and the presence of aneuploids in the populations. Chromosome separation errors mostly occurred in wheat chromosomes due to their monosomy. Cytological instability and aneuploidy in octoploid and hexaploid wheat-rye allopolyploids have been an issue from the start (Shkutina, Khvostova, 1971; Kaltsikes, 1974; Weimarck, 1974; Lukaszewski, Gustafson, 1987), but secondary triticales turned out to be more cytogenetically stable than primary ones (Kaltsikes, 1974). Cytological study of triticales showed that the interaction between wheat and rye genomes in the cells of the same plant resulted in physiological defects in cells persisting for decades at least. Meiotic and mitotic errors were found both in the triticale obtained by Rimpau in 1888 (Levitsky, 1978; Müntzing, 1974) and in the triticales obtained later. Despite the complete set of chromosomes, meiotic univalents were found in triticales with different ploidy levels (Shkutina, Khvostova, 1971; Kaltsikes, 1974; Lukaszewski, Gustafson, 1987; Olesczuk, Lukaszewski, 2014; Orlovskaya et al., 2015). The study into meiotic chromosomal behavior in triticales in the present and earlier papers demonstrated that while only bivalents were present at the diakinesis stage, univalents appeared in metaphase I as a result of to produce lagging chromosomes at the equator (Shkutina, Khvostova, 1971). Presumably, aneuploid cells in triticales may emerge as a result of asynchronous separation of rye and wheat chromosomes and their lagging in anaphase and telophase (Shkutina, Khvostova, 1971). The presumed cause of meiotic instability in the obtained am-

phidiploids is an imbalance in the genetic system of meiotic pairing and the differences in cell cycles between wheat and rye (Müntzing, 1974).

It is known for a fact that chromosome separation depends on kinetochore function (Sanei et al., 2011). It was found that kinetochore protein CENH3 produced by one of the parents maintained the function of other parental kinetochores in stable hybrids, despite the differences between DNA sequences of centromere regions in parental lines (Ishii et al., 2016). It was also shown that cytogenetic stability in triticales could also be linked to increased expression of rye-specific CENH3 forms in a hybrid genome (Evtushenko et al., 2019).

Another possible cause of cytogenetic instability of the obtained triticales could be a nuclear-cytoplasmic incompatibility, because the hybrids are bred using an intergeneric F_1 hybrid of emmer lines (*T. dicoccum* (Schrank) Schuebl.) (L133 × PKK) × k-25516 (genome AABB) as the maternal form. The AABB genomes in paternal triticale forms originate from soft wheat, while the rye genome is used as the base. Selection of genotypes of wheat varieties during triticale backcrossing may lead to recovery of fertility and cytogenetic stability in new forms, as exemplified by alloplasmic soft wheat lines (*H. vulgare*)-*T. aestivum* (Pershina et al., 1999, 2018; Trubacheeva et al., 2021).

Conclusion

The result of this work was to create and study hybrids of emmer *T. dicoccum* (Schrank) Schuebl. with hexaploid triticale. According to the characteristics of productivity and test weight values of the grain, hexaploid prebreeding forms F_6 were isolated – DT 4/168, DT 5/176 and DT 6/186. Thus, hexaploid forms DT 4/168 and DT 5/176 with well-filled, smooth grains and high test weights that can be used as the source of said traits in food-grade triticale breeding have been obtained. The DT 6/186 line shows promise in terms of breeding for high yields. As a result of the analysis of meiotic division in DT 6/186 plants, this sample is unstable for wheat chromosomes, therefore, aneuploids will occur in the offspring. To restore fertility and cytogenetic stability of new forms of triticale with emmer cytoplasm and chromosomes, it is necessary to select the genotypes of wheat varieties during backcrossing.

References

Abdelkawy R.N.F., Turbayev A.Zh., Solov'iev A.A. Technological properties of spring triticale grain grown in the Central Economic Region, Russia. *Vestnik Kurskoy Gosudarstvennoy Selskokhozyastvennoy Akademii = Bulletin of the Kursk State Agricultural Academy.* 2020;5:87-97. (in Russian)

Ammar K., Mergoum M., Rajaram S. The history and evolution of triticale. In: Triticale Improvement and Production (FAO Plant Production and Protection Paper. 179). Rome, 2004;1-10.

Ayalew H., Kumssa T.T., Butler T.J., Ma X.-F. Triticale improvement for forage and cover crop uses in the Southern Great Plains of the United States. *Front. Plant Sci.* 2018;9:1130. DOI 10.3389/fpls.2018.01130.

Badaev N.S., Badaeva E.D., Bolsheva N.L., Maximov N.G., Zelenin A.V. Cytogenetic analysis of forms produced by crossing hexaploid triticale with common wheat. *Theor. Appl. Genet.* 1985;70(5):536-541. DOI 10.1007/BF00305987.

Balyan H.S., Fedak G. Meiotic study of hybrids between barley (*Hordeum vulgare* L.) and Triticale (× *Triticosecale* Wittmack). *J. Hered.* 1989;80(6):460-463. DOI 10.1093/oxfordjournals.jhered.a110898.

Bocharnikova O.G., Gorbunov V.N., Shevchenko V.E. Assessment of spring triticale varieties in performance and grain quality. *Vestnik Voronezhskogo Gosudarstvennogo Agrarnogo Universita = Vestnik of Voronezh State Agrarian University.* 2017;2(53):23-30. DOI 10.17238/issn2071-2243.2017.2.23. (in Russian)

Cheng Z.J., Murata M. Loss of chromosomes 2R and 5RS in octoploid triticale selected for agronomic traits. *Genes Genet. Syst*. 2002;77(1):23-29. DOI 10.1266/ggs.77.23.

Dospekhov B.A. Methodology of Field Experiments. Moscow: Agropromizdat Publ., 1985. (in Russian)

Dou Q., Tanaka H., Nakata N., Tsujimoto H. Molecular cytogenetic analyses of hexaploid lines spontaneously appearing in octoploid Triticale. *Theor. Appl. Genet.* 2006;114(1):41-47. DOI 10.1007/s00122-006-0408-x.

Evtushenko E.V., Lipikhina Y.A., Stepochkin P.I., Vershinin A.V. Cytogenetic and molecular characteristics of rye genome in octoploid triticale (× *Triticosecale* Wittmack). *Comp. Cytogenet*. 2019;13(4):423-434. DOI 10.3897/CompCytogen.v13i4.39576.

Francki M.G. Identification of Bilby, a diverged centromeric T1-copia retrotransposon family from cereal rye (*Secale cereale* L.). *Genome*. 2001;44(2):266-274. DOI 10.1139/g00-112.

Fu S., Ren Z., Chen X., Yan B., Tan F., Fu T., Tang Z. New wheat-rye 5DS-4RS·4RL and 4RS-5DS·5DL translocation lines with powdery mildew resistance. *J. Plant Res.* 2014;127(6):743-753. DOI 10.1007/s10265-014-0659-6.

Grabovets A.I., Krokhmal A.V., Dremucheva G.F., Karchevskaya O.E. Breeding of triticale for baking purposes. *Russian Agricultural Sciences*. 2013;39(3):197-202.

Gruszecka D., Tarkowski C., Stefanowska G., Marciniak K. Cytological analysis of F_1 Aegilops hybrids with Triticosecale. In: Guedes-Pinto H., Darvey N., Carnide V.P. (Eds.) Triticale: Today and Tomorrow. Developments in Plant Breeding. Vol. 5. Dordrecht: Springer, 1996;195-201. DOI 10.1007/978-94-009-0329-6_24.

Gupta P.K., Fedak G. Intergeneric hybrids between × *Triticosecale* cv. Welsh (2*n* = 42) and three genotypes of *Agropyron intermedium* (2*n* = 42). *Can. J. Genet. Cytol.* 1986a;28(2):176-179. DOI 10.1139/g86-024.

Gupta P.K., Fedak G. Variation in induction of homoeologous pairing among chromosomes of 6*x Hordeum parodii* as a result of three triticale (× *Triticosecale* Wittmack) cultivars. *Can. J. Genet. Cytol.* 1986b;28(3):420-425. DOI 10.1139/g86-062.

Gustafson J.P., Bennett M.D. Preferential selection for wheat-rye substitutions in 42-chromosome triticale. *Crop Sci.* 1976;16(5):688-693. DOI 10.2135/cropsci1976.0011183X001600050022x.

Hao M., Luo J., Zhang L., Yuan Z., Yang Y., Wu M., Chen W., Zheng Y., Zhang H., Liu D. Production of hexaploid triticale by a synthetic hexaploid wheat-rye hybrid method. *Euphytica.* 2013;193:347-357. DOI 10.1007/s10681-013-0930-2.

Ishii T., Karimi-Ashtiyani R., Houben A. Haploidization via chromosome elimination: means and mechanisms. *Annu. Rev. Plant Biol*. 2016;67:421-438. DOI 10.1146/annurev-arplant-043014-114714.

Ivanova Yu.N., Solovey L.A., Loginova D.B., Miroshnikova E.E., Dubovets N.I., Silkova O.G. The creation and characterization of the bread wheat line with a centric translocation T2DL.2RL. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2019;23(7):846-855. DOI 10.18699/VJ19.558. (in Russian)

Jenkins B.C. History of the development of some presently promising hexaploid *Triticales*. *Wheat Inform. Serv*. 1969;28:18-20.

Kaltsikes P.J. Univalency in triticale. In: Macintyre R., Campbell M. (Eds.) Triticale: Proceedings of an international symposium. El Batan, Mexico: Int. Develop. Res. Centre, 1974;159-167.

Kang H., Wang H., Huang J., Wang Y., Li D., Diao C., Zhu W., Tang Y., Wang Y., Fan X., Zeng J., Xu L., Sha L., Zhang H., Zhou Y. Di-

vergent development of hexaploid triticale by a wheat–rye–*Psathyrostachys huashanica* trigeneric hybrid method. *PloS One*. 2016; 11(5):e0155667. DOI 10.1371/journal.pone.0155667.

Krokhmal A.V., Grabovets A.I. Triticale grain quality formation. *Izvestiya Orenburgskogo Gosudarstvennogo Agrarnogo Universiteta = Proceedings of the Orenburg State Agrarian University.* 2015;2(52): 46-48. (in Russian)

Krolow K.D. Aneuploidie und Fertilitat bei amphidiploiden Weizen-Roggen-Bastarden (*Triticale*). I. Aneuploidie un Selection auf Fertilität bei oktoploiden Triticale-Formen. *Z. Pflanzenzucht*. 1962;48(2): 177-196.

Kwiatek M., Belter J., Majka M., Wiśniewska H. Allocation of the S-genome chromosomes of *Aegilops variabilis* Eig. carrying powdery mildew resistance in triticale (× *Triticosecale* Wittmack). *Protoplasma.* 2016;253:329-343. DOI 10.1007/s00709-015-0813-6.

Kwiatek M., Majka M., Wiśniewska H., Apolinarska B., Belter J. Effective transfer of chromosomes carrying leaf rust resistance genes from *Aegilops tauschii* Coss. into hexaploid triticale (× *Triticosecale* Witt.) using *Ae. tauschii* × *Secale cereale* amphiploid forms. *J. Appl. Genet.* 2015;56:163-168. DOI 10.1007/s13353-014-0264-3.

Levitsky G.A. On the history of prolific intermediate constant wheat-rye hybrids. In: Levitsky G.A. Plant Cytogenetics. Moscow: Nauka Publ., 1978;251-253. (in Russian)

Levitsky G.A., Benetzkaja G.K. Cytology of the wheat-rye amphidiploids. In: Levitsky G.A. Plant Cytogenetics. Moscow: Nauka Publ., 1978;224-250. (in Russian)

Li H., Guo X., Wang C., Ji W. Spontaneous and divergent hexaploid triticales derived from common wheat × rye by complete elimination of D-genome chromosomes. *PLoS One*. 2015;10(3):e0120421. DOI 10.1371/journal.pone.0120421.

Lukaszewski A.J., Gustafson J.P. Translocations and modifications of chromosomes in triticale × wheat hybrids. *Theor. Appl. Genet.* 1983; 63(1):49-55. DOI 10.1007/BF00303771.

Lukaszewski A.J., Gustafson J.P. Cytogenetics of triticale. In: Janick J. (Ed.) Plant Breeding Reviews. 1987;5:41-93. DOI 10.1002/9781118061022.ch3.

McGoverin C.M., Snyders F., Muller N., Botes W., Fox G., Manley M. A review of triticale uses and the effect of growth environment on grain quality. *J. Sci. Food Agric*. 2011;91(7):1155-1165. DOI 10.1002/jsfa.4338.

Meister G.K. Natural hybridization of wheat and rye in Russia. *J. Hered.* 1921;12(10):467-470. DOI 10.1093/oxfordjournals.jhered.a102049.

Meleshkina E.P., Pankratieva I.A., Politukha O.V., Chirkova L.V., Zhiltsova N.S. Evaluation of triticale grain quality. *Khleboprodukty = Grain Products.* 2015;2:48-49. (in Russian)

Mergoum M., Singh P.K., Peña R.J., Lozano-del Río A.J., Cooper K.V., Salmon D.F., Gómez-Macpherson H. Triticale: a "new" crop with old challenges. In: Carena M. (Ed.) Cereals. Handbook of Plant Breeding. Vol. 3. New York: Springer, 2009. DOI 10.1007/978-0-387-72297-9_9.

Merker A. Chromosome composition of hexaploid triticale. *Hereditas.* 1975;80(1):41-52. DOI 10.1111/j.1601-5223.1975.tb01498.x.

Müntzing A. Genetic Research. Stockholm: Lts Förlag, 1961. (Russ. ed.: Muentzing A. Geneticheskie Issledovaniya. Moscow, 1963. (in Russian)

Müntzing A. Historical review of the development of triticale. In: Macintyre R., Campbell M. (Eds.) Triticale: Proceedings of an international symposium. El Batan, Mexico, 1–3 October 1973. Int. Develop. Res. Centre, 1974;13-30.

Neumann M., Kison H.U. Hybridization between hexaploid triticale and *Triticum monococcum* L. II. The F_1 generation A^uA^bBR. *Hereditas*. 1992;116(3):291-294. DOI 10.1111/j.1601-5223.1992.tb00157.x.

Oettler G. The fortune of a botanical curiosity – Triticale: past, present and future. *J. Agric. Sci.* 2005;143(5):329-346. DOI 10.1017/S0021859605005290.

Olesczuk S., Lukaszewski A.J. The origin of unusual chromosome constitutions among newly formed allopolyploids. *Am. J. Bot.* 2014; 101(2):318-326. DOI 10.3732/ajb.1300286.

Orlovskaya O.A., Leonova I.N., Adonina I.G., Salina E.A., Khotyleva L.V., Shumny V.K. Molecular cytogenetic analysis of triticale and wheat lines with introgressions of the genetic material of Triticeae tribe species. *Russ. J. Genet. Appl. Res.* 2016;6(5):527-536. DOI 10.1134/S2079059716050087.

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

Pershina L.A., Numerova O.M., Belova L.I., Devyatkina E.P., Shumny V.K. Restoration of fertility in backcross progeny of barley-wheat hybrids *Hordeum vulgare* L. ($2n = 14$) × *Triticum aestivum* L. ($2n = 42$) in relation to wheat genotypes involved in backcrosses. *Russ. J. Genet.* 1999;35(2):176-183.

Pisarev V.E. Breeding of Grain Crops. Moscow, 1964. (in Russian)

Rakha A., Åman P., Andersson R. Dietary fiber in triticale grain: variation in content, composition, and molecular weight distribution of extractable components. *J. Cereal Sci.* 2011;54(3):324-331. DOI 10.1016/j.jcs.2011.06.010.

Sanei M., Pickering R., Kumke K., Nasuda S., Houben A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proc. Natl. Acad. Sci. USA*. 2011;108(33):E498-E505. DOI 10.1073/pnas.1103190108.

Shishkina A.A., Badaeva E.D., Soloviev A.A. Features of karyotype organization in some forms of spring triticale. *Izvestiya Timiryazevskoy Selskokhozyaystvennoy Akademii = Proceedings of the Timiryazev Agricultural Academy*. 2009;2:123-130. (in Russian)

Shkutina F.M., Khvostova V.V. Cytological investigation of *Triticale. Theor. Appl. Genet.* 1971;41:109-119. DOI 10.1007/BF00277752.

Shulyndin A.F. Genetic grounds for the synthesis of various triticales and their improvement by breeding. In: Triticale: Study and Breeding. Leningrad: VIR Publ., 1975;53-56. (in Russian)

Silkova O.G., Ivanova Y.N., Loginova D.B., Solovey L.A., Sycheva E.A., Dubovets N.I. Karyotype reorganization in wheat–rye hybrids obtained via unreduced gametes: is there a limit to the chromosome number in triticale? *Plants.* 2021;10(10):2052. DOI 10.3390/plants10102052.

State Register for Selection Achievements Admitted for Usage (National List). Vol. 1. Plant varieties (official publication). Moscow: Rosinformagrotech, 2022. (in Russian)

State Standard 34142-2017. Tritical Flour: Specifications. 2018. Available at: https://docs.cntd.ru/document/1200146040 (in Russian)

Stepochkin P.I. The appearance of 6*x* triticale plants among the C_2 offspring of homogenomic 8*x* triticale. *Genetika = Genetics* (*Moscow*).1978;14(9):1658-1659. (in Russian)

Stepochkin P.I., Vladimirov N.S. Chromosome number, seed set, and winter hardiness in C_1 winter lines of homogenomic 8x triticale. *Genetika = Genetics* (*Moscow*). 1978;14(9):1597-1603. (in Russian)

Stepochkina N.I., Stepochkin P.I. Use of microchondrometer for determination of grain nature of single plants of triticale and wheat. *Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC*. 2015;29(11):39-40. (in Russian)

Torikov V.Ye., Shpilev N.S., Mameyev V.V., Yatsenkov I.N. Comparative characteristics of grain quality of winter triticale varieties grown in the south-west of Russia. *Vestnik Altayskogo Gosudarstvennogo Agrarnogo Universiteta = Bulletin of the Altai State Agricultural University.* 2019;2(172):49-56. (in Russian)

Trubacheeva N.V., Divashuk M.G., Chernook A.G., Pershina L.A. The effect of chromosome arm 1BS on the fertility of alloplasmic

recombinant lines in bread wheat with the *Hordeum vulgare* cytoplasm. *Plants* (*Basel*). 2021;10(6):1120. DOI 10.3390/plants10061120.

Turbayev A.Z., Sergaliev N.K., Soloviev A.A. Comparative study of some genotypes of spring triticale for grain qualities. *Izvestiya Timiryazevskoy Selskokhozyaystvennoy Akademii = Izvestiya of Timiryazev Agricultural Academy*. 2019;1:19-33. (in Russian)

Weimarck A. Elimination of wheat and rye chromosomes in a strain of octoploid *Triticale* as revealed by Giemsa banding technique. *Hereditas.* 1974;77(2):281-286. DOI 10.1111/j.1601-5223.1974.tb00940.x.

Yerzhebayeva R.S., Tajibaev D., Abugalieva A.I. Grain quality of samples of the spring triticale collection (× *Triticosecale* Wittmack)]. *Sibirskii Vestnik Sel'skokhozyaistvennoi Nauki = Siberian Herald of Agricultural Science*. 2020;50(3):111-121. DOI 10.26898/0370-8799-2020-3-12. (in Russian)

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

Zhou J., Zhang H., Yang Z., Li G., Hu L., Lei M. Characterization of a new T2DS.2DL-?R translocation triticale ZH-1 with multiple resistances to diseases. *Genet. Resour. Crop Evol.* 2012;59:1161-1168. DOI 10.1007/s10722-011-9751-0.

Zhu F. Triticale: nutritional composition and food uses. *Food Chem*. 2018;241:468-479. DOI 10.1016/j.foodchem.2017.09.009.

ORCID ID

O.G. Silkova orcid.org/0000-0003-3299-2975
Yu.N. Ivanova orcid.org/0000-0002-9655-4539

Acknowledgements. This study was supported by ICG SB RAS budget project FWNR-2022-0017.

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

Received July 26, 2022. Revised March 29, 2023. Accepted March 30, 2023.

Вавиловский журнал генетики и селекции. 2023;27(4):333-341
DOI 10.18699/VJGB-23-40

Studying sex differences in responses to fibroblast growth factor 21 administration in obese mice consuming a sweet-fat diet

N.M. Bazhan[1, 3] ✉, T.V. Jakovleva[1], A.Yu. Kazantseva[1], N.E. Kostina[1], P.E. Orlov[3], N.Yu. Balybina[1], K.O. Baranov[2], E.N. Makarova[1]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
✉ bazhan-nm@yandex.ru

Abstract. In animals, obesity caused by consumption of a sweet-fat diet (SFD) is the most adequate mouse model of human diet-induced obesity. Fibroblast growth factor 21 (FGF21) reduces body weight, beneficially affects taste preferences, and corrects glucose metabolism in obese mice. Sex is known to influence FGF21 effects in different models of diet-induced and hereditary obesity. In mice with SFD-induced obesity, the effects of FGF21 have been studied only in males. The aim of this study was to compare the effects of FGF21 on body weight, food preferences and glucose and lipid metabolism in C57Bl/6J male and female mice with SFD-induced obesity. Mice were fed with a diet consisting of standard chow, lard and cookies for 10 weeks, then they were injected with FGF21 (1 mg per 1 kg) or vehicle for 7 days. Body weight, weights of different types of food, blood parameters, glucose tolerance, gene and protein expression in the liver, gene expression in the white, brown adipose tissues, and the hypothalamus were assessed. FGF21 administration reduced body weight, did not alter total energy consumption, and activated orexigenic pathways of hypothalamus in mice of both sexes. However, sex dimorphism was found in the realization of the orexigenic FGF21 action at the transcriptional level in the hypothalamus. Metabolic effects of FGF21 were also sex-specific. Only in males, FGF21 exerted beneficial antidiabetic action: it reduced fatty acid and leptin plasma levels, improved glucose-tolerance, and upregulated hepatic expression of *Ppargc1*, *Fasn*, *Accα*, involved in lipid turnover, gene *Insr* and protein glucokinase, involved in insulin action. Only in obese females, FGF21 induced preference of standard diet to sweet food. Thus, in mouse model of obesity induced by consumption of a sweet-fat diet, the catabolic effect of FGF21 was not sex-specific and hormonal, transcriptional and behavioral effects of FGF21 were sex-specific. These data suggest elaboration of different approaches to use FGF21 analogs for correction of metabolic consequences of obesity in different sexes.
Key words: FGF21; obesity; sex differences; liver; adipose tissue; hypothalamus; food intake; gene expression.

For citation: Bazhan N.M., Jakovleva T.V., Kazantseva A.Yu., Kostina N.E., Orlov P.E., Balybina N.Yu., Baranov K.O., Makarova E.N. Studying sex differences in responses to fibroblast growth factor 21 administration in obese mice consuming a sweet-fat diet. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):333-341. DOI 10.18699/VJGB-23-40

Изучение половых различий ответов на введение фактора роста фибробластов 21 у мышей с ожирением, вызванным потреблением сладко-жирной диеты

Н.М. Бажан[1, 3] ✉, Т.В. Яковлева[1], А.Ю. Казанцева[1], Н.Е. Костина[1], П.Е. Орлов[3], Н.В. Балыбина[1], К.О. Баранов[2], Е.Н. Макарова[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ bazhan-nm@yandex.ru

Аннотация. У животных ожирение, вызванное потреблением сладко-жирной диеты, служит наиболее адекватной моделью развития алиментарного ожирения у человека. У мышей с ожирением фактор роста фибробластов 21 (FGF21) снижает массу тела, благоприятно влияет на вкусовые предпочтения, улучшает метаболизм глюкозы и липидов. В различных моделях ожирения показано, что пол влияет на эффекты FGF21. У мышей с ожирением, вызванным сладко-жирной диетой, эффекты FGF21 изучены только у самцов. Цель работы заключалась в сравнении влияния FGF21 на массу тела, выбор пищи, углеводно-жировой обмен у самцов и самок мышей C57Bl/6J с ожирением, вызванным потреблением сладко-жирной диеты. В течение 10 недель мышей кормили диетой, состоящей из стандартного лабораторного корма, свиного сала и сладкого печенья, затем 7 дней они получали инъекции FGF21 (1 мг на 1 кг) или растворителя. Измерены масса тела, масса потребленных продуктов, показатели крови, толерантность к глюкозе, экспрессия в печени генов и белков, а в жировой ткани и гипоталамусе – только генов. У мышей обоих полов введение FGF21 снизило массу тела, не изменило общее количество потребленной энергии и активировало орексигенные пути в гипоталамусе. Однако на транскрип-

ционном уровне выявлены половые различия в реализации орексигенного действия FGF21 в гипоталамусе. Метаболическое действие FGF21 также зависело от пола. Только у самцов FGF21 вызывал благоприятное антидиабетическое воздействие: снижал уровни жирных кислот и лептина в крови, улучшал толерантность к глюкозе, повышал экспрессию в печени генов *Ppargc1*, *Fasn*, *Accα*, вовлеченных в жировой обмен, и гена *Insr*, а также белка глюкокиназы, которые опосредуют действие инсулина. Только у самок FGF21 увеличивал потребление энергии со стандартным кормом и снижал – с печеньем. Таким образом, у мышей в модели ожирения, вызванного сладко-жирной диетой, FGF21 уменьшал массу тела особей обоего пола, оказывал антидиабетический эффект только у самцов и менял вкусовые предпочтения только у самок. Эти результаты указывают на необходимость разработки специальных подходов в использовании FGF21 и его аналогов для лечения метаболических последствий ожирения у представителей разных полов.
Ключевые слова: FGF21; ожирение; половые различия; печень; жировая ткань; гипоталамус; потребление пищи; экспрессия генов.

Introduction

In the human population, there is a significant increase in the number of people suffering from obesity and associated metabolic diseases such as type 2 diabetes, cardiovascular diseases and non-alcoholic fatty liver. Intensive research is underway to create drugs for normalization of carbohydrate-lipid metabolism at obesity. Fibroblast growth factor 21 (FGF21) is purported to be one of the most promising candidates for such aims (Talukdar, Kharitonenkov, 2021). FGF21 is synthesized and secreted into the circulation mainly by the liver (Fisher et al., 2011) in response to metabolic stresses such as food deprivation (Zhang Y. et al., 2012; Bazhan et al., 2019a), cold exposure (Dutchak et al., 2012), and obesity (Chukijrungroat et al., 2017; Bazhan et al., 2019b). FGF21 functions to restore homeostasis by coordinating metabolic responses from brown and white adipose tissues, muscles, liver and hypothalamus (Martínez-Garza et al., 2019; Makarova et al., 2021a, b). The effect of FGF21 on metabolic phenotype is partially mediated by its effect on gene expression in these tissues (Hale et al., 2012; Keinicke et al., 2020). FGF21 administration has potent beneficial effects on obesity and diabetes in humans, monkeys, and rodents (Kharitonenkov, Adams, 2013), it reduces body weight, increases insulin sensitivity, normalizes blood glucose levels (Kharitonenkov et al., 2005; Coskun et al., 2008; Xu et al., 2009).

High energy density diets: high fat diet (HFD) and sweet fat diet (SFD) are most commonly used to induce obesity in mice. It has recently been shown that the anti-diabetic effect of FGF21 in genetically obese mice (ob/ob and Ay mice) appears only in males (Berglund et al., 2009; Makarova et al., 2020; Makarova et al., 2021b), whereas in mice with obesity caused by HFD it is manifested both in females and males. In mice with SFD-induced obesity (SFDIO), the effects of FGF21 have been shown in males (Coskun et al., 2008), while in females, they have not been studied. Elucidation of the role of FGF21 in the regulation of metabolic processes in females fed SFD is relevant, since SFD mimics the human Western diet and is the most adequate model of human diet-induced obesity (Sampey et al., 2011).

In addition, FGF21 may influence food choices. In mice with normal body weight, FGF21 administration has been shown to increase protein intake (Hill et al., 2020) and reduce sugar intake (Talukdar et al., 2016b). In ovariectomized obese female mice fed mixed diet (standard chows, lard and sweet cookies), FGF21 administration increases the intake of calories with standard chows, which contain more proteins than other types of food (Jakovleva et al., 2022). It is not known whether FGF21 influences the intake of various types of food and the expression of hypothalamic genes involved in the regulation of food intake in SFDIO mice with normal gonadal function.

Thus, the aim of this work was to study in male and female mice with obesity caused by the consumption of a sweet-fat diet the effect of FGF21 on weight of body, liver, adipose tissues, tolerance to glucose, blood levels of hormones and metabolites, the amount of energy consumed with various types of food, and expression of genes involved in the regulation of metabolic processes in the liver, white, brown adipose tissue and in the regulation of eating behavior in the hypothalamus.

Materials and methods

The experiment was performed according to the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (Council of Europe No 123, Strasbourg 1985) and 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 of the Siberian Branch of the Russian Academy of Sciences on November 8, 2021.

Animals. Male and female C57BL mice bread in the vivarium of the Institute of Cytology and Genetics were used. At the age of four weeks, mice were separated from mothers and placed in groups of three per cage under a 12/12-h light-dark regime (light from 07:30 to 19:30) at an ambient temperature of 22–24 °C. The animals were provided ad libitum ac cess to commercial mouse chow (Assortiment Agro, Turakovo Village, Moscow region, Russia) and water. At the age of 10 weeks, lard and sweet butter cookies were added to the standard chow in order to induce the development of obesity. The animals fed SFD for 13 weeks until they reached marked obesity (the body weight was 41.6 ± 1.0 g, mean $\pm$ SE, $n = 20$). To measure the food intake of each mouse, they were placed in separate cages and kept individually until the start of the experiment. Solitary maintenance is emotional stress, reducing body weight. Three days after the cage change, the body weight decreased to 38.9 ± 0.9 g. Two weeks after the cage change, the weight of the mice recovered to 41.8 ± 0.8 g. This body weight was considered as the initial.

Mice of both sexes were randomized into the control group injected with vehicle – phosphate-bicarbonate solution – and the FGF21 group injected with mouse recombinant FGF21 (1 mg per 1 kg). Each group consisted of 5–7 mice. Substances were administered subcutaneously at the end of the light period (17:00–17:30) for seven days. The protocol of the expression

and purification of mouse FGF21 was described previously (Makarova et al., 2021a). Mice were weighed daily. In the part of the animals, various food components were also weighed daily and the amount of energy intake was calculated based on the fact that the calorie content of lard is 8 kcal/g, the standard chow is 2.5 kcal/g, and the biscuit is 4.58 kcal/g. The percentage of consumed energy in relation to total consumed energy was calculated for each type of food.

A day after the last injection of FGF21, some of the mice were sacrificed by decapitation, and the others were tested for glucose tolerance and were sacrificed the day after testing. Trunk blood was collected in test tubes with EDTA after decapitation, centrifuged and plasma was stored at –20 °C until the assay of hormones and metabolites. Liver, subcutaneous and abdominal white adipose tissue (WAT), and brown adipose tissue (BAT) from the interscapular region were weighed. Samples of the liver, BAT, perigonadal WAT and hypothalamus were collected and snap-frozen in liquid nitrogen to measure gene expression. In the liver, protein expression was also measured by Western Blot Analysis.

Glucose tolerance test. Before the test, food was removed from the animals at 08:00, and the test started at 15:00. Animals were injected with glucose (AO "REACHEM", Moscow, Russia) intraperitoneally at the dose of 1 g/kg body weight. Blood glucose concentrations were measured using a Lifescan One Touch Basic Plus glucometer (LifeScan Inc., Switzerland) before glucose administration (fasting glucose) and 15, 30, 60, and 120 minutes after glucose administration.

Assay of plasma biochemical parameters. Concentrations of insulin, leptin, and adiponectin were measured using a Rat/Mouse insulin ELISA Kit, a Mouse leptin ELISA Kit (EMD Millipore, St. Charles, MO, USA), and a Mouse adiponectin ELISA Kit (EMD Millipore, Billerica, MA, USA), respectively. Concentrations of glucose, triglycerides, and cholesterol were measured colorimetrically using Fluitest GLU, Fluitest TG, and Fluitest CHOL (Analyticon® Biotechnologies AGAm Mühlenberg 10, 35104 Lichtenfels, Germany), respectively. Concentrations of free fatty acids were measured using NEFA FS DiaSys kits (DiaSys Diagnostic Systems GmbH, Holzheim, Germany).

Relative quantitation real-time PCR. RNA was isolated from tissue samples using an ExtractRNA kit (Evrogen, Moscow, Russia) according to the manufacturer's recommendations. 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, USA) were used for relative quantitation real-time PCR. The genes tested involved acetyl-CoA carboxylase alpha/beta (*Acacα*/β, Mm01304257_m1/ Mm01204671_m1), adipose triglyceride lipase (*Atgl*, Mm00503040_m1), agouti related neuropeptide (*Agrp*, Mm00475829_g1), beta-actin (*Actb*, Mm00607939_s1), carnitine palmitoyltransferase 1A/1B (*Cpt1*α/β, Mm01231183_m1/Mm00487191_g1), corticotrophin releasing hormone (*Crh*, Mm01293920_s1), fatty acid synthase (*Fasn*, Mm00662319_m1), fibroblast growth factor 21 (*Fgf21*, Mm00840165_g1), glucose-6-phosphatase (*G6pc*, Mm00839363_m1), Glucokinase (*Gck*, Mm00439129_m1), Insulin receptor (*Insr*, Mm01211875_m1), hormone-sensitive lipase (*Lipe*, Mm00495359_m1), klotho beta (*Klb*, Mm00473122_m1), neuropeptide Y (*Npy*, Mm01410146_m1), peroxisome proliferator-activated receptor alpha/gamma (*Pparα/γ*, Mm00440939_m1/Mm00440940_m1), peroxisome proliferator-activated receptor gamma coactivator (*Ppargc1α*, Mm01208835_m1), phosphoenolpyruvate carboxykinase (*Pck*, Mm01247058_m1), Pyruvate kinase liver and red blood cell (*Pklr*, Mm00443090_m1), proopiomelanocortin (*Pomc*, Mm00435874_m1), solute carrier family 2 member 1/2/4 (*Slc2a1/Slc2a2 Slc2a4*, Mm00441480_m1/Mm00446229_m1/Mm00436615_m1), uncoupling protein 1/3 (*Ucp1/Ucp3*, Mm01244861_m1/ Mm01163394_m1). Beta-actin was used as endogenous controls. The PCR and fluorescence detection were performed on an Applied Biosystems VIIA 7 Real-Time PCR System. Relative quantification was performed by the comparative threshold cycle (CT) method (i. e. $2^{-\Delta\Delta Ct}$ method).

Western Blot Analysis of protein levels. Expression of hepatic proteins was measured as described previously (Iakovleva et al., 2020). Primary polyclonal rabbit antibodies (Santa Cruz Biotechnology, USA, breeding 1:2000) were used: Insulin receptor antibody (IRa, N-20) and Glucokinase antibody (GK H-88). The results referred to the total amount of protein.

Statistical analysis. Each result is presented as mean ± SE for a sample size (i. e., number of mice) indicated. A repeated measures ANOVA with the factors "sex" (male, female), "experiment" (phosphate-bicarbonate solution, FGF21), and "day of experiment" was used to analyze FGF21 effects on body weight loss and total daily energy intake. Other parameters were compared with Student's *t-test*. Significance was determined as $p < 0.05$. The STATISTICA 6 software package (StatSoft Inc., USA) was used for analysis.

Results

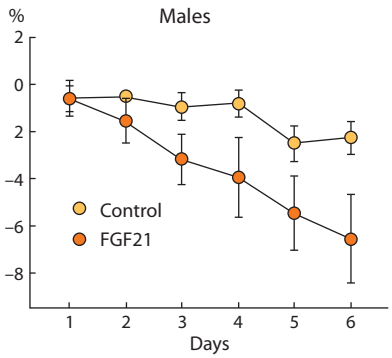
Weight characteristics. Injections of vehicle and FGF21 induced weight loss in both male and female mice ($p < 0.05$, factor "experiment"; Fig. 1). However, weight loss was more pronounced in FGF21 than in the control group ($p < 0.001$, interaction of "experiment" and "day of experiment" by repeated measures ANOVA) regardless of sex.

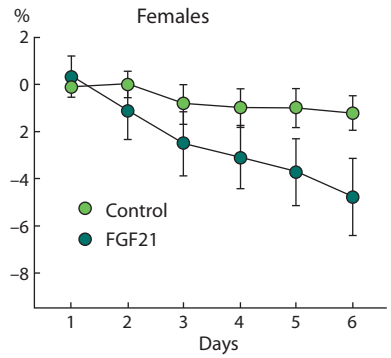
Sex and FGF21 administration did not affect indexes of the liver and adipose tissues (Table 1).

Energy consumption. In both groups, males and females did not differ in total energy consumption with different types of food. FGF21 administration did not affect total energy intake or individual food choices in males, but increased energy intake with standard food and decreased it with cookies ($p < 0.05$ in both cases) in females (Fig. 2, *a, b*). At the same time, in females treated with FGF21, the total energy intake from all types of food did not differ from that in the control group.

Blood biochemical parameters, glucose tolerance test. In the control group, no sex differences were found for most of the hormonal and metabolic blood parameters, only the adiponectin blood level in the females was higher than that of males ($p < 0.001$).

FGF21 administration reduced blood free fatty acid and leptin concentrations ($p < 0.05$ in both cases), tended to reduce insulin concentration ($p < 0.08$) (Fig. 3, *a*) and increased glucose tolerance only in males (see Fig. 3, *b*). In the glucose tolerance test in males treated with FGF21, the blood glucose curve was lower than in the control, and at the 15th and 30th





Repeated ANOVA

Effect	*p*
Sex	0.397
Exp	0.045
Sex*Exp	0.808
Day	0.001
Day*Sex	0.193
Day*Exp	0.001
Day*Sex*Exp	0.969

Fig. 1. Body weight loss (% of initial BW) in mice with sweet fat diet-induced obesity received vehicle (control) or FGF21.
Group size 5–7 animals.

Table 1. Weight characteristics in male and female obese mice that received vehicle (control) or FGF21 for 7 days

Tissue	Males		Females	
	Control	FGF21	Control	FGF21
Liver index × 10	0.35 ± 0.03	0.38 ± 0.03	0.37 ± 0.02	0.35 ± 0.01
scWAT index × 10	3.12 ± 0.04	2.74 ± 0.29	3.12 ± 0.18	3.10 ± 0.06
abdWAT index × 10	3.02 ± 0.04	2.81 ± 0.11	2.89 ± 0.17	2.89 ± 0.13
BAT index × 10	0.53 ± 0.02	0.53 ± 0.02	0.43 ± 0.04	0.59 ± 0.08

Note. scWAT, subcutaneous white adipose tissue; abdWAT, abdominal white adipose tissue; BAT, brown adipose tissue.

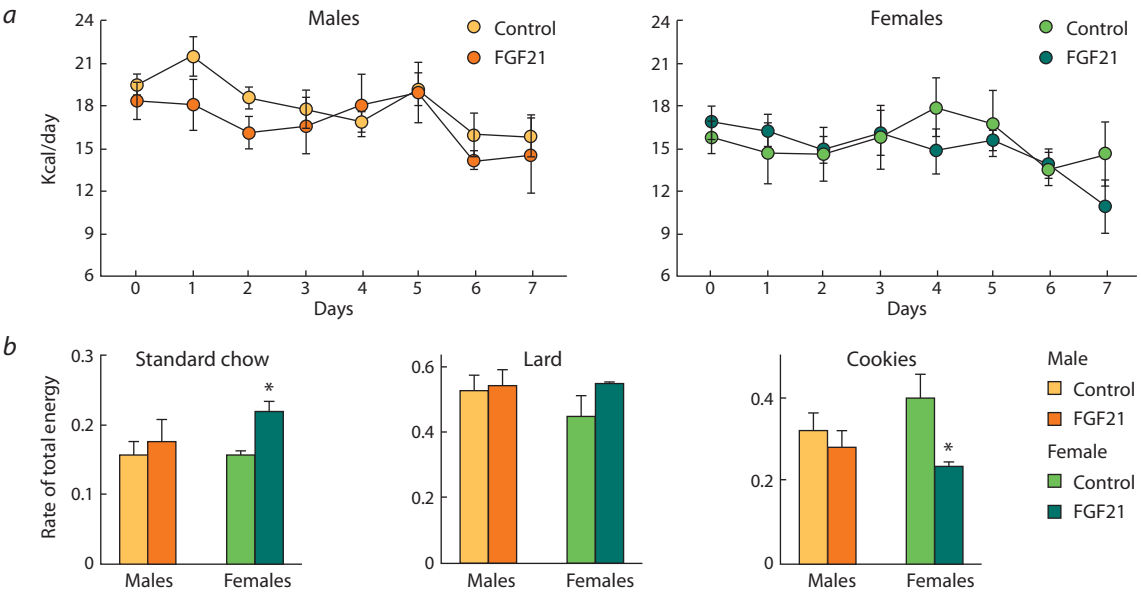


Fig. 2. Total daily energy intake with all types of food (*a*) and energy intake with each type of food for 7 days (*b*) in obese mice that received vehicle (control) or FGF21 for 7 days.
5–7 mice/group. * $p < 0.05$ vs. control in the same sex by Student's *t*-test.

minute of the test, the differences with the control were significant (* $p < 0.05$ vs. control in both cases).

Gene expression. In the control group, the expression of many hepatic genes in females was higher than in males (Fig. 4, 5). These include genes encoding peroxisome proliferator-activated receptor gamma coactivator (*Ppargc1*), carnitine palmitoyltransferase 1A (*Cpt1α*), fatty acid synthase (*Fasn*), acetyl-CoA carboxylase beta (*Acacβ*) and klotho beta (*Klb*) ($p < 0.01$ for *Acacβ*, and $p < 0.05$ for other genes). In addition, the expression of adipose triglyceride lipase (*Pnpla2*), solute carrier family 2 member 2 (*Slc2a2*), insulin receptor (*Insr*) genes in control females was higher than in SFDIO males at the trend level ($p = 0.06$ for *Pnpla2*, and $p = 0.07$ for *Insr* and *Slc2a2*) (see Fig. 4).

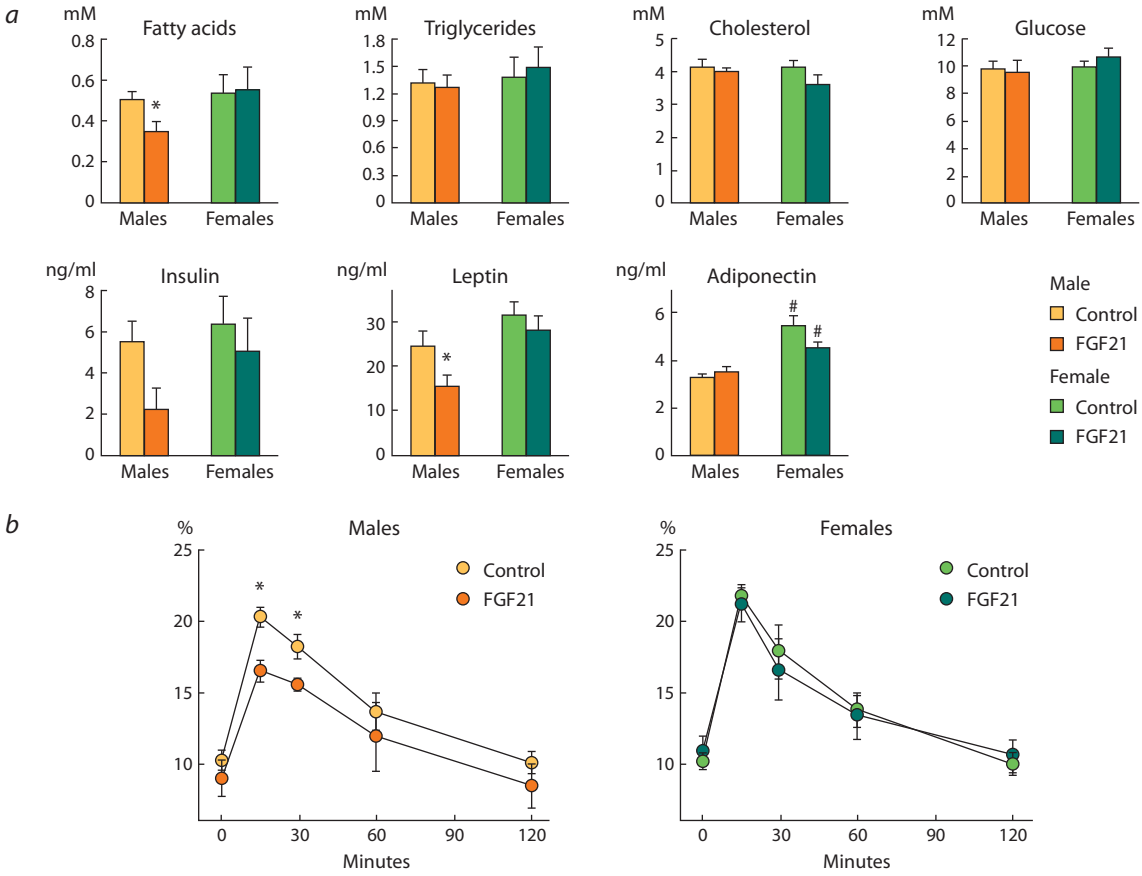


Fig. 3. Biochemical blood parameters (*a*) and blood glucose levels in the glucose tolerance test (*b*) in obese mice that received vehicle (control) or FGF21.

Here and in Fig. 4–7: 5–7 mice/group. * $p < 0.05$ vs. control in the same sex, # $p < 0.05$ vs. males in the same experimental group by Student's *t*-test.

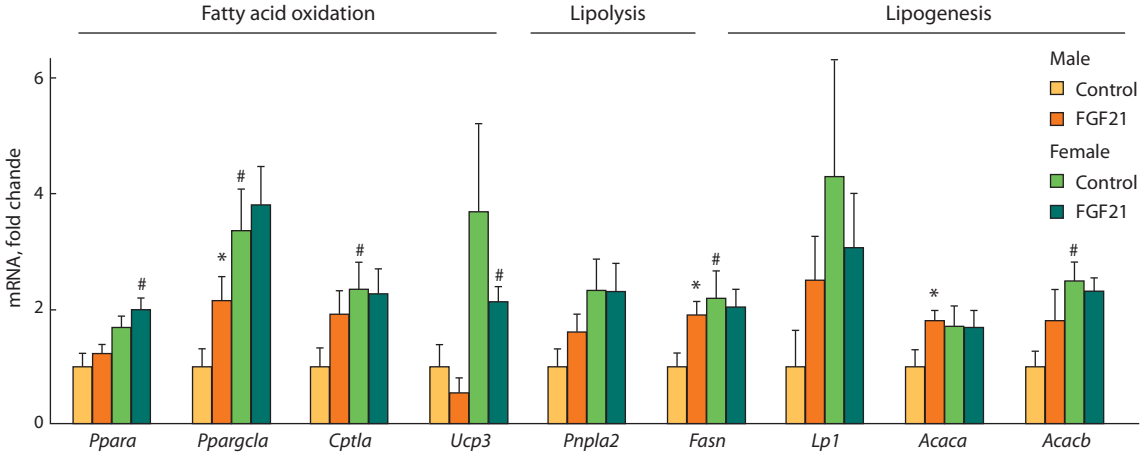


Fig. 4. Relative expression of hepatic genes involved in lipid metabolism in obese mice that received vehicle (control) or FGF21.

Only in males, administration of FGF21 increased expression of hepatic genes involved in fatty acid oxidation (*Ppargc1*), lipogenesis (*Fasn*, acetyl-CoA carboxylase α, *Acacα*) and insulin sensitivity (*Insr*) ($p < 0.05$ for all genes) (see Fig. 4, 5). In addition, in SFDIO males, FGF21 increased, on a tendency level, the expression of genes related to glucose oxidation (glucokinase, *Gck*, pyruvate kinase, *Pklr*) ($p < 0.06$ for *Pklr*, and $p < 0.07$ for *Gck*). After FGF21 administration, the expression of genes involved in fatty acid oxidation (peroxisome proliferator-activated receptor alpha, *Pparα* and uncoupling protein 3, *Ucp3*) was higher in females than in males ($p < 0.05$ for both genes).

As far as FGF21 administration upregulated the expression of genes encoding insulin receptor (IR) and glucokinase (GK), hepatic expression of these proteins was also measured (see Fig. 5, *b*). In the control group, IR expression did not differ

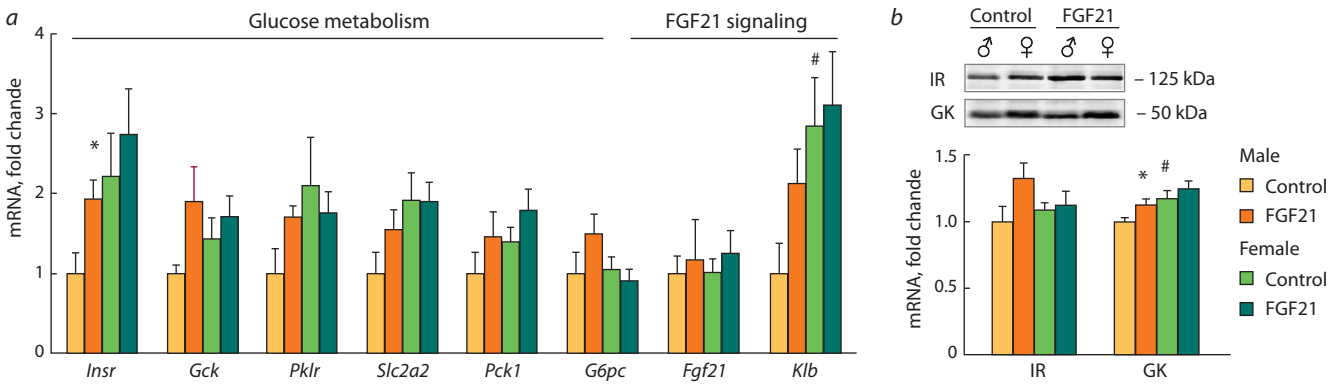


Fig. 5. Relative expression of hepatic genes (*a*) and proteins (*b*) involved in glucose metabolism and FGF21 signaling in male and female obese mice that received vehicle (control) or FGF21.

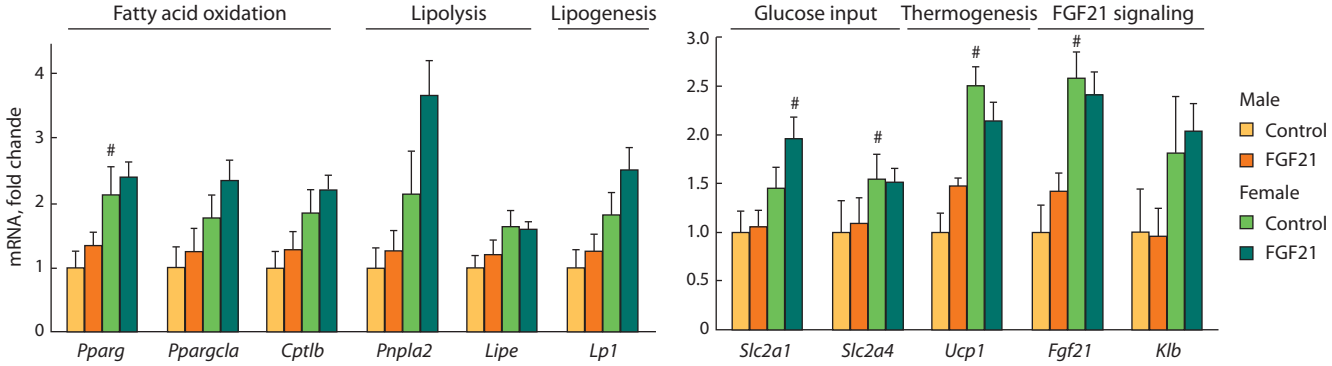


Fig. 6. Relative BAT gene expression in male and female obese mice that received vehicle (control) or FGF21.

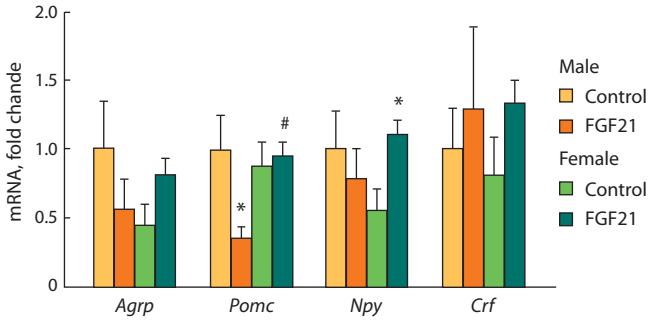


Fig. 7. Relative hypothalamic gene expression in male and female obese mice that received vehicle (control) or FGF21.

between females and males and GK expression was higher in females than in males ($p < 0.05$). FGF21 administration increased expression of GK ($p < 0.05$) and IR (tendency $p < 0.06$) only in males.

In obese mice from the control group, expression of BAT genes involved in fatty acid oxidation (peroxisome proliferator-activated receptor gamma, *Pparγ*), thermogenesis (uncoupling protein 1, *Ucp1*) and gene encoding FGF21 (*Fgf21*) was higher in females than in males ($p < 0.001$ for *Fgf21*, $p < 0.01$ for *Ucp1*, and $p < 0.05$ for *Pparγ*) (Fig. 6). FGF21 administration did not affect BAT gene expression. After FGF21 administration expression of *Slc2a1* encoding Solute carrier family 2 member 1 was significantly higher in females than in males ($p < 0.05$). In subcutaneous WAT, expression of all measured genes (*Pparγ*, *Ppargc1*, *Cpt1β*, *Pnpla2*, *Lipe*, *Lpl*, *Fasn*, *Slc2a1*, *Slc2a4*, *Ucp1*, *Insr*) was independent of sex and FGF21 administration (the data are not presented).

In the control group, males and females did not differ in the expression of hypothalamic genes involved in the regulation of food intake (Fig. 7). FGF21 administration increased the expression of the gene encoding orexigenic neuropeptide NPY in SFDIO females and decreased the expression of gene encoding anorexigenic neuropeptide POMC in SFDIO males ($p < 0.05$ for both genes). In the FGF21 group, the expression of the *Pomc* gene in females was higher than in males ($p < 0.05$).

Discussion

In this work, we evaluated the pharmacological effects of FGF21 in male and female mice in a model of obesity induced by consumption of SFD that mimics the human Western diet and is the most adequate model of human diet-induced obesity (Sampey et al., 2011). Results of this work demonstrated that in C57Bl mice with SFDIO, FGF21 exerted catabolic effect regardless of sex, and antidiabetic effect – differently in male and female mice. FGF21 administration exerted anti-diabetic therapeutic effects only in SFDIO males and increased consumption of standard chow and decreased consumption of sweet cakes only in SFDIO females.

The ability of FGF21 to reduce body weight, regardless of sex, was previously demonstrated in obese mice that consumed HFD (Makarova et al., 2021a). However, in Ay mice with genetic melanocortin obesity, the catabolic effect of FGF21 was manifested only in males (Makarova et al., 2020). Thus, FGF21-treated obese male mice exert reduced BW regardless of the type of obesity (hereditary, HFD- or SFD-induced obesity), and FGF21-treated obese female mice – only at alimentary form of obesity.

In SFDIO mice, the catabolic effect of FGF21 was not associated with a decrease in total energy intake, suggesting that it was due to an increase in energy expenditure. Indeed, T. Coskun et al. demonstrated that FGF21-treated male mice with obesity induced by consumption of a sweet-fat diet had a significantly higher energy expenditure rate, oxygen consumption as well as increased core body temperature, compared with the vehicle-treated mice (Coskun et al., 2008). The authors hypothesized that the FGF21-induced increase in energy expenditure could be due to increased thermogenesis in BAT, since the expression of genes involved in thermogenesis and fatty acid oxidation was increased in FGF21-treated male mice. In our experiment, FGF21 administration did not alter expression of genes (*Ucp1* and *Dio2*) related to thermogenesis in BAT and WAT in SFDIO mice. It can be assumed that in SFDIO mice, FGF21 increased energy expenditure through other mechanisms, unrelated to BAT and WAT thermogenesis.

It is known that FGF21 enhances mitochondrial biogenesis and fatty acid oxidation not only in adipose tissues, but also in muscles (Sun et al., 2021). In addition, FGF21 can increase energy expenditure by activating sympathetic nerves (Owen et al., 2014), which is accompanied by an increase in body temperature and physical activity (Owen et al., 2014). We demonstrated earlier that FGF21 increases locomotor activity equally in male and female Ay mice with genetic obesity (Makarova et al., 2021a). It can be assumed that the catabolic effect of FGF21 in SFDIO mice was due to activation of thermogenesis in muscles (Sun et al., 2021), and motor activity (Makarova et al., 2021a). Weight loss in FGF21-treated SFDIO mice is indicative of yet other mechanisms whereby energy expenditure may be increased, which remain to be explored.

The antiobesity effect of FGF21 was associated with activation of hypothalamic orexigenic mechanisms in mice of both sexes. However, sex dimorphism was found in the realization of the orexigenic FGF21 action at the transcriptional level: FGF21 upregulated the expression of the orexigenic *Npy* in SFDIO females, and downregulated the expression of the anorexigenic *Pomc* in SFDIO males. These data suggest that specific pathways for the orexigenic action of FGF21 may differ between males and females. Obviously, transcriptional changes found at the level of the hypothalamus do not contribute to body weight loss in response to FGF21 administration. It can be assumed that they are part of a counter-regulatory mechanism necessary for limiting the catabolic effect of FGF21.

Our data demonstrated that FGF21 had a beneficial antidiabetic effect only in SFDIO males: it decreased plasma FFA, leptin levels, tended to reduce plasma insulin level, and increased glucose tolerance. Sex dimorphism in response to FGF21 was reported earlier in obese Ay mice: FGF21 administration decreases hyperinsulinemia and hepatic lipid accumulation, increases muscle expression of genes involved in fatty acid oxidation and insulin signaling only in obese Ay males (Makarova et al., 2021b). It is possible that in SFDIO males, the FGF21-induced improvement in glucose tolerance was associated with a decrease in plasma concentration of leptin and fatty acids which are key risk factors for insulin resistance at obesity (Yang et al., 2018; Zhang Q. et al., 2019). Zhao and coauthors, using genetic approaches and a leptin neutralizing antibody, demonstrated that in obese mice, a partial reduction of plasma leptin levels restores hypothalamic leptin sensitivity and effectively enhances glucose tolerance and insulin sensitivity (Zhao et al., 2019).

In SFDIO males, the antidiabetic effect of FGF21 is most likely due to its action on the liver, which is a highly dimorphic target organ for FGF21 and plays a key role in the regulation of carbohydrate and lipid metabolism (Fisher et al., 2011; Torre et al., 2017). In our study, FGF21-treated SFDIO males exerted upregulated expression of hepatic genes (*Pclr*, *Gck* – both tendency*),* and proteins (GK and IR – tendency) involved in insulin signaling and genes related to fatty acid oxidation (*Ppargc1*) and lipogenesis *(Fasn*, *Acaca*), suggesting increased glucose and lipid turnover.

In SFDIO females, FGF21 administration did not increase the expression of any hepatic gene, which could be due to a ceiling effect. In vehicle-treated SFDIO mice, expression of many hepatic genes involved in lipid metabolism (*Ppargc1a*, *Cpt1*, *Fasn*, *Accb*) and expression of GK was higher in females than in males. It can be assumed that the increased expression of hepatic genes in vehicle-treated SFDIO females is at least partially due to the influence of estradiol, the expression of its receptors in the liver of female mice is significantly higher than in male mice (Torre et al., 2017). Estradiol administration to ovariectomized obese female mice increases expression of the insulin receptor gene (*Insr*) and uncoupling protein 2 gene (*Ucp2*) in the liver (Jakovleva et al., 2022). The relatively high initial level of gene expression could not be further enhanced by FGF21 administration (ceiling effect).

In addition, FGF21 and estradiol can interact with each other at the level of intracellular signal transduction pathways. According to the available data, estradiol and FGF21 have different receptors and the same signaling pathways (Fisher et al., 2011; Vrtačnik et al., 2014). Recently we demonstrated that in ovariectomized obese female mice, FGF21 reduced plasma insulin level, expression of *Pklr* and upregulated expression of *Irs2* in the liver. FGF21 did not affect these parameters if it was administered together with estradiol (Jakovleva et al., 2022). Thus, lack of the effect of FGF21 on hepatic gene expression in SFDIO females may be associated with crosstalk between estradiol and FGF21 in their actions.

In normal-weight male mice, FGF21 is known to increase protein intake, decrease sugar and alcohol intake, and have no effect on total energy intake (Talukdar et al., 2016a; Hill et al., 2020). Until now, it was unknown to what extent the effect of FGF21 on taste preferences is reproduced in mice with SFD-induced obesity. We demonstrated that FGF21 did not affect the total calorie intake in SFDIO males and females, and affected taste preferences only in SFDIO females: it increased the intake of calories with standard food, decreased it with cookies and did not alter it with lard. The standard chow contained the maximum amount of proteins, compared

with lard and cookies, and cookies contained a large amount of both carbohydrates and fats. Most likely, FGF21 reduced carbohydrate intake rather than lipid intake in SFDIO females, as it reduces sugar intake, but does not affect fat intake in experiments with a choice between two diets (Hill et al., 2020). Summing up, it can be assumed that FGF21 increased protein intake and reduced sweets intake in SFDIO female mice.

The selective influence of FGF21 on the feeding behavior of SFDIO females could be due to estrogen action. Recently we demonstrated that in ovariectomized SFDIO female mice, FGF21 upregulates the expression of hypothalamic genes encoding leptin receptor (*Lepr*) and its own co-receptor β-klotho (*Klb*) and estradiol increases its stimulating effect (Jakovleva et al., 2022). These data suggest that estradiol may enhance the stimulatory effect of FGF21 on hypothalamic sensitivity to leptin and FGF21. It was reported that hypothalamic circuitry connects with CNS circuitry modulating feeding behavior and leptin inhibits motivated behavior for palatable food (Figlewicz et al., 2003).

Summing up the data of the present work and our previous studies, one can say that different high-energy obesogenic diets create different metabolic backgrounds, which may affect the effectiveness of FGF21 treatment in obese female mice: the antidiabetic FGF21 action was manifested if they were fed a high fat diet (Makarova et al., 2021b) and was not manifested if they were fed a diet containing a sweet component along with fat. The mechanisms of diet-dependent action of FGF21 in females require special study, since they may affect the tactics of using FGF21 in individuals of different sexes.

Conclusion

Our study demonstrated that in SFDIO mice, FGF21 reduced body weight, did not alter total energy consumption, and activated orexigenic pathways of hypothalamus in mice of both sexes. However, sex dimorphism was found in the realization of the orexigenic FGF21 action at the transcriptional level in the hypothalamus. Hormonal, transcriptional and behavioral responses to FGF21 were also sex-specific. Only in SFDIO males, therapeutic effects of FGF21 administration – a decrease in fatty acid and leptin plasma levels, improvement of glucose tolerance, and upregulation of expression of genes involved in insulin signaling and lipid turnover in the liver – were observed. Only in SFDIO females, FGF21 induced preference of standard diet to sweet food.

References

Bazhan N., Jakovleva T., Feofanova N., Denisova E., Dubinina A., Sitnikova N., Makarova E. Sex differences in liver, adipose tissue, and muscle transcriptional response to fasting and refeeding in mice. *Cells*. 2019a;8(12):1529. DOI 10.3390/cells8121529.

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.* 2019b;19(1):28-36. DOI 10.3844/ojbsci.2019.28.36.

Berglund E.D., Li C.Y., Bina H.A., Lynes S.E., Michael M.D., Shanafelt A.B., Kharitonenkov A., Wasserman D.H. Fibroblast growth factor 21 controls glycemia via regulation of hepatic glucose flux and insulin sensitivity. *Endocrinology*. 2009;150(9):4084-4093. DOI 10.1210/en.2009-0221.

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.

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.

Dutchak P.A., Katafuchi T., Bookout A.L., Choi J.H., Yu R.T., Mangelsdorf D.J., Kliewer S.A. Fibroblast growth factor-21 regulates PPARγ activity and the antidiabetic actions of thiazolidinediones. *Cell*. 2012;148(3):556-567. DOI 10.1016/j.cell.2011.11.062.

Figlewicz D.P., Evans S.B., Murphy J., Hoen M., Baskin D.G. Expression of receptors for insulin and leptin in the ventral tegmental area/substantia nigra (VTA/SN) of the rat. *Brain Res*. 2003;964(1):107-115. DOI 10.1016/s0006-8993(02)04087-8.

Fisher F.M., Estall J.L., Adams A.C., Antonellis P.J., Bina H.A., Flier J.S., Kharitonenkov A., Spiegelman B.M., Maratos-Flier E. Integrated regulation of hepatic metabolism by fibroblast growth factor 21 (FGF21) *in vivo*. *Endocrinology*. 2011;152(8):2996-3004. DOI 10.1210/en.2011-0281.

Hale C., Chen M.M., Stanislaus S., Chinookoswong N., Hager T., Wang M., Véniant M.M., Xu J. Lack of overt FGF21 resistance in two mouse models of obesity and insulin resistance. *Endocrinology*. 2012;153(1):69-80. DOI 10.1210/en.2010-1262.

Hill C.M., Qualls-Creekmore E., Berthoud H.R., Soto P., Yu S., McDougal D.H., Münzberg H., Morrison C.D. FGF21 and the physiological regulation of macronutrient preference. *Endocrinology*. 2020;161(3):bqaa019. DOI 10.1210/endocr/bqaa019.

Iakovleva T.V., Kostina N.E., Makarova E.N., Bazhan N.M. Effect of gonadectomy and estradiol on the expression of insulin signaling cascade genes in female and male mice. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2020;24(4):427-434. DOI 10.18699/VJ20.635.

Jakovleva T.V., Kazantseva A.Y., Dubinina A.D., Balybina N.Y., Baranov K.O., Makarova E.N., Bazhan N.M. Estradiol-dependent and independent effects of FGF21 in obese female mice. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(2):159-168. DOI 10.18699/VJGB-22-20.

Keinicke H., Sun G., Mentzel C.M.J., Fredholm M., John L.M., Andersen B., Raun K., Kjaergaard M. FGF21 regulates hepatic metabolic pathways to improve steatosis and inflammation. *Endocr. Connect*. 2020;9(8):755-768. DOI 10.1530/EC-20-0152.

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):1627-1635. DOI 10.1172/JCI23606.

Kharitonenkov A., Adams A.C. Inventing new medicines: The FGF21 story. *Mol. Metab*. 2013;3(3):221-229. DOI 10.1016/j.molmet.2013.12.003.

Makarova E., Kazantseva A., Dubinina A., Denisova E., Jakovleva T., Balybina N., Bgatova N., Baranov K., Bazhan N. Fibroblast growth factor 21 (FGF21) administration sex-specifically affects blood insulin levels and liver steatosis in obese A^y mice. *Cells*. 2021a;10(12):3440. DOI 10.3390/cells10123440.

Makarova E., Kazantseva A., Dubinina A., Jakovleva T., Balybina N., Baranov K., Bazhan N. The same metabolic response to FGF21 administration in male and female obese mice is accompanied by sex-specific changes in adipose tissue gene expression. *Int. J. Mol. Sci*. 2021b;22(19):10561. DOI 10.3390/ijms221910561.

Makarova E.N., Yakovleva T.V., Balyibina N.Y., 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.

Martínez-Garza Ú., Torres-Oteros D., Yarritu-Gallego A., Marrero P.F., Haro D., Relat J. Fibroblast growth factor 21 and the adaptive response to nutritional challenges. *Int. J. Mol. Sci*. 2019;20(19):4692. DOI 10.3390/ijms20194692.

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.

Sampey B.P., Vanhoose A.M., Winfield H.M., Freemerman A.J., Muehlbauer M.J., Fueger P.T., Newgard C.B., Makowski L. Cafeteria diet is a robust model of human metabolic syndrome with liver and adipose inflammation: comparison to high-fat diet. *Obesity (Silver Spring)*. 2011;19(6):1109-1117. DOI 10.1038/oby.2011.18.

Sun H., Sherrier M., Li H. Skeletal muscle and bone – emerging targets of fibroblast growth factor-21. *Front. Physiol*. 2021;12:625287. DOI 10.3389/fphys.2021.625287.

Talukdar S., Kharitonenkov A. FGF19 and FGF21: In NASH we trust. *Mol. Metab*. 2021;46:101152. DOI 10.1016/j.molmet.2020.101152.

Talukdar S., Owen B.M., Song P., Hernandez G., Zhang Y., Zhou Y., Scott W.T., Paratala B., Turner T., Smith A., Bernardo B., Müller C.P., Tang H., Mangelsdorf D.J., Goodwin B., Kliewer S.A. FGF21 regulates sweet and alcohol preference. *Cell Metab*. 2016a;23(2):344-349. DOI 10.1016/j.cmet.2015.12.008.

Talukdar S., Zhou Y., Li D., Rossulek M., Dong J., Somayaji V., Weng Y., Clark R., Lanba A., Owen B.M., Brenner M.B., Trimmer J.K., Gropp K.E., Chabot J.R., Erion D.M., Rolph T.P., Goodwin B., Calle R.A. A long-acting FGF21 molecule, PF-05231023, decreases body weight and improves lipid profile in non-human primates and type 2 diabetic subjects. *Cell Metab*. 2016b;23(3):427-440. DOI 10.1016/j.cmet.2016.02.001.

Torre D., Lolli F., Ciana P., Maggi A. Sexual Dimorphism and Estrogen Action in Mouse Liver [published correction appears in Adv. Exp. Med. Biol. 2017;1043:E1]. *Adv. Exp. Med. Biol*. 2017;1043:141-151. DOI 10.1007/978-3-319-70178-3_8.

Vrtačnik P., Ostanek B., Mencej-Bedrač S., Marc J. The many faces of estrogen signaling. *Biochem. Med. (Zagreb)*. 2014;24(3):329-342. DOI 10.11613/BM.2014.035.

Xu J., Stanislaus S., Chinookoswong N., Lau Y.Y., Hager T., Patel J., Ge H., Weiszmann J., Lu S.C., Graham M., Busby J., Hecht R., Li Y.S., Li Y., Lindberg R., Véniant M.M. Acute glucose-lowering and insulin-sensitizing action of FGF21 in insulin-resistant mouse models–association with liver and adipose tissue effects. *Am. J. Physiol. Endocrinol. Metab*. 2009;297(5):E1105-E1114. DOI 10.1152/ajpendo.00348.2009.

Yang Q., Vijayakumar A., Kahn B.B. Metabolites as regulators of insulin sensitivity and metabolism. *Nat. Rev. Mol. Cell Biol*. 2018;19(10):654-672. DOI 10.1038/s41580-018-0044-8.

Zhang Q., Kong X., Yuan H., Guan H., Li Y., Niu Y. Mangiferin improved Palmitate-induced-insulin resistance by promoting free fatty acid metabolism in HepG2 and C2C12 cells via PPAR*α*: mangiferin improved insulin resistance. *J. Diabetes Res*. 2019;2019:2052675. DOI 10.1155/2019/2052675.

Zhang Y., Xie Y., Berglund E.D., Coate K.C., He T.T., Katafuchi T., Xiao G., Potthoff M.J., Wei W., Wan Y., Yu R.T., Evans R.M., Kliewer S.A., Mangelsdorf D.J. The starvation hormone, fibroblast growth factor-21, extends lifespan in mice. *Elife*. 2012;1:e00065. DOI 10.7554/eLife.00065.

Zhao S., Zhu Y., Schultz R.D., Li N., He Z., Zhang Z., Caron A., Zhu Q., Sun K., Xiong W., Deng H., Sun J., Deng Y., Kim M., Lee C.E., Gordillo R., Liu T., Odle A.K., Childs G.V., Zhang N., Kusminski C.M., Elmquist J.K., Williams K.W., An Z., Scherer P.E. Partial leptin reduction as an insulin sensitization and weight loss strategy. *Cell Metab*. 2019;30(4):706-719.e6. DOI 10.1016/j.cmet.2019.08.005.

ORCID ID

N.M. Bazhan orcid.org/0000-0002-7246-4758
T.V. Jakovleva orcid.org/0000-0001-7628-5856
A.Yu. Kazantseva orcid.org/0000-0003-4192-2020
N.E. Kostina orcid.org/0000-0001-6137-0492
P.E. Orlov orcid.org/0000-0001-6574-0286
N.Yu. Balybina orcid.org/0000-0001-7168-1438
K.O. Baranov orcid.org/0000-0002-4029-6518
E.N. Makarova orcid.org/0000-0002-6417-9893

Acknowledgements. This work was supported by the Russian Science Foundation, project 17-15-01036-П. The use of the equipment of the Department of Experimental Animal Genetic Resources, ICG, was supported by the Russian Ministry of Science and Higher Education, project RFMEFI62117X0015.

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

Received July 29, 2022. Revised September 30, 2022. Accepted September 30, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):342-348
DOI 10.18699/VJGB-23-41

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

Effects of central administration of the human Tau protein on the *Bdnf*, *Trkb*, *p75*, *Mapt*, *Bax* and *Bcl-2* genes expression in the mouse brain

A.S. Oreshko*, A.Ya. Rodnyy*, D.V. Bazovkina, V.S. Naumenko

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
naumenko2002@mail.ru

Abstract. Alzheimer's disease is the most common form of dementia, affecting millions of people worldwide. Despite intensive work by many researchers, the mechanisms underlying Alzheimer's disease development have not yet been elucidated. Recently, more studies have been directed to the investigation of the processes leading to the formation of neurofibrillary tangles consisting of hyperphosphorylated microtubule-associated Tau proteins. Pathological aggregation of this protein leads to the development of neurodegeneration associated with impaired neurogenesis and apoptosis. In the present study, the effects of central administration of aggregating human Tau protein on the expression of the *Bdnf*, *Ntrk2*, *Ngfr*, *Mapt*, *Bax* and *Bcl-2* genes in the brain of C57Bl/6J mice were explored. It was found that five days after administration of the protein into the fourth lateral ventricle, significant changes occurred in the expression of the genes involved in apoptosis and neurogenesis regulation, e. g., a notable decrease in the mRNA level of the gene encoding the most important neurotrophic factor BDNF (brain-derived neurotrophic factor) was observed in the frontal cortex which could play an important role in neurodegeneration caused by pathological Tau protein aggregation. Central administration of the Tau protein did not affect the expression of the *Ntrk2*, *Ngfr*, *Mapt*, *Bax* and *Bcl-2* genes in the frontal cortex and hippocampus. Concurrently, a significant decrease in the expression of the *Mapt* gene encoding endogenous mouse Tau protein was found in the cerebellum. However, no changes in the level or phosphorylation of the endogenous Tau protein were observed. Thus, central administration of aggregating human Tau protein decreases the expression of the *Bdnf* gene in the frontal cortex and the *Mapt* gene encoding endogenous mouse Tau protein in the cerebellum of C57Bl/6J mice.
Key words: Alzheimer's disease; Tau protein; *Bdnf*; neurogenesis; apoptosis; mice.

For citation: Oreshko A.S., Rodnyy A.Ya., Bazovkina D.V., Naumenko V.S. Effects of central administration of the human Tau protein on the *Bdnf*, *Trkb*, *p75*, *Mapt*, *Bax* and *Bcl-2* genes expression in the mouse brain. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):342-348. DOI 10.18699/VJGB-23-41

Эффекты центрального введения Tau-белка человека на экспрессию генов *Bdnf*, *Trkb*, *p75*, *Mapt*, *Bax* и *Bcl-2* в мозге мышей

А.С. Орешко*, А.Я. Родный*, Д.В. Базовкина, В.С. Науменко

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
naumenko2002@mail.ru

Аннотация. Болезнь Альцгеймера – это наиболее распространенная форма деменции, вызывающая прогрессирующую утрату когнитивных способностей и поражающая миллионы людей во всем мире. Несмотря на интенсивную работу множества исследовательских групп, механизмы, лежащие в основе развития болезни Альцгеймера, до сих пор не выяснены. В последнее время все больше усилий направлено на изучение механизмов, приводящих к формированию внутриклеточных нейрофибриллярных клубков, состоящих из гиперфосфорилированного Tau-белка, ассоциированного с микротрубочками. Патологическая агрегация Tau-белка, как известно, приводит к развитию нейродегенерации, связанной с нарушением нейрогенеза и апоптоза. В данном исследовании мы рассмотрели эффекты центрального введения агрегирующего Tau-белка человека на паттерны экспрессии генов *Bdnf*, *Ntrk2*, *Ngfr*, *Mapt*, *Bax* и *Bcl-2* в мозге мышей линии C57Bl/6J. Обнаружено, что через пять дней после введения Tau-белка человека в левый боковой желудочек мозга мыши происходят существенные изменения в паттернах экспрессии генов, принимающих участие в регуляции апоптоза и нейрогенеза. Так, было показано значительное снижение уровня мРНК гена *Bdnf*, кодирующего важнейший нейротрофический фактор мозга (brain-derived neurotrophic factor), во фронтальной коре мозга мышей экспериментальной группы, что может играть важную роль в нейродегенерации, вызываемой патоло-

* These authors contributed equally to this work.

гической агрегацией Tau-белка. В то же время центральное введение Tau-белка человека не повлияло на экспрессию генов *Ntrk2*, *Ngfr*, *Mapt*, *Bax* и *Bcl-2* во фронтальной коре и гиппокампе мышей. При этом в мозжечке было обнаружено существенное снижение экспрессии гена *Mapt*, кодирующего эндогенный Tau-белок мыши. Однако изменений в уровне белка и фосфорилировании эндогенного Tau-белка в исследованных структурах мозга не выявлено. Таким образом, центральное введение агрегирующего Tau-белка человека приводит к снижению экспрессии гена *Bdnf* во фронтальной коре и гена эндогенного Tau-белка (*Mapt*) в мозжечке мышей линии C57Bl/6J.

Ключевые слова: болезнь Альцгеймера; Tau-белок; *Bdnf*; нейрогенез; апоптоз; мыши.

Introduction

Alzheimer's disease (AD) is the most common cause of dementia with a prevalence of 24 million and an incidence of up to 5 million cases per year (Ferri et al., 2005). Russia is among the nine countries with the highest number of people suffering from AD (Prince et al., 2013; Collaborators, 2019). The annual death rate from AD and other forms of dementia in Russia in 2016 reached 35.7 per 100,000 inhabitants (https://www.who.int/healthinfo/global_burden_disease/estimates/en/), however, since many cases of AD remain unreported, this number is likely to be greatly underestimated.

AD is characterized by two main histopathological features: (1) extracellular amyloid plaques formed by insoluble aggregates of hydrophobic beta-amyloid peptides and (2) intracellular neurofibrillary tangles composed of hyperphosphorylated microtubule-associated Tau proteins. The accumulation of these two major types of aggregates leads to irreversible neurodegeneration that slowly spreads throughout the brain and causes progressive memory loss, cognitive decline, severe dementia and, finally, death (Breijyeh, Karaman, 2020). Although many generations of researchers have tried to unravel the mechanisms underlying this disease, they are still far from being fully understood. In recent years, the attention of an increasing number of scientists has been directed to studying the mechanisms leading to Tau pathology development.

Tau protein is a member of the microtubule-associated protein (MAP) family. Physiologically, this protein is involved in the formation and stabilization of microtubules in neurons and in the regulation of axonal transport and axon growth (Avila et al., 2004). The protein's main function is the regulation of tubulin polymerization, but it has also been shown to have DNA/RNA protection and signaling functions as well as to play a role in transcription regulation (Mandelkow E.M., Mandelkow E., 2012; Tapia-Rojas et al., 2019; Wegmann et al., 2021; Giovannini et al., 2022). Under pathological conditions, the accumulation of the protein's insoluble aggregates leads to the development of neurodegeneration, which is obviously leads to deteriorated neurogenesis and apoptosis, e. g., it has been shown that the level of Tau protein expression negatively correlates with the expression of BDNF (Wei et al., 2022) playing an important role in neuron development and support (Lu, Figurov, 1997; Benarroch, 2015; Gulyaeva, 2017). A similar correlation has been demonstrated for pathological Tau protein hyperphosphorylation (Yuan et al., 2022). Increased Tau protein expression also leads to a decrease in the BDNF level of blood plasma (Alvarez et al., 2022). Various actions that increase BDNF expression suppress the expression and pathological hyperphosphorylation of Tau protein (Li et al., 2022; Lin et al., 2022). Medications, including those of plant origin, that improve the performance of cognitive tasks in various models, reduce the expression of Tau protein and the proapoptotic BAX protein gene, which is accompanied by an increase in the expression of the antiapoptotic BCL-2 protein (Huang et al., 2022; Tu et al., 2022; Zhang et al., 2022), an increase in BDNF expression, as well as an increase in the expression and phosphorylation of the TrkB receptors mediating the positive effects of BDNF (Zhao et al., 2021; Liu et al., 2022; Nandini et al., 2022; Saikia et al., 2022; Wang et al., 2022). TrkB receptor activation has led to a decrease in Tau protein phosphorylation both *in vitro*, in cell culture, and in animal models (Chiang et al., 2021; Liao et al., 2021; Gonzalez et al., 2022).

The association between the nonspecific p75 receptor mediating the proapoptotic effects of the BDNF precursor (Guo et al., 2016; Hashimoto, 2016) and Tau pathology is less clear. Some studies have demonstrated that the adverse effects of aging and inflammation may be mediated at least partially by increased expression of the p75 receptor (Xie et al., 2021). At the same time, p75 receptor blockade suppresses proNGF (nerve growth factor precursor)-induced Tau protein phosphorylation (Shen et al., 2018). LM11A-31, a p75 receptor antagonist, also suppresses hyperphosphorylation and pathological aggregation of Tau protein in a mouse AD model (Yang et al., 2020).

However, it remains unclear what effect exerts introduction of aggregating human Tau proteins on the expression patterns of the genes involved in the processes of neurogenesis and apoptosis in the mouse brain. The aim of this study was to investigate the possibility of using standard C57Bl/6J mice after administration of aggregating human Tau protein into the left lateral ventricle as a model for studying Tau pathology mechanisms. In particular, we planned to evaluate the effects of the Tau protein administration on the expression patterns of the *Bdnf*, *Ntrk2* (encodes the TrkB receptor), *Ngfr* (encodes the p75 receptor), *Mapt* (encodes endogenous Tau), *Bax*, and *Bcl-2* genes in the mouse brain as well as on the level and phosphorylation of endogenous mouse Tau protein.

Materials and methods

Experimental animals. The investigation was carried out on C57Bl/6J inbred male mice of 10–12 weeks old weighing 27 ± 0.3 g at Center for Genetic Resources of Laboratory Animals of the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences (RFMEFI62119X0023).

In all experimental series, the animals were kept under the standard conditions of the vivarium that included artificial 14-hour lighting, 60 % humidity, temperature of 23 °C and a free access to balanced food and water. All the procedures involving experimental animals were performed in accordance with the international rules for the treatment of animals

(Directive 2010/63/EU) and Order of the Ministry of Health of the Russian Federation on Approval of the Rules of Good Laboratory Practice of 04/01/2016 No. 199n (registered on 08/15/2016 No. 43232).

Intraventricular administration of Tau protein. The human Tau protein was synthesized at Convergence Research Center for Diagnosis, Treatment and Care System of Dementia, Brain Science Institute, Korea Institute of Science and Technology (KIST) and kindly provided by the Director of the Institute, Dr. Yun Kyung Kim.

The protein was diluted in DMSO to a concentration of 2 mg/ml and then diluted with saline to a concentration of 0.2 μg/μl to be microinjected into the left lateral ventricle of the mice's brain (i.c.v.), AP: –0.5, L: –1.6 mm, DV: 2 mm (Slotnick, Leonard, 1975) under stereotaxic control (TSE, Germany). Before the injection, the mice had been narcotized for 20–30 sec with diethyl ether (Kondaurova et al., 2012). Mice in the control group received an injection of the solvent of the same composition. The volume of centrally injected fluids was 5 μl. Three days after the injection, the animals were placed in individual cages to remove group effects. After 46–48 hours, the mice were decapitated, their frontal cortex, hippocampus and cerebellum (as a control brain structure that is less involved in the implementation of the hyperphosphorylation and aggregation effects of Tau protein on cognition) were frozen in liquid nitrogen and stored at –80 °C prior total RNA isolation and western blotting.

RT-PCR. Total RNA was isolated using the TRIzol reagent (ThermoScientific, USA) and 1 μg of mRNA was used for synthesizing cDNA with random hexanucleotide primer. PCR was performed as in our previous studies (Naumenko et al., 2013a, b; Kondaurova et al., 2020). Real-time quantitative PCR was performed using the primers described in the Table. Gene expression was presented as the number of cDNA copies relative to 100 copies of *Polr2a* cDNA (Kulikov et al., 2005; Naumenko, Kulikov, 2006; Naumenko et al., 2008).

The primer sequences, annealing temperatures, and PCR product lengths

Gene	Nucleotide sequence	$T_{ann.}$, °C	PCR product length, bp
Bdnf	F5′-tagcaaaaagagaattggctg-3′ R5′-tttcaggtcatggatatgtcc-3′	59	255
Ntrk2	F5′-cattcactgtgagaggcaacc-3′ R5′-atcagggtgtagtctccgttatt-3′	63	175
Ngfr	F5′-acaacacccagcacccagga-3′ R5′-cacaaccacagcagccaaga-3′	62	171
Mapt	F5′-ccaagaaggtggcagtggtc-3′ R5′-agagccaatcttcgacctgac-3′	63	119
Bax	F5′-catctttgtggctggagtcctc-3′ R5′-aagtggacctgaggtttattggc-3′	64	216
Bcl-2	F5′-agagaggagaacgcaggtagtg-3′ R5′-cctcgcttcactgcctccttag-3′	64	187
Polr2a	F5′-tgtgacaactccatacaatgc-3′ R5′-ctctcttagtgaatttgcgtact-3′	61	194

Western blotting. The analysis was performed in the way it had been done in our previous works (Ilchibaeva et al., 2018; Popova et al., 2020). In brief, the total protein fraction was isolated from brain samples. The samples were then separated by 10 % SDS-PAGE and transferred to a nitrocellulose membrane. The membrane was then blocked with 5 % skimmed milk powder or 5 % BSA (for Phospho-Tau Thr18) for 1 hour and then incubated with primary antibodies to Tau proteins (5A6, 1:1000, DSHB, USA) and GAPDH (CAB932Hu01, 1:2500, Cloud-Clone Corp., USA) in 5 % milk powder with TBS-T or in 5 % FBS with TBS-T for Phospho-Tau Thr181 (AT270, 1:1000, Thermo Fisher Scientific, USA) for 16 hours at 4 °C. For protein detection, the membranes were incubated with horseradish peroxidase conjugated with secondary antibodies (anti-mouse Ig ab6728, 1:20000, Invitrogen, USA, Abcam, UK) in 5 % FBS with TBS-T for 1 hour at room temperature. Protein bands were visualized in a C-DiGit chemiluminescent blot scanner (LI-COR, USA) using the Clarity Western ECL substrate (Bio-Rad., USA). The bands were quantified using the Image Studio software (LICOR, USA). Target protein levels were normalized to that of GAPDH expression, which is constitutive of brain cells, and presented as a percentage of control animals. The number of analyzed samples was $n \geq 8$.

Statistical analysis. The results were presented as m ± SEM, where m is the mean and SEM is the standard error of the mean. The samples were compared using a one-way ANOVA. The differences were considered significant at $p < 0.05$. The normality of the variances was tested using the Kolmogorov–Smirnov and Shapiro–Wilk tests. Dixon's test was applied to identify and exclude extreme deviations from the analysis.

Results

Central administration of the human Tau protein resulted in a change in *Bdnf* gene expression. A significant decrease in the expression of this gene was found in the frontal cortex of the mice of the experimental group ($F_{1,13} = 7.2$, $p < 0.05$) (Fig. 1, *a*).

At the same time, no changes were found in the expression of genes encoding BDNF receptors ($F_{1,14} = 0.08$ for *Ntrk2* and $F_{1,14} = 1.9$, $p > 0.05$ for *Ngfr*, see Fig. 1, *b*, *c*). Also, the Tau protein had no effect on the expression of the genes encoding proapoptotic factor BAX ($F_{1,13} = 0.08$) and antiapoptotic factor BCL-2 ($F_{1,14} = 0.06$, see Fig. 1, *d*, *e*). Tau protein administration did not lead to significant changes in the expression of endogenous Tau protein both at the mRNA ($F_{1,14} = 0.2$) and protein levels ($F_{1,14} = 0.034$, see Fig. 1, *f*, *g*). The phosphorylation of endogenous Tau protein also did not change ($F_{1,14} = 0.273$ for the phospho-Tau level and $F_{1,14} = 0.393$ for the phospho-Tau/Tau ratio, see Fig. 1, *h*, *i*).

In hippocampus, Tau protein administration did not cause any significant changes in the expression pattern of the studied genes ($F_{1,13} = 1.2$, $p > 0.05$ for *Bdnf*; $F_{1,14} = 0.8$ for *Ntrk2*; $F_{1,13} = 1.0$, $p > 0.05$ for *Ngfr*; $F_{1,12} = 0.004$ for *Bcl-2*; $F_{1,10} = 0.1$ for *Bax*) (Fig. 2, *a*–*e*).

Also, no significant changes were found in the expression of the endogenous Tau protein both at the mRNA ($F_{1,14} = 1.3$, $p > 0.05$) and protein levels ($F_{1,14} = 0.508$, see Fig. 2, *f*, *g*). The phosphorylation of endogenous Tau protein did not change

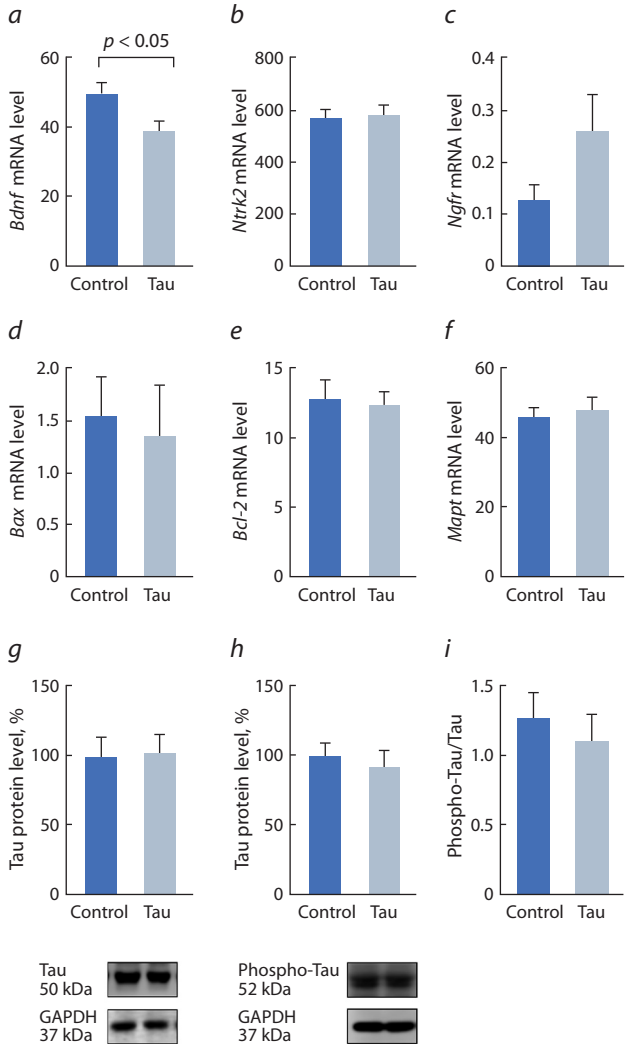


Fig. 1. Effect of central administration of human Tau protein on the expression of *Bdnf* (*a*), *Ntrk2* (*b*), *Ngfr* (*c*), *Bax* (*d*) and *Bcl-2* (*e*) genes, as well as on the *Mapt* gene mRNA level (*f*), Tau protein (*g*), phosphorylated Tau protein (*h*) and the ratio of phosphorylated Tau protein to Tau protein (*i*) in the frontal cortex of C57Bl/6J mice.

Here and in Fig 2 and 3: the gene expression is presented as the number of cDNA copies of the corresponding gene per 100 copies of *Polr2a* cDNA. The protein level is presented in relative units of the chemiluminescent signal and normalized to the level of GAPDH protein. $n \geq 8$.

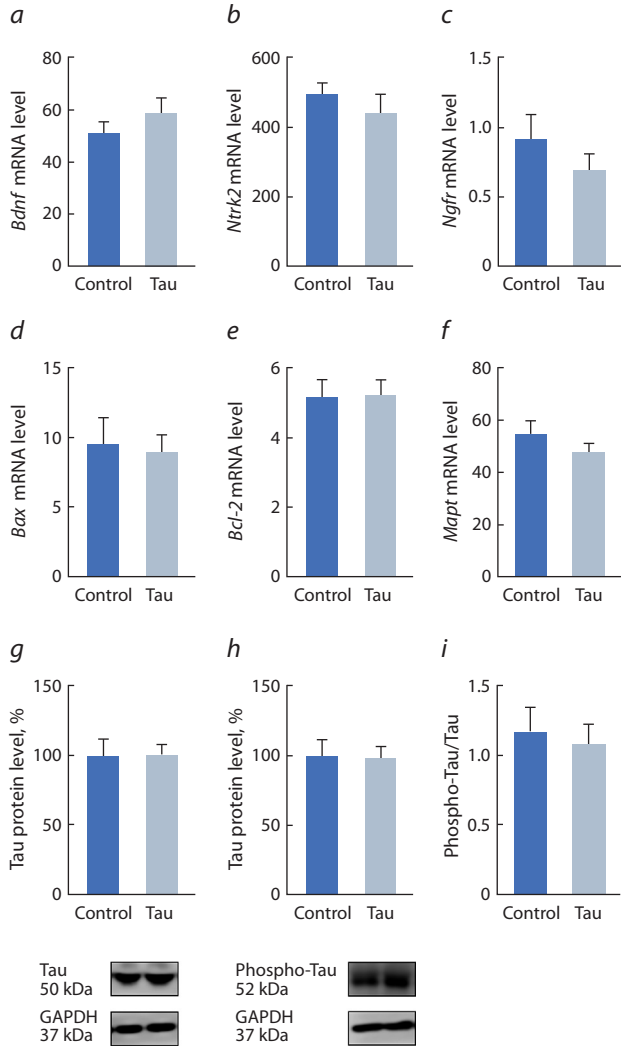


Fig. 2. Effect of central administration of human Tau protein on the expression of *Bdnf* (*a*), *Ntrk2* (*b*), *Ngfr* (*c*), *Bax* (*d*) and *Bcl-2* (*e*) genes, as well as on the *Mapt* gene mRNA level (*f*), Tau protein (*g*), phosphorylated Tau protein (*h*) and the ratio of phosphorylated Tau protein to Tau protein (*i*) in the hippocampus of C57Bl/6J mice.

as well ($F_{1,14} = 0.012$ for the phospho-Tau level; $F_{1,13} = 0.015$ for the phospho-Tau/Tau ratio, see Fig. 2, *h*, *i*).

In the cerebellum, the human Tau protein also did not affect the expression of *Bdnf* ($F_{1,14} = 0.3$), *Ntrk2* ($F_{1,14} = 0.5$), *Ngfr* ($F_{1,14} = 2.4$, $p > 0.05$), and *Bax* ($F_{1,11} = 1.4$, $p > 0.05$) genes (Fig. 3, *a*–*d*). At the same time, a trend towards a decrease in the anti-apoptotic *Bcl-2* gene was found ($F_{1,14} = 3.8753$, $p = 0.076$, see Fig. 3, *e*). Interestingly, Tau protein administration had a significant effect on the expression of the gene encoding endogenous Tau protein in the cerebellum ($F_{1,11} = 9.7$, $p > 0.01$, see Fig. 3, *f*). However, no significant changes were found in the level of endogenous Tau protein ($F_{1,13} = 0.043$) as well as in the level of its phosphorylation ($F_{1,13} = 0.107$ for the phospho-Tau level; $F_{1,13} = 0.011$ for the phospho-Tau/Tau ratio, see Fig. 3, *g*–*i*).

Discussion

AD is one of the most common causes of dementia that affects millions of people. It also leads to a large burden for the health and care systems. Despite intensive research worldwide, AD mechanisms remain unclear and only symptomatic treatment is available to date.

AD is characterized by the formation of two types of protein aggregates leading to the development of neurodegeneration. These are accumulations of extracellular amyloid plaques and intracellular neurofibrillary tangles, consisting of hyperphosphorylated microtubule-associated Tau proteins. Since long-term studies of the amyloid pathology have not brought the desired results, in recent years more and more research groups have directed their attention to investigating the mechanisms underlying the Tau pathology.

In this study, we investigated how central administration of human Tau protein in mice affects the expression patterns of the genes involved in the processes of neurogenesis and apop-

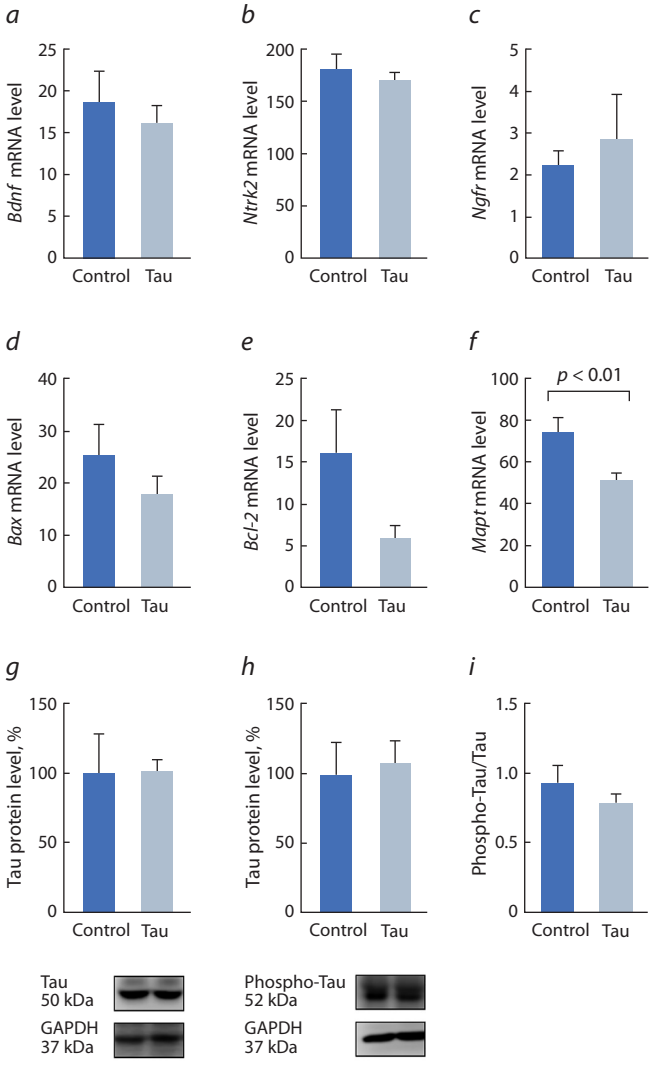


Fig. 3. Effect of central administration of human Tau protein on the expression of *Bdnf* (*a*), *Ntrk2* (*b*), *Ngfr* (*c*), *Bax* (*d*) и *Bcl-2* (*e*) genes, as well as on the *Mapt* gene mRNA level (*f*), Tau protein (*g*), phosphorylated Tau protein (*h*) and the ratio of phosphorylated Tau protein to Tau protein (*i*) in the cerebellum of C57Bl/6J mice.

tosis, as well as the level and phosphorylation of endogenous Tau protein. It was shown that Tau protein administration into the lateral ventricle led to a significant decrease in the expression of the gene encoding BDNF in the frontal cortex. Considering the critical role of this factor in neurons development and support (Lu, Figurov, 1997; Benarroch, 2015; Gulyaeva, 2017), it can be assumed that a decrease in *Bdnf* gene expression can lead to the development of neurodegeneration. Here, it has to be emphasized that the most pronounced neurodegenerative changes, as well as cell function changes in AD, have been observed precisely in the frontal cortex and hippocampus (Guevara et al., 2022; Lee et al., 2022). This is due, among other things, to the enhanced accumulation of Tau protein aggregates in these brain structures (Shimada et al., 2020) and the role of these brain structures in cognitive function regulation.

It is noteworthy that Tau protein administration also had its effect on the cerebellum that is not that much involved in cognitive processes. Our study demonstrated that the expression of the *Mapt* gene encoding the endogenous Tau protein was reduced in the cerebellum of the mice of the experimental group. In part, these results are consistent with the data on accumulation of Tau protein aggregates in this brain structure (Guevara et al., 2022). However, the detected changes in the Tau protein mRNA level in the cerebellum did not lead to significant changes in the level of endogenous Tau protein and its phosphorylation. Also, no changes were observed in the levels of mRNA, protein, and phosphorylation of the endogenous Tau protein in all structures studied. The central administration of Tau protein did not significantly affect the expression of other studied genes.

In general, our data agree with those on a negative correlation between Tau protein expression (Wei et al., 2022) and hyperphosphorylation (Yuan et al., 2022) with BDNF expression. However, despite the well-documented relationship between Tau protein expression and that of pro- and anti-apoptotic genes (Huang et al., 2022; Tu et al., 2022; Zhang et al., 2022), as well as the gene encoding the TrkB receptor (Zhao et al., 2021; Liu et al., 2022; Nandini et al., 2022; Saikia et al., 2022; Wang et al., 2022), the administration of Tau protein did not affect the expression patterns of these genes. An assumption can be made that the exogenous Tau protein introduced into the intercellular space penetrates poorly inside the neurons or is quickly catabolized there, not having the time to initiate a process of cells degeneration. Probably, for a more thorough investigation of the effects of Tau protein aggregation on the brain function, it is necessary to ensure endogenous expression of the pathologically phosphorylated Tau proteins in neurons.

Conclusion

The results obtained in this study indicate that central administration of human Tau protein to C57Bl/6J mice has a very weak effect on the expression of the investigated genes involved in neurogenesis and apoptosis. Nevertheless, Tau protein administration has led to a decrease in the expression of *Bdnf* gene in the frontal cortex and endogenous Tau protein-encoding gene in the cerebellum of C57Bl/6J mice without affecting the level and phosphorylation of endogenous Tau protein in the studied brain structures.

References

Alvarez X.A., Winston C.N., Barlow J.W., Sarsoza F.M., Alvarez I., Aleixandre M., Linares C., Garcia-Fantini M., Kastberger B., Winter S., Rissman R.A. Modulation of amyloid-β and Tau in Alzheimer's disease plasma neuronal-derived extracellular vesicles by Cerebrolysin® and donepezil. *J. Alzheimers Dis*. 2022;90(2):705-717. DOI 10.3233/JAD-220575.

Avila J., Lucas J.J., Perez M., Hernandez F. Role of tau protein in both physiological and pathological conditions. *Physiol. Rev*. 2004; 84(2):361-384. DOI 10.1152/physrev.00024.2003.

Benarroch E.E. Brain-derived neurotrophic factor: regulation, effects, and potential clinical relevance. *Neurology*. 2015;84(16):1693-1704. DOI 10.1212/WNL.0000000000001507.

Breijyeh Z., Karaman R. Comprehensive review on Alzheimer's disease: causes and treatment. *Molecules*. 2020;25(24):5789. DOI 10.3390/molecules25245789.

Chiang N.N., Lin T.H., Teng Y.S., Sun Y.C., Chang K.H., Lin C.Y., Hsieh-Li H.M., Su M.T., Chen C.M., Lee-Chen G.J. Flavones 7,8-DHF, quercetin, and apigenin against Tau toxicity *via* activation of TRKB signaling in ΔK280 Tau_{RD}-DsRed SH-SY5Y cells.

Front. Aging Neurosci. 2021;13:758895. DOI 10.3389/fnagi.2021.758895.

Collaborators G.B.D.N. Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2019;18(5):459-480. DOI 10.1016/S1474-4422(18)30499-X.

Ferri C.P., Prince M., Brayne C., Brodaty H., Fratiglioni L., Ganguli M., Hall K., Hasegawa K., Hendrie H., Huang Y., Jorm A., Mathers C., Menezes P.R., Rimmer E., Scazufca M., Alzheimer's Disease International. Global prevalence of dementia: a Delphi consensus study. *Lancet*. 2005;366(9503):2112-2117. DOI 10.1016/S0140-6736(05)67889-0.

Giovannini J., Smeralda W., Jouanne M., Sopkova-de Oliveira Santos J., Catto M., Voisin-Chiret A.S. Tau protein aggregation: key features to improve drug discovery screening. *Drug Discov. Today*. 2022;27(5):1284-1297. DOI 10.1016/j.drudis.2022.01.009.

Gonzalez S., McHugh T.L.M., Yang T., Syriani W., Massa S.M., Longo F.M., Simmons D.A. Small molecule modulation of TrkB and TrkC neurotrophin receptors prevents cholinergic neuron atrophy in an Alzheimer's disease mouse model at an advanced pathological stage. *Neurobiol. Dis*. 2022;162:105563. DOI 10.1016/j.nbd.2021.105563.

Guevara E.E., Hopkins W.D., Hof P.R., Ely J.J., Bradley B.J., Sherwood C.C. Epigenetic ageing of the prefrontal cortex and cerebellum in humans and chimpanzees. *Epigenetics*. 2022;17(12):1774-1785. DOI 10.1080/15592294.2022.2080993.

Gulyaeva N.V. Interplay between brain BDNF and glutamatergic systems: a brief state of the evidence and association with the pathogenesis of depression. *Biochemistry* (*Mosc*.). 2017;82(3):301-307. DOI 10.1134/S0006297917030087.

Guo J., Ji Y., Ding Y., Jiang W., Sun Y., Lu B., Nagappan G. BDNF pro-peptide regulates dendritic spines via caspase-3. *Cell Death Dis*. 2016;7:e2264. DOI 10.1038/cddis.2016.166.

Hashimoto K. Regulation of brain-derived neurotrophic factor (BDNF) and its precursor proBDNF in the brain by serotonin. *Eur. Arch. Psychiatry Clin. Neurosci*. 2016;266(3):195-197. DOI 10.1007/s00406-016-0682-9.

Huang J., Xu Z., Chen H., Lin Y., Wei J., Wang S., Yu H., Huang S., Zhang Y., Li C., Zhou X. Shen Qi Wan ameliorates learning and memory impairment induced by STZ in AD rats through PI3K/AKT pathway. *Brain Sci*. 2022;12(6):758. DOI 10.3390/brainsci12060758.

Ilchibaeva T.V., Tsybko A.S., Kozhemyakina R.V., Kondaurova E.M., Popova N.K., Naumenko V.S. Genetically defined fear-induced aggression: focus on BDNF and its receptors. *Behav. Brain Res*. 2018;343:102-110. DOI 10.1016/j.bbr.2018.01.034.

Kondaurova E.M., Naumenko V.S., Popova N.K. Effect of chronic activation of 5-HT_3 receptors on 5-HT_3, 5-HT_{1A} and 5-HT_{2A} receptors functional activity and expression of key genes of the brain serotonin system. *Neurosci. Lett*. 2012;522(1):52-56. DOI 10.1016/j.neulet.2012.06.015.

Kondaurova E.M., Rodnyy A.Y., Ilchibaeva T.V., Tsybko A.S., Eremin D.V., Antonov Y.V., Popova N.K., Naumenko V.S. Genetic background underlying 5-HT_{1A} receptor functioning affects the response to fluoxetine. *Int. J. Mol. Sci*. 2020;21(22):8784. DOI 10.3390/ijms21228784.

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

Lee Y., Miller M.R., Fernandez M.A., Berg E.L., Prada A.M., Ouyang Q., Schmidt M., Silverman J.L., Young-Pearse T.L., Morrow E.M. Early lysosome defects precede neurodegeneration with amyloid-β and tau aggregation in NHE6-null rat brain. *Brain*. 2022;145(9):3187-3202. DOI 10.1093/brain/awab467.

Li R., Ding X., Geetha T., Fadamiro M., St. Aubin C.R., Shim M., Al-Nakkash L., Broderick T.L., Babu J.R. Effects of genistein and exercise training on brain damage induced by a high-fat high-sucrose diet in female C57BL/6 mice. *Med. Cell. Longev*. 2022;2022:1560435. DOI 10.1155/2022/1560435.

Liao J., Chen C., Ahn E.H., Liu X., Li H., Edgington-Mitchell L.E., Lu Z., Ming S., Ye K. Targeting both BDNF/TrkB pathway and delta-secretase for treating Alzheimer's disease. *Neuropharmacology*. 2021;197:108737. DOI 10.1016/j.neuropharm.2021.108737.

Lin D.T., Kao N.J., Cross T.L., Lee W.J., Lin S.H. Effects of ketogenic diet on cognitive functions of mice fed high-fat-high-cholesterol diet. *J. Nutr. Biochem*. 2022;104:108974. DOI 10.1016/j.jnutbio.2022.108974.

Liu S., Fan M., Xu J.X., Yang L.J., Qi C.C., Xia Q.R., Ge J.F. Exosomes derived from bone-marrow mesenchymal stem cells alleviate cognitive decline in AD-like mice by improving BDNF-related neuropathology. *J. Neuroinflammation*. 2022;19(1):35. DOI 10.1186/s12974-022-02393-2.

Lu B., Figurov A. Role of neurotrophins in synapse development and plasticity. *Rev. Neurosci*. 1997;8(1):1-12. DOI 10.1515/revneuro.1997.8.1.1.

Mandelkow E.M., Mandelkow E. Biochemistry and cell biology of Tau protein in neurofibrillary degeneration. *Cold Spring Harb. Perspect. Med.* 2012;2(7):a006247. DOI 10.1101/cshperspect.a006247.

Nandini H.S., Krishna K.L., Apattira C. Combination of *Ocimum sanctum extract* and Levetiracetam ameliorates cognitive dysfunction and hippocampal architecture in rat model of Alzheimer's disease. *J. Chem. Neuroanat*. 2022;120:102069. DOI 10.1016/j.jchemneu.2021.102069.

Naumenko V.S., Bazovkina D.V., Semenova A.A., Tsybko A.S., Il'chibaeva T.V., Kondaurova E.M., Popova N.K. Effect of glial cell line-derived neurotrophic factor on behavior and key members of the brain serotonin system in mouse strains genetically predisposed to behavioral disorders. *J. Neurosci. Res*. 2013a;91(12):1628-1638. DOI 10.1002/jnr.23286.

Naumenko V.S., Kozhemyakina R.V., Plyusnina I.F., Kulikov A.V., Popova N.K. Serotonin 5-HT_{1A} receptor in infancy-onset aggression: comparison with genetically defined aggression in adult rats. *Behav. Brain Res*. 2013b;243:97-101. DOI 10.1016/j.bbr.2012.12.059.

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

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

Popova N.K., Kulikov A.V., Naumenko V.S. Spaceflight and brain plasticity: spaceflight effects on regional expression of neurotransmitter systems and neurotrophic factors encoding genes. *Neurosci. Biobehav. Rev*. 2020;119:396-405. DOI 10.1016/j.neubiorev.2020.10.010.

Prince M., Bryce R., Albanese E., Wimo A., Ribeiro W., Ferri C.P. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement*. 2013;9(1):63-75.e2. DOI 10.1016/j.jalz.2012.11.007.

Saikia B., Buragohain L., Barua C.C., Sarma J., Tamuli S.M., Kalita D.J., Barua A.G., Barua I.C., Elancheran R. Evaluation of antiamnesic effect of *Conyza bonariensis* in rats. *Indian J. Pharmacol*. 2022;54(2):102-109. DOI 10.4103/ijp.ijp_201_19.

Shen L.L., Manucat-Tan N.B., Gao S.H., Li W.W., Zeng F., Zhu C., Wang J., Bu X.L., Liu Y.H., Gao C.Y., Xu Z.Q., Bobrovskaya L., Lei P., Yu J.T., Song W., Zhou H.D., Yao X.Q., Zhou X.F., Wang Y.J. The ProNGF/p75NTR pathway induces tau pathology and is a therapeutic target for FTLD-tau. *Mol. Psychiatry*. 2018;23(8):1813-1824. DOI 10.1038/s41380-018-0071-z.

Shimada H., Minatani S., Takeuchi J., Takeda A., Kawabe J., Wada Y., Mawatari A., Watanabe Y., Shimada H., Higuchi M., Suhara T., Tomiyama T., Itoh Y. Heavy tau burden with subtle amyloid β accumulation in the cerebral cortex and cerebellum in a case of familial Alzheimer's disease with APP Osaka mutation. *Int. J. Mol. Sci*. 2020;21(12):4443. DOI 10.3390/ijms21124443.

Slotnick B.M., Leonard C.M. A Stereotaxic Atlas of the Albino Mouse Forebrain. Rockville, Maryland: U.S. Dept. of Health, Education, and Welfare, 1975.

Tapia-Rojas C., Cabezas-Opazo F., Deaton C.A., Vergara E.H., Johnson G.V.W., Quintanilla R.A. It's all about tau. *Prog. Neurobiol*. 2019;175:54-76. DOI 10.1016/j.pneurobio.2018.12.005.

Tu Y., Chen Q., Guo W., Xiang P., Huang H., Fei H., Chen L., Yang Y., Peng Z., Gu C., Tan X., Liu X., Lu Y., Chen R., Wang H., Luo Y., Yang J. MiR-702-5p ameliorates diabetic encephalopathy in db/db mice by regulating 12/15-LOX. *Exp. Neurol*. 2022;358:114212. DOI 10.1016/j.expneurol.2022.114212.

Wang Y., Zhang J.J., Hou J.G., Li X., Liu W., Zhang J.T., Zheng S.W., Su F.Y., Li W. Protective effect of ginsenosides from stems and leaves of *Panax ginseng* against scopolamine-induced memory damage via multiple molecular mechanisms. *Am. J. Chin. Med*. 2022; 50(4):1113-1131. DOI 10.1142/S0192415X22500458.

Wegmann S., Biernat J., Mandelkow E. A current view on Tau protein phosphorylation in Alzheimer's disease. *Curr. Opin. Neurobiol*. 2021;69:131-138. DOI 10.1016/j.conb.2021.03.003.

Wei Y.D., Chen X.X., Yang L.J., Gao X.R., Xia Q.R., Qi C.C., Ge J.F. Resveratrol ameliorates learning and memory impairments induced by bilateral hippocampal injection of streptozotocin in mice. *Neurochem. Int*. 2022;159:105385. DOI 10.1016/j.neuint.2022.105385.

Xie Y., Seawell J., Boesch E., Allen L., Suchy A., Longo F.M., Meeker R.B. Small molecule modulation of the p75 neurotrophin receptor suppresses age- and genotype-associated neurodegeneration in HIV gp120 transgenic mice. *Exp. Neurol*. 2021;335:113489. DOI 10.1016/j.expneurol.2020.113489.

Yang T., Tran K.C., Zeng A.Y., Massa S.M., Longo F.M. Small molecule modulation of the p75 neurotrophin receptor inhibits multiple amyloid beta-induced tau pathologies. *Sci. Rep*. 2020;10(1):20322. DOI 10.1038/s41598-020-77210-y.

Yuan M., Wang Y., Wen J., Jing F., Zou Q., Pu Y., Pan T., Cai Z. Dietary salt disrupts tricarboxylic acid cycle and induces tau hyperphosphorylation and synapse dysfunction during aging. *Aging Dis*. 2022;13(5):1532-1545. DOI 10.14336/AD.2022.0220.

Zhang H., Lu F., Liu P., Qiu Z., Li J., Wang X., Xu H., Zhao Y., Li X., Wang H., Lu D., Qi R. A direct interaction between RhoGDIα/Tau alleviates hyperphosphorylation of Tau in Alzheimer's disease and vascular dementia. *J. Neuroimmune Pharmacol*. 2022. DOI 10.1007/s11481-021-10049-w.

Zhao H.L., Cui S.Y., Qin Y., Liu Y.T., Cui X.Y., Hu X., Kurban N., Li M.Y., Li Z.H., Xu J., Zhang Y.H. Prophylactic effects of sporoderm-removed *Ganoderma lucidum* spores in a rat model of streptozotocin-induced sporadic Alzheimer's disease. *J. Ethnopharmacol*. 2021;269:113725. DOI 10.1016/j.jep.2020.113725.

ORCID ID

A.Ya. Rodnyy orcid.org/0000-0001-5739-4176
D.V. Bazovkina orcid.org/0000-0001-5786-7965
V.S. Naumenko orcid.org/0000-0001-7196-4729

Acknowledgements. The study was supported by the Russian Science Foundation (Grant No. 22-15-00011). The cost of animals housing was partly covered by basic research project FWNR-2022-0010.

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

Received November 16, 2022. Revised January 16, 2023. Accepted January 17, 2023.

Вавиловский журнал генетики и селекции. 2023;27(4):349-356
DOI 10.18699/VJGB-23-42

Comparison of the evolutionary patterns of DNA repeats in ancient and young invertebrate species flocks of Lake Baikal

Wang Yuxiang[1], T.E. Peretolchina[1], E.V. Romanova[1], D.Y. Sherbakov[1, 2, 3] ✉

[1] Limnological institute of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Irkutsk State University, Irkutsk, Russia
✉ sherb@lin.irk.ru

Abstract. DNA repeat composition of low coverage (0.1–0.5) genomic libraries of four amphipods species endemic to Lake Baikal (East Siberia) and four endemic gastropod species of the fam. Baicaliidae have been compared to each other. In order to do so, a neighbor joining tree was inferred for each quartet of species (amphipods and mollusks) based on the ratio of repeat classes shared in each pair of species. The topology of this tree was compared to the phylogenies inferred for the same species from the concatenated protein-coding mitochondrial nucleotide sequences. In all species analyzed, the fraction of DNA repeats involved circa half of the genome. In relatively more ancient amphipods (most recent common ancestor, MRCA, existed approximately sixty millions years ago), the most abundant were species-specific repeats, while in much younger Baicaliidae (MRCA equal to ca. three millions years) most of the DNA repeats were shared among all four species. If the presence/absence of a repeat is regarded as a separate independent trait, and the ratio of shared to total numbers of repeats in a species pair is used as the measure of distance, the topology of the NJ tree is the same as the quartet phylogeny inferred for the mitogenomes protein coding nucleotide sequences. Meanwhile, in each group of species, a substantial number of repeats were detected pointing to the possibility of non-neutral evolution or a horizontal transfer between species occupying the same biotope. These repeats were shared by non-sister groups while being absent in the sister genomes. On the other hand, in such cases some traits of ecological significance were also shared.
Key words: DNA repeats; Lake Baikal; phylogeny; Baicaliidae; amphipods; evolution of repeats; repeatome.

For citation: Wang Yuxiang, Peretolchina T.E., Romanova E.V., Sherbakov D.Y. Comparison of the evolutionary patterns of DNA repeats in ancient and young invertebrate species flocks of Lake Baikal. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):349-356. DOI 10.18699/VJGB-23-42

Сравнение эволюционных паттернов ДНК-повторов у представителей древних и молодых букетов видов из озера Байкал

Ван Юйсян[1], Т.Е. Перетолчина[1], Е.В. Романова[1], Д.Ю. Щербаков[1, 2, 3] ✉

[1] Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Иркутский государственный университет, Иркутск, Россия
✉ sherb@lin.irk.ru

Аннотация. Исследованы ДНК повторы, присутствующие в геномных библиотеках с низким покрытием (0.1–0.5) четырех видов амфипод, эндемичных для озера Байкал (Восточная Сибирь), и четырех эндемичных видов брюхоногих моллюсков семейства Baicaliidae. Для этого были построены деревья методом объединения ближайших соседей для каждого квартета видов (амфиподы и моллюски) на основе соотношения повторяющихся классов, общих для каждой пары видов. Топология этих деревьев была сопоставлена с филогениями, полученными для тех же видов на основе сцепленных белок-кодирующих митохондриальных нуклеотидных последовательностей. У всех проанализированных видов в долю повторов ДНК вовлечено около половины генома. У относительно более древних амфипод (самый последний общий предок, MRCA, существовал приблизительно шестьдесят миллионов лет назад) наиболее распространенными были видоспецифичные повторы, тогда как у гораздо более молодых байкалиид (MRCA приблизительно равен трем миллионам лет) большинство повторов ДНК были общими для всех четырех видов. Если наличие/отсутствие повтора рассматривать как отдельный независимый признак, а отношение общего числа повторов в паре видов использовать в качестве меры расстояния, топология дерева NJ такая же, как и филогения квартета, выведенная для белков митогеномов, кодирующих нуклеотидные последовательности. Между тем в каждой группе видов было обнаружено значительное количество повторов, указывающих на возможность ненейтральной эволюции или горизонтального переноса между видами, занимающими один и тот же биотоп. Эти повторы были общими для неродственных групп, но отсутствовали в сестринских геномах. С другой стороны, в таких случаях некоторые черты, имеющие экологическое значение, также были общими.
Ключевые слова: повторы ДНК; озеро Байкал; филогения; Baicaliidae; амфиподы; эволюция повторов; репитом.

Introduction

In Metazoa, approximately half of all genomic DNA is made up of repeated DNA sequences, which are otherwise called "non-genic DNA" (Cavalier-Smith, Beaton, 1999; Bird et al., 2006) or repeatome (Titievsky et al., 2021). The already known functions of this fraction of the genome are very diverse. Most of it is satellite DNA (Biscotti et al., 2015; Silva et al., 2019; Thakur et al., 2021). A significant proportion of DNA repeats account for mobile elements belonging to different classes. There is evidence that highly repeated mobile elements may play a certain role in the regulation of genetic activity (see for example (Rocha et al., 2022)), their distribution must also be taken into account in the epigenetic analysis (Lerat et al., 2019). It is important to note that evidence is accumulating about the important role that repeated mobile elements may play in horizontal gene transfer between phylogenetically distant species (Ahmad et al., 2021; Athanasouli, Rödelsperger, 2022; Kejnovsky, Jedlicka, 2022). Dodsworth et al. (2015) have shown that a set of repeated elements contains a significant phylogenetic signal. They also noted the presence of a repeat fraction, which was inconsistent with the phylogenies inferred from individual nucleotide sequences and from the repeats, but treated this fraction more like an obstacle rather than an interesting phenomenon. It has been shown that horizontal gene transfer is a common mechanism of transmission of traits involved in adaptation processes in bacteria (Lee et al., 2022) and fungi (Steensels et al., 2021). Recently, there has been more and more evidence that similar mechanisms are likely to be involved in the adaptive evolution of Metazoa (Boto, 2014; Chen et al., 2017; Ahmad et al., 2021; Li et al., 2022). Since then their work though the main focus of studies of the repeated DNA shifted mostly towards their potential structural role and was performed mostly of plant models (see for example (Titievsky et al., 2021)).

Here we apply the repeats analysis to the two species flocks of Baikalian invertebrates. Genetic studies of invertebrates from Lake Baikal allow to unravel many problems of their evolutionary history (Romanova et al., 2016; Peretolchina et al., 2020), mechanisms of speciation (Naumenko et al., 2017; Gurkov et al., 2019; Drozdova et al., 2022) and adaptation (Lipaeva et al., 2021), diversity and conservation (Butina et al., 2019; Yakhnenko, Itskovich, 2020). We use the features of the evolution of two species flocks of endemic Baikal invertebrates – amphipods (Bazikalova, 1945; Kamaltynov, 1999; Takhteev, 2019) and gastropods of the family Baicaliidae (Sitnikova et al., 2001; Hausdorf et al., 2003; Peretolchina et al., 2020) to study the evolution of the maximum diversity of repeats in their genomes. These two groups are attractive models for this kind of research for the following reasons:

1. Both groups of organisms have been well and comprehensively studied (see (Kozhov, 1963)).
2. Both groups evolved within Baikal, therefore all possible genome transformations were minimally, if at all, dependent on the introduction of genetic information from outside the ecosystem, and speciation processes occurred mainly by sympatric mechanisms.
3. The evolutionary histories of amphipods and Baicaliidae in Baikal are fundamentally different: if the former is represented by at least two branches that independently penetrated Baikal, the common ancestor of which existed at least 60 million years ago, then the maximum age of the common ancestor of Baikal species is at least 3 million years (Sherbakov, 1999; Mats et al., 2011).

Thus, the above-mentioned properties of evolutionary histories allow us to conduct a comparative analysis of sets of DNA repeats in two species-rich groups of invertebrates and assess the potential benefits of such a comparison for a deeper understanding of the evolutionary mechanisms that have shaped their modern diversity.

Materials and methods

In this work, the genomic libraries of gastropods were used: *Baicalia turriformis*, *Maackia herderiana*, *Korotnewia korotnewi* and *Godlewskia godlewskia*, the collection of samples and genome-wide sequencing of which is described in Peretolchina et al. (2020). Obtaining the genomic libraries of amphipods *Acanthogammarus victorii*, *Brachyuropus grewingkii*, *Garjajewia cabanisi* and *Macrohectopus branickii* is described in (Romanova et al., 2016).

Random sets of reads were prepared from the source libraries using Seqtk-1.3 (r106) (Shen et al., 2016) on the Galaxy (Jalili et al., 2020) platform. The size of a library was set to 5×10^5 reads. The search for repeating genetic elements was performed using the RepeatExplorer (Novák et al., 2013) pipeline implemented on the Galaxy platform.

Quality control and library filtering were performed using the standard Galaxy FastQC (de Sena Brandine, Smith, 2019) tool. Cluster analysis requires files containing sequences of reads in FASTA format as input data.

The search for repeated sequences was performed using RepeatExplorer2 clustering.

Launch Parameters:

Paired-end reads True, Read sampling false, Sample size 0,
Select taxon and protein domain database version (REXdb) Viridiplantae version 3.0,
Advanced options false,
Select queue basic_fast_queue,
Modify parameters (optional) –l select=1:ncpus=10:mem=32gb:scratch_local=50gb –l walltime=48:00:00 –q elixirre@pbs.elixir-czech.cz –v TAREAN_MAX_MEM=4000000,TAREAN_CPU=4

The search for repeats was limited to those that occur more often than 0.01 % of the input reads. In addition to satellite repeats, the output data contains LTR-retrotrans-posons, 45S, 5S rDNA and all other repeats, the number of which exceeds the threshold value.

Comparative analysis of the composition of repeats was performed using a set of original scripts in Python 3.10 and Biopython ver. 1.79. nblast (Costa et al., 2016) was used to compare the nucleotide sequences of DNA repeat contigs from different species.

Results

Each of the groups of organisms is represented in this paper by four species selected in such a way that they cover the maximum range of evolutionary distances within their branch. The phylogenetic relationships of four gastropod species (*Baicalia turriformis*, *Maackia herderiana*, *Korotnewia korotnewi* and *Godlewskia godlewskia*) are shown in Fig. 1, *b*. The lifetime of their common ancestor does not exceed 3 million years

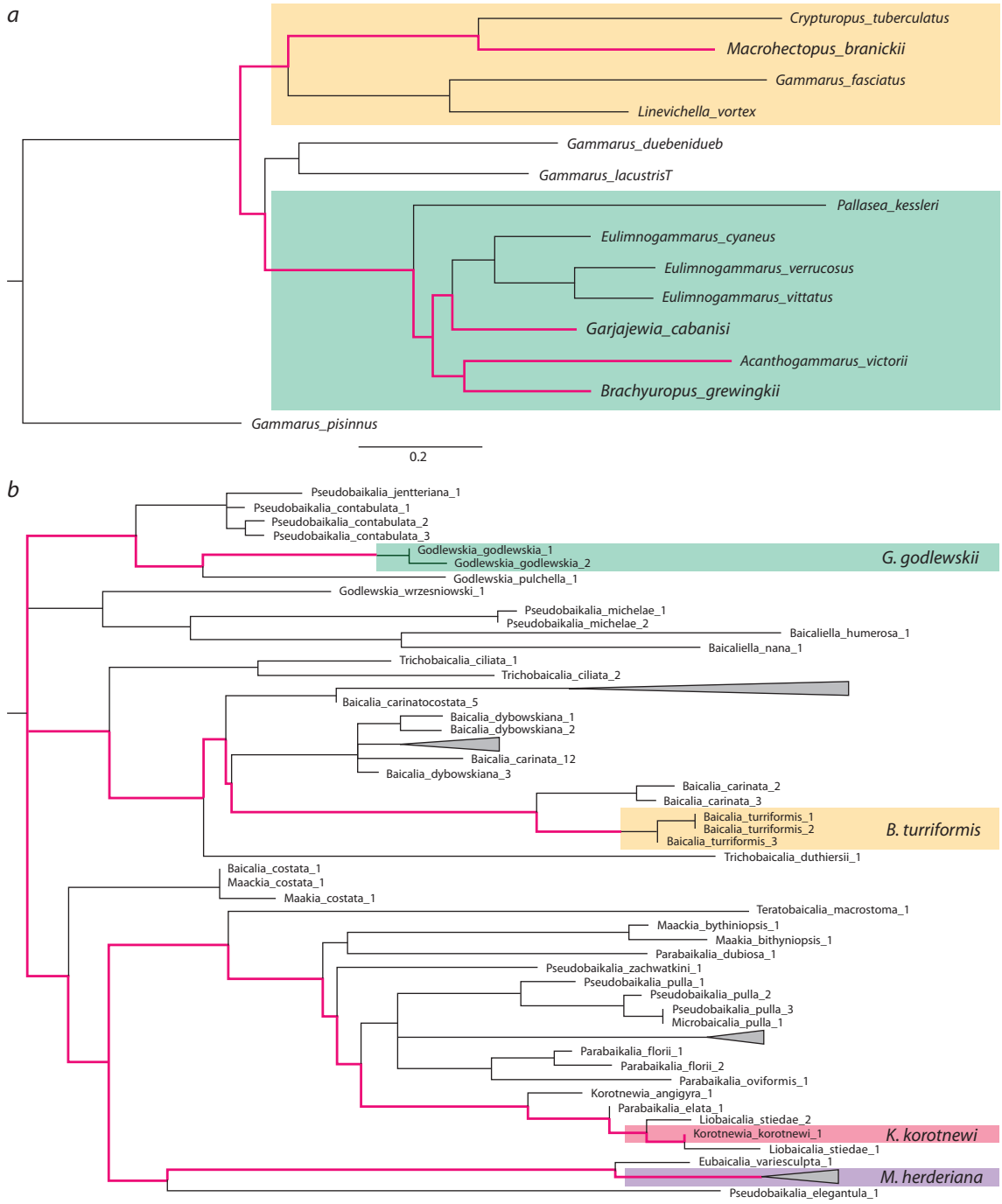


Fig. 1. *a*, A phylogenetic tree of Baikal amphipods and some representatives of the genus *Gammarus*, obtained on the basis of a comparison of concatenated nucleotide sequences on which the species studied in this work are isolated. Two branches represented in Baikal are highlighted in color, as well as phylogenetic relationships between *Acanthogammarus victorii*, *Brachyuropus grewingkii*, *Garjajewia cabanisi* and *Macrohectopus branickii*.
b, A phylogenetic tree of Baicaliidae, rooted at the midpoint, built on the basis of the analysis of the sequences of the Folmer fragment. The phylogenetic relationships between *Baicalia turriformis*, *Maackia herderiana*, *Korotnewia korotnewi* and *Godlewskia godlewskia* are highlighted in color.

(Zubakov et al., 1997; Sherbakov, 1999), and the connections between the selected four species pass through the root of the tree if the tree is rooted at the midpoint. It should be noted that the selected species differ dramatically in their important ecological characteristics and distribution.

Amphipods in Baikal belong to at least two large branches, both within the genus *Gammarus* (Sherbakov, 1999; Hou, Sket, 2016; Romanova et al., 2016; Naumenko et al., 2017). Of the species selected for the study, only *M. branickii* – a representative of the monotypic family – belongs to the branch 'Micruropus', the rest belong to the branch 'Acanthogammarus' – the most diverse in both species and ecology. *B. grewingkii* and *G. cabanisi* are abyssal species, *A. victorii* lives at shallow and medium depths, and *M. branickii* is a unique

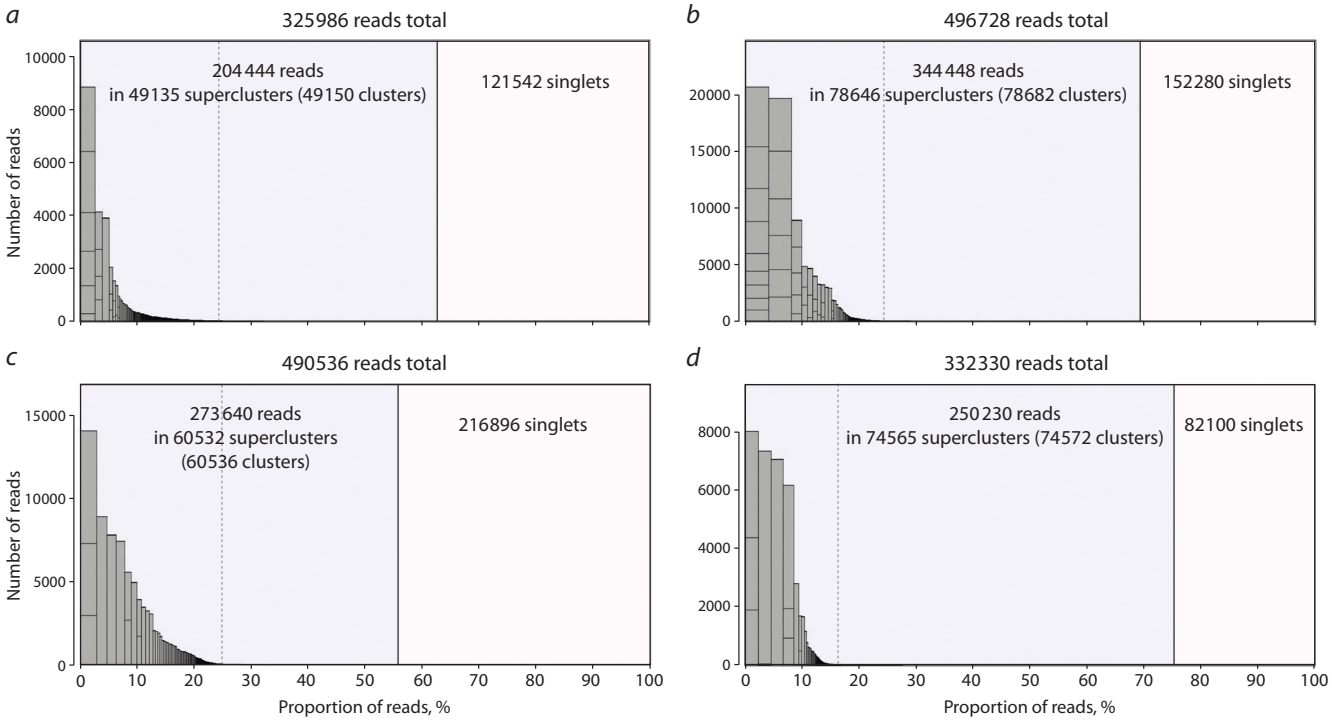


Fig. 2. Representation of the repeats among the species of amphipods sorted by their abundance: *a*, *Acanthogammarus victorii*; *b*, *Brachyuropus grewingkii*; *c*, *Garjajewia cabanisi*, and *d*, *Macrohectopus branickii*.

pelagic species distributed throughout the water column of the lake, including the maximum depths (Bazikalova, 1945) and as part of the quartet of species considered in this work is a distant outer group (see Fig. 1, *a*).

Libraries of repeated contigs were constructed from genome-wide libraries of four species of amphipods and four species of gastropods of the Baikal endemic family Baicaliidae, the production of which is described in (Romanova et al., 2016) and (Peretolchina et al., 2020), respectively. 0.5×10^6 reads were randomly selected from each library and without return, resulting in depleted libraries with a coverage degree of less than 0.5, as a result of which the representation of unique sequences in them turned out to be very low. These subsets of genomic libraries were used to search and annotate DNA repeats using repeatexplorer (Novák et al., 2013). In all cases, the repetitions included approximately 50 % of the reads, which accounted for from 5×10^3 to 10^4 of unique contigs (Fig. 2).

The distributions of contigs by representation in genomes were also approximately the same in all cases; however, if *A. victorii* and *G. cabanisi* had a single dominant repeat (refers to simple DNA repeats, SSR) (see Fig. 2, *a*, *c*), then *B. grewingkii* and *M. branickii* had several dominant repeats (see Fig. 2, *b*, *d*).

The comparison of the compositions of repeats in the species within each of the groups was carried out by concatenating the output files – lists of contigs resulting from a search through genomic libraries. To distinguish between contigs belonging to different species, prefix indexing was used, specific to each of the species. After converting a copy of the concatenated list into a library in blast format, we used nblast to search in the "all against all" mode. At the same time, deletions/insertions were allowed and the similarity threshold of sequences was set to 80 %.

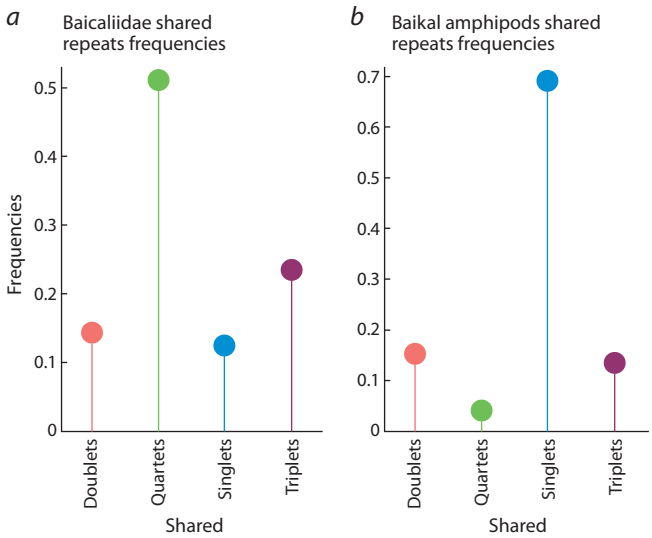


Fig. 3. The proportions of repeats common to several species and found as a result of the blast search for gastropods Baicaliidae and amphipods.

For each of the studied groups, libraries of repeat contigs were concatenated after adding species-specific tags to sequence names, then groups consisting of at least five sequences were selected and a nblast search was performed "all against all". The search conditions allowed 20 % differences and indels. The analysis of intragroup distributions of repeats revealed significant differences between amphipods and molluscs (Fig. 3, see the Table). In a much younger group of Bai-

Patterns of repeats shared in amphipods and gastropods (Baicaliidae)

Species	Total number of contigs	Unique contigs (ratio)	Fully shared contigs (ratio)
	Amphipods		
Brachyuropus grewingkii	21070.0	0.233	0.135
Acanthogammarus victorii	30695.0	0.315	0.093
Garjajewia cabanisi	58335.0	0.605	0.049
Macrohectopus branickii	10674.0	0.455	0.267
	Baicaliidae		
Baicalia turriformis	58491.0	0.036	0.626
Maackia herderiana	57525.0	0.044	0.637
Korotnewia korotnewi	51841.0	0.040	0.707
Godlewskia godlewskia	57872.0	0.037	0.633

kal, the repeats that occur in all four species turned out to be the most represented. Conversely, a relatively small proportion is accounted for by species-specific sequences (see Fig. 3, *a*). The representation of repetitions common to several species also turned out to be very similar in different Baikal species.

In amphipods, on the contrary, most of the repeats are unique (species-specific), and there are very few common ones for all four species (see Fig. 3, *a*). Repeats common to two, three and four species are also not equally represented in different genomes of amphipods (see the Table and Fig. 3, *b*). Interestingly, the largest proportion of common repeats ('quartets') was found in the genome of *M. branickii*, which is a very remote external group in relation to the other three species and, unlike the rest of the Baikal amphipods, lives in the pelagic zone of the lake.

In general, it should be noted that all possible patterns of repeat propagation are present in the genomes of both groups of species: there are both those present in only two species in all possible combinations, and all variants of absence in only one of the species (see Fig. 3).

A comparison of the distribution of repeats belonging to different classes according to their distribution in Figure 4 also does not reveal any interspecific variation in Baikal and rather significant differences between amphipod species. However, in both, all possible combinations of the two species are detected, which have repeated elements in common only for them (up to the sensitivity of detection and identification conditions).

Sets of common repeats were used to cluster species. To do this, as a measure of the distance between species, we used

$$d_{ij} = 1 - \frac{N_{\text{shared}}}{N_i + N_j},$$

where d_{ij} is the distance between species (genomes), i and j are species or genomes numbers, N_{shared}, N_i and N_j are the numbers of repeat types in the respective species. Note that the abundances of repeats of each type are not taken into account, but the denominator N_{common} involves all types of repeats found in the pulled repeats library of the species compared. These distances were used to construct the distance matrix, and it, in turn, was used to build a tree by combining the nearest neighbors (Saitou, Nei, 1987).

For the same species and both groups, maximum likelihood trees were inferred based on a comparison of concatenated protein-coding nucleotide sequences of mitochondrial genomes. The topologies of the trees coincided, but the ratio of nucleotide distances and distances calculated by common repetition is not linear (data not shown).

A more detailed examination of pairs of species with common repeats (Fig. 5) shows that a noticeable, albeit relatively small number of repeats is shared by species that are not sister species and thus are not consistent with the phylogeny. This property is present in both amphipods and gastropods.

Discussion

The libraries of the NGS reads of four species of Baikal endemic gastropods (*Baicalia turriformis*, *Maackia herderiana*, *Korotnewia korotnewi* and *Godlewskia godlewskia*) belonging to the Baikal endemic Baicaliidae, and four Baikal amphipods (*Acanthogammarus victorii*, *Brachyuropus grewingkii*, *Garjajewia cabanisi* and *Macrohectopus branickii*) were used to *de novo* search for repeated DNA elements using the repeatexplorer algorithm. All taxa whose genomic libraries are analyzed in this work, despite various evolutionary histories, evolved within the limits of the reservoir that continuously existed on the site of modern Baikal.

The gastropods of Baicaliidae are a relatively young group, the time of the most recent common ancestors (t_{MRCA}) of modern species is no more than 2.5 million years old. They are found at depths of no more than 100 m on a variety of soil types (Zubakov et al., 1997; Sitnikova, 2006). The amphipods in Baikal are represented by at least two large branches that diverged no earlier than about 60 million years ago (Sherbakov, 1999; Mats et al., 2011; Naumenko et al., 2017). The variety of ecological niches occupied by them is exceptionally large, they are found at all depths.

By their distribution between species, all theoretically possible combinations of repeat classes were found. They ranged from species-specific ones to those found in all the genomes studied. This circumstance made it possible to use the distribution of repeats between genomes as a tool for clustering the corresponding species and comparing the topology of the obtained quartets with the results of clustering of the same species based on a comparison of the nucleotide sequences of concatenated protein-coding fragments of mitochondrial genomes. The topologies coincided, but the ratio of the lengths

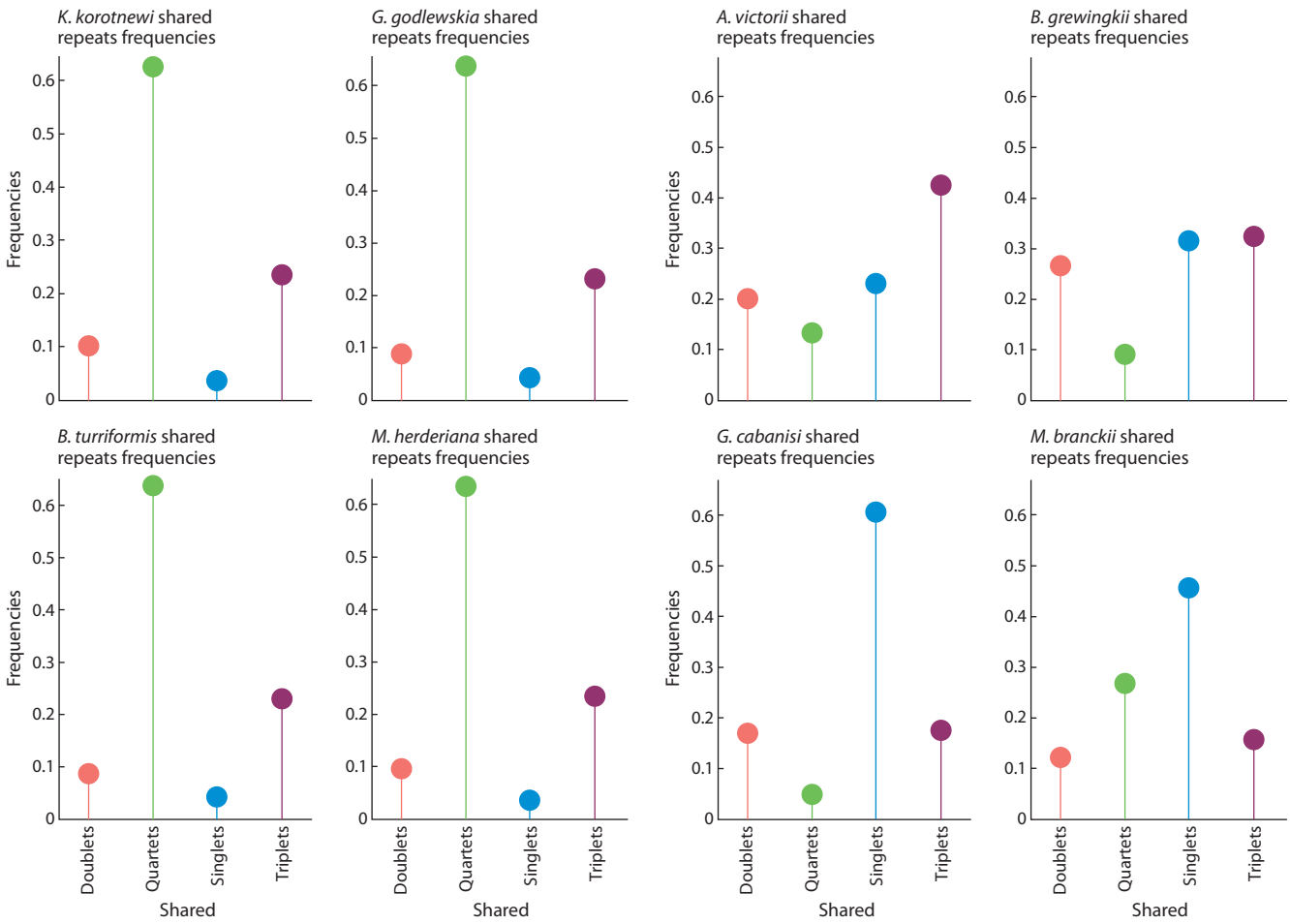


Fig. 4. Species-specific patterns of repeats shared in gastropods (the left four panels) and gastropods (the right four panels).

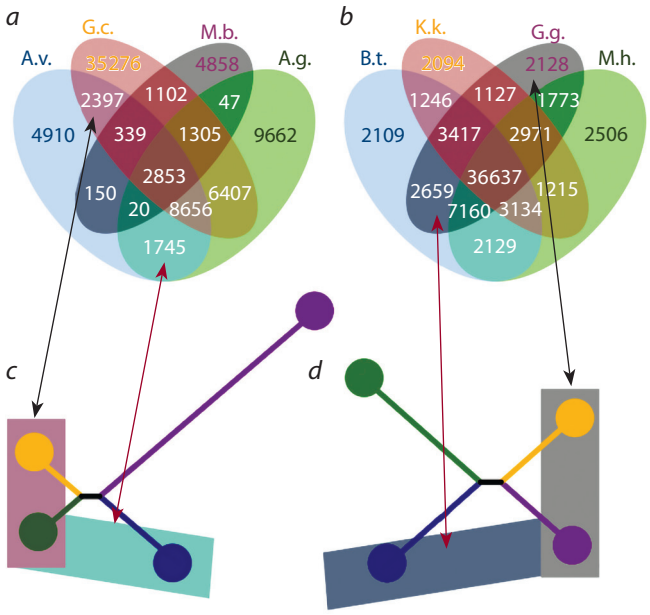


Fig. 5. Venn diagrams of the distribution of repeated DNA between amphipod (*a*) and Baicaliidae (*b*) species. *c*, *d* are unrooted trees inferred from concatenated protein-coding mitochondrial DNA of the corresponding species.

The colors of the branches correspond to the colors of the ovals on the Venn diagrams.

of the branches turned out to be different. In other words, the proportions of common repeats and the degree of differences in nucleotide sequences turned out to be independent, albeit partially correlated features.

The method we used to identify highly repeated sequences and a set of search parameters allow us to identify those that are repeated in the genome with at least 50–100 copies per haploid genome. The detected repeats make up approximately 50 % of the genome and are very diverse ($1 \times 10^4 \ldots 6 \times 10^4$ varieties per genome, see the Table). Therefore, at the present stage of the study, we focused on the integral characteristics of this repetition and the comparison of these characteristics in two flocks of invertebrate species.

The distance tree was inferred from the repeats data using the distance metric calculated from the presence/absence of a repeat class in a sample as justified by blast search under a mild set of parameters. This differs from the parsimony approach employed by (Dodsworth et al., 2015). Its advantage was in avoiding the assumption of strict homology. Nevertheless, like in their study, we obtained the same tree topology to the one inferred from mitogenome sequences for both animal groups studied. Although the topologies were the same, the ratios in branch lengths differed dramatically. We believe that these differences result from the peculiarities of the evolution of the presence/absence of repeats in genomes. The main feature is

that in order to appear in the genome as a repeat, the nucleotide sequence starts as a single copy and must be amplified to such an extent that it can be detected by the repeatexplorer algorithm. The loss of repetition should also go through a gradual decrease in the number of copies of it.

Over time, differences in the compositions of repeated sequences accumulate. This confirms the spread of species-specific sequences in amphipods compared to Baicaliidae and vice versa, with a decrease in the proportion of repeats that occur in all four species. Therefore, the comparison of repeat spectra in a large number of species can be an interesting tool for phylogenetic analysis due to the high diversity of repeats and the fact that a large proportion of the genome is used in such an analysis, which gives hope for obtaining a more adequate and stable picture of evolution. A more detailed examination of pairs of species with common repeats (see Fig. 5) shows that a noticeable, albeit relatively small number of repeats, is common between species that are not sister species. This fraction, if "inconsistent", is present in both amphipods and gastropods.

Conclusion

Of particular interest are the repeats, the distribution of which between species contradicts the topology of phylogenetic trees, but corresponds to the ecological or geographical confinement of species. Such repeats are found in both groups (see the Table and Fig. 5), and in a significant (from hundreds to thousands) amount. From the point of view of phylogenetic analysis, they reduce its resolution but allow us to make an intriguing assumption that some part of them is involved in horizontal transfer between sympatrically inhabiting species. This requires an annotation of this part of the contigs, the results of which will be described elsewhere.

References

Ahmad S.F., Singchat W., Panthum T., Srikulnath K. Impact of repetitive DNA elements on snake genome biology and evolution. *Cells.* 2021;10(7):1707. DOI 10.3390/cells10071707.

Athanasouli M., Rödelsperger C. Analysis of repeat elements in the *Pristionchus pacificus* genome reveals an ancient invasion by horizontally transferred transposons. *BMC Genomics.* 2022;23(1):523. DOI 10.1186/s12864-022-08731-1.

Bazikalova A.Ya. Amphipoda of Lake Baikal. In: Proceedings of the Baikal Limnological Station. Vol. 11. Moscow, 1945. (in Russian)

Bird C.P., Stranger B.E., Dermitzakis E.T. Functional variation and evolution of non-coding DNA. *Curr. Opin. Genet. Dev.* 2006;16(6): 559-564. DOI 10.1016/j.gde.2006.10.003.

Biscotti M.A., Olmo E., Heslop-Harrison J. Repetitive DNA in eukaryotic genomes. *Chromosome Res.* 2015;23(3):415-420. DOI 10.1007/s10577-015-9499-z.

Boto L. Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc. Biol. Sci.* 2014;281(1777):20132450. DOI 10.1098/rspb.2013.2450.

Butina T.V., Bukin Yu.S., Khanaev I.V., Kravtsova L.S., Maikova O.O., Tupikin A.E., Kabilov M.R., Belikov S.I. Metagenomic analysis of viral communities in diseased Baikal sponge *Lubomirskia baikalensis*. *Limnol. Freshw. Biol.* 2019;1:155-162. DOI 10.31951/2658-3518-2019-A-1-155.

Cavalier-Smith T., Beaton M. The skeletal function of non-genic nuclear DNA: new evidence from ancient cell chimaeras. *Genetica.* 1999;106(1-2):3-13. DOI 10.1023/a:1003701925110.

Chen S.C., Sun G.X., Rosen B.P., Zhang S.Y., Deng Y., Zhu B.K., Rensing C., Zhu Y.G. Recurrent horizontal transfer of arsenite methyltransferase genes facilitated adaptation of life to arsenic. *Sci. Rep.* 2017;7(1):7741. DOI 10.1038/s41598-017-08313-2.

Costa M., Manton J.D., Ostrovsky A.D., Prohaska S., Jefferis G.S. NBLAST: rapid, sensitive comparison of neuronal structure and construction of neuron family databases. *Neuron.* 2016;91(2):293-311. DOI 10.1016/j.neuron.2016.06.012.

de Sena Brandine G., Smith A.D. Falco: high-speed FastQC emulation for quality control of sequencing data. *F1000Res.* 2019;8:1874. DOI 10.12688/f1000research.21142.2.

Dodsworth S., Chase M.W., Kelly L.J., Leitch I.J., Macas J., Novák P., Piednoël M., Weiss-Schneeweiss H., Leitch A.R. Genomic repeat abundances contain phylogenetic signal. *Syst. Biol*. 2015;64(1):112-126. DOI 10.1093/sysbio/syu080.

Drozdova P., Saranchina A., Madyarova E., Gurkov A., Timofeyev M. Experimental crossing confirms reproductive isolation between cryptic species within *Eulimnogammarus verrucosus* (Crustacea: Amphipoda) from Lake Baikal. *Int. J. Mol. Sci*. 2022;23(18):10858. DOI 10.3390/ijms231810858.

Gurkov A., Rivarola-Duarte L., Bedulina D., Fernández Casas I., Michael H., Drozdova P., Nazarova A., Govorukhina E., Timofeyev M., Stadler P.F., Luckenbach T. Indication of ongoing amphipod speciation in Lake Baikal by genetic structures within endemic species. *BMC Evol. Biol*. 2019;19(1):138. DOI 10.1186/s12862-019-1470-8.

Hausdorf B., Röpstorf P., Riedel F. Relationships and origin of endemic Lake Baikal gastropods (Caenogastropoda: Rissooidea) based on mitochondrial DNA sequences. *Mol. Phylogenet. Evol*. 2003;26(3): 435-443. DOI 10.1016/s1055-7903(02)00365-2.

Hou Z., Sket B. A review of gammaridae (crustacea: Amphipoda): the family extent, its evolutionary history, and taxonomic redefinition of genera. *Zool. J. Linn. Soc.* 2016;176(2):323-348. DOI 10.1111/zoj.12318.

Jalili V., Afgan E., Gu Q., Clements D., Blankenberg D., Goecks J., Taylor J., Nekrutenko A. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res*. 2020;48(W1):W395-W402. DOI 10.1093/nar/gkaa434.

Kamaltynov R.M. On the higher classification of Lake Baikal amphipods. *Crustaceana.* 1999;72(8):933-944.

Kejnovsky E., Jedlicka P. Nucleic acids movement and its relation to genome dynamics of repetitive DNA: is cellular and intercellular movement of DNA and RNA molecules related to the evolutionary dynamic genome components? *BioEssays.* 2022;44(4):e2100242. DOI 10.1002/bies.202100242.

Kozhov M. Lake Baikal and Its Life. Monographiae Biologicae. Vol. 11. Dordrecht: Springer, 1963. DOI 10.1007/978-94-015-7388-7.

Lee I.P.A., Eldakar O.T., Gogarten J.P., Andam C.P. Bacterial cooperation through horizontal gene transfer. *Trends Ecol. Evol*. 2022; 37(3):223-232. DOI 10.1016/j.tree.2021.11.006.

Lerat E., Casacuberta J., Chaparro C., Vieira C. On the importance to acknowledge transposable elements in epigenomic analyses. *Genes.* 2019;10(4):258. DOI 10.3390/genes10040258.

Li Y., Liu Z., Liu C., Shi Z., Pang L., Chen C., Chen Y., Pan R., Zhou W., Chen X.X., Rokas A., Huang J., Shen X.X. HGT is widespread in insects and contributes to male courtship in lepidopterans. *Cell.* 2022;185(16):2975-2987.e10. DOI 10.1016/j.cell.2022.06.014.

Lipaeva P., Vereshchagina K., Drozdova P., Jakob L., Kondrateva E., Lucassen M., Bedulina D., Timofeyev M., Stadler P., Luckenbach T. Different ways to play it cool: transcriptomic analysis sheds light on different activity patterns of three amphipod species under long-term cold exposure. *Mol. Ecol*. 2021;30(22):5735-5751. DOI 10.1111/mec.16164.

Mats V.D., Shcherbakov D.Y., Efimova I.M. Late Cretaceous–Cenozoic history of the Lake Baikal depression and formation of its unique biodiversity. *Stratigr. Geol. Correl.* 2011;19(4):404-423. DOI 10.1134/S0869593811040058.

Naumenko S.A., Logacheva M.D., Popova N.V., Klepikova A.V., Penin A.A., Bazykin G.A., Etingova A.E., Mugue N.S., Kondrashov A.S., Yampolsky L.Y. Transcriptome-based phylogeny of endemic Lake Baikal amphipod species flock: fast speciation accom-

panied by frequent episodes of positive selection. *Mol. Ecol.* 2017; 26(2):536-553. DOI 10.1111/mec.13927.

Novák P., Neumann P., Pech J., Steinhaisl J., Macas J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics.* 2013;29(6):792-793. DOI 10.1093/bioinformatics/btt054.

Peretolchina T., Sitnikova T.Y., Sherbakov D.Y. The complete mitochondrial genomes of four Baikal molluscs from the endemic family Baicaliidae (Caenogastropoda: Truncatelloida). *J. Molluscan Stud.* 2020;86(3):201-209. DOI 10.1093/mollus/eyaa004.

Rocha A., Dalgarno A., Neretti N. The functional impact of nuclear reorganization in cellular senescence. *Brief. Funct. Genomics*. 2022; 21(1):24-34. DOI 10.1093/bfgp/elab012.

Romanova E.V., Aleoshin V.V., Kamaltynov R.M., Mikhailov K.V., Logacheva M.D., Sirotinina E.A., Gornov A.Y., Anikin A.S., Sherbakov D.Y. Evolution of mitochondrial genomes in Baikalian amphipods. *BMC Genomics*. 2016;17(Suppl.14):1016. DOI 10.1186/s12864-016-3357-z.

Saitou N., Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 1987;4(4):406-425. DOI 10.1093/oxfordjournals.molbev.a040454.

Shen W., Le S., Li Y., Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PloS One.* 2016;11(10): e0163962. DOI 10.1371/journal.pone.0163962.

Sherbakov D.Y. Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. *Trends Ecol. Evol.* 1999;14(3):92-95. DOI 10.1016/s0169-5347(98)01543-2.

Silva B.S.M.L., Heringer P., Dias G.B., Svartman M., Kuhn G.C.S. *De novo* identification of satellite DNAs in the sequenced genomes of *Drosophila virilis* and *D. americana* using the RepeatExplorer and TAREAN pipelines. *PLoS One.* 2019;14(12):e0223466. DOI 10.1371/journal.pone.0223466.

Sitnikova T.Y. Endemic gastropod distribution in Baikal. *Hydrobiologia.* 2006;568(1):207-211. DOI 10.1007/s10750-006-0313-y.

Sitnikova T., Roepstorf P., Riedel F. Reproduction, duration of embryogenesis, egg capsules and protoconchs of gastropods of the family Baicaliidae (Caenogastropoda) endemic to Lake Baikal. *Malacologia.* 2001;43(1-2):59-85.

Steensels J., Gallone B., Verstrepen K.J. Interspecific hybridization as a driver of fungal evolution and adaptation. *Nat. Rev. Microbiol.* 2021; 19(8):485-500. DOI 10.1038/s41579-021-00537-4.

Takhteev V. On the current state of taxonomy of the Baikal Lake amphipods (Crustacea: Amphipoda) and the typological ways of constructing their system. *Arthropoda Selecta.* 2019;28(3):374-402. DOI 10.15298/arthsel.28.3.03.

Thakur J., Packiaraj J., Henikoff S. Sequence, chromatin and evolution of satellite DNA. *Int. J. Mol. Sci.* 2021;22(9):4309. DOI 10.3390/ijms22094309.

Titievsky A., Putintseva Y.A., Taranenko E.A., Baskin S., Oreshkova N.V., Brodsky E., Sharova A.V., Sharov V.V., Panov J., Kuzmin D.A., Brodsky L., Krutovsky K.V. Comparative genomics analysis of repetitive elements in ten gymnosperm species: "dark repeatome" and its abundance in conifer and gnetum species. *Life.* 2021;11(11):1234. DOI 10.3390/life11111234.

Yakhnenko A., Itskovich V. Analysis of mtDNA variability in closely related Baikal sponge species for new barcoding marker development. *Limnology.* 2020;21(1):49-57. DOI 10.1007/s10201-019-00599-7.

Zubakov D.Y., Shcherbakov D.Y., Sitnikova T.Y. Phylogeny of the endemial Baicaliidae molluscs inferred from partial nucleotide sequences of the CO1 mitochondrial gene. *Mol. Biol.* 1997;31(6): 935-939.

ORCID ID

T.E. Peretolchina orcid.org/0000-0002-2950-9762
D.Y. Sherbakov orcid.org/0000-0002-1410-392X

Acknowledgements. The research was supported by the State Project No. 0279-2021-0010.

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

Received December 23, 2022. Revised February 20, 2023. Accepted February 20, 2023.

Вавиловский журнал генетики и селекции. 2023;27(4):357-365
DOI 10.18699/VJGB-23-43

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

Body composition as an indicator of metabolic changes in mice obtained by *in vitro* fertilization

M.V. Anisimova[1], Yanli Gon[1], G.V. Kontsevaya[1], A.V. Romashchenko[1], N.V. Khotskin[1], A.K. Stanova[1], L.A. Gerlinskaya[1] ✉, M.P. Moshkin[1, 2]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Tomsk State University, Department of Vertebrate Zoology and Ecology, Tomsk, Russia
✉ lgerlinskaya@gmail.com

Abstract. To identify body systems subject to epigenetic transformation during *in vitro* fertilization (IVF), comparative morphological and functional studies were performed on sexually mature offspring of outbred CD1 mice, specific-pathogen-free (SPF), obtained by IVF (experiment) and natural conception (control). The studies included assessment of age-related changes in body weight and composition, energy intake and expenditure, and glucose homeostasis. To level the effects caused by the different number of newborns in the control and in the experiment, the size of the fed litters was halved in the control females. Males obtained using the IVF procedure were superior in body weight compared to control males in all age groups. As was shown by analysis of variance with experiment/control factors, gender, age (7, 10 and 20 weeks), the IVF procedure had a statistically significant and unidirectional effect on body composition. At the same time, IVF offspring outperformed control individuals in relative fat content, but were behind in terms of lean mass. The effect of the interaction of factors was not statistically significant. IVF offspring of both sexes had higher fat to lean mass ratios (FLR). Since adipose tissue contributes significantly less to total energy intake compared to muscle, the main component of lean mass, it is not surprising that at the same level of IVF locomotor activity offspring consumed less food than controls. When converted to one gram of body weight, this difference reached 19 %. One of the consequences of reduced utilization of IVF energy substrates by offspring is a decrease in their tolerance to glucose loading. The integral criterion for the effectiveness of restoring the initial glucose level is the area under the curve (AUC), the value of which was 2.5 (males) and 3.2 (females) times higher in IVF offspring compared to the corresponding control. Thus, the totality of our original and literature data shows an increase in the risk of metabolic disorders in IVF offspring, which is confirmed by epidemiological studies of a relatively young cohort of people born using assisted reproductive technologies.

Key words: *in vitro* fertilization; mature offspring; epigenetic transformation; body composition; feed consumption; glucose tolerance.

For citation: Anisimova M.V., Gon Yanli, Kontsevaya G.V., Romashchenko A.V., Khotskin N.V., Stanova A.K., Gerlinskaya L.A., Moshkin M.P. Body composition as an indicator of metabolic changes in mice obtained by *in vitro* fertilization. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):357-365. DOI 10.18699/VJGB-23-43

Композиция тела как индикатор метаболических изменений у мышей, полученных путем оплодотворения *in vitro*

М.В. Анисимова[1], Л. Гон[1], Г.В. Концевая[1], А.В. Ромащенко[1], Н.В. Хоцкин[1], А.К. Станова[1], Л.А. Герлинская[1] ✉, М.П. Мошкин [1, 2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Томский государственный университет, кафедра зоологии позвоночных и экологии, Томск, Россия
✉ lgerlinskaya@gmail.com

Аннотация. Для выявления систем организма, подверженных эпигенетической трансформации при оплодотворении *in vitro* (IVF), были выполнены сравнительные морфофункциональные исследования половозрелых потомков мышей аутбредной линии CD1, свободных от патогенов (SPF-статус), полученных путем IVF (опыт) и при естественном зачатии (контроль). Исследования включали в себя оценку возрастной динамики массы и композиций тела, потребления и расходования энергии, а также глюкозный гомеостаз. Для нивелирования эффектов, обусловленных разным числом новорожденных в контроле и в опыте, у контрольных самок сокращали вдвое размеры выкармливаемых пометов. Самцы, полученные с использованием процедуры IVF, превосходили по массе тела контрольных самцов во всех возрастных группах. Как показал дисперсионный анализ с факторами «опыт/контроль», «пол», «возраст» (7, 10 и 20 недель), процедура IVF статистически значимо и однонаправленно влияла на композиционный состав тела. При этом IVF потомки превосходили контрольных особей по относительному содержанию жира, но проигрывали им по значениям тощей массы. Эффект взаимодействия факторов был

статистически не значимым. У IVF потомков обоего пола были отмечены большие значения отношений жира к тощей массе (FLR). Поскольку жировая ткань вносит значительно меньший вклад в общее потребление энергии по сравнению с мышцами – основным компонентом тощей массы, то неудивительно, что при одинаковом уровне двигательной активности IVF потомки потребляли меньше корма, чем контрольные животные. При пересчете на один грамм массы тела эта разница достигала 19 %. Одним из следствий пониженной утилизации энергетических субстратов IVF потомками является снижение их толерантности к нагрузке глюкозой. Интегральным критерием эффективности восстановления исходного уровня глюкозы служит площадь под кривой (AUC), величина которой была в 2.5 раза (самцы) и 3.2 раза (самки) выше у IVF потомков по сравнению с соответствующим контролем. Таким образом, совокупность собственных и литературных данных показывает увеличение риска метаболических нарушений у IVF потомков, что подтверждают эпидемиологические исследования сравнительно молодой когорты людей, рожденных с применением вспомогательных репродуктивных технологий.
Ключевые слова: фертилизация *in vitro*; половозрелые потомки; эпигенетическая трансформация; композиция тела; потребление корма; глюкозотолерантность.

Introduction

One of the civilizational problems significantly affecting the health of new generations is the increasing spread of assisted reproductive technologies (ART), including *in vitro* fertilization (IVF), as well as intracytoplasmic sperm injection (ICSI), *in vitro* cultivation of preimplantation embryos and embryo transfer to surrogate mothers. In the 44 years since the first successful IVF pregnancy, the number of people conceived *in vitro* has exceeded 10 million and accounts for about 2 % of all newborns in developed countries (Wyns et al., 2018). In Russia, their number exceeds 130,000, of whom 90,000 have been born in the last five years, and the percentage of successful pregnancies with the use of ART has now increased approximately threefold (Russian Association…, 2019).

Despite the youthfulness of the generation of people born with ART, this group of offspring has a higher risk of diabetes, metabolic disorders, arterial hypertension, and neuropsychiatric disorders, compared to those observed in the same-age groups of offspring of natural conception (Hart, Norman, 2013; Hyrapetian et al., 2014; Duranthon, Chavatte-Palmer, 2018; Halliday et al., 2019). These results allow us to predict a more rapid development of age-related pathologies, which may become a real public health problem. According to data from specialized clinics, IVF is used not only by couples where the inability to conceive is due to the age of one or both partners, but also by patients with health disorders, in particular overweight, disorders of psycho-emotional status and other diseases (Cauldwell et al., 2017; Farquhar et al., 2019). Therefore, based on clinical observations alone, it is difficult to differentiate between the contribution to potential health impairments of the IVF procedure itself and the genetic and physiological characteristics of the parents.

The most adequate approach to assess the phenotypic effects of IVF and ART is experimental studies performed under controlled conditions on standardized laboratory animals. It is the experiment that can reveal the pros and cons of *in vitro* fertilization in solving the demographic problems of modern society. It should be noted that the experimental data available in the literature support the phenotypic significance of IVF (Roy et al., 2017; La Rovere et al., 2019). One of the actively developed aspects of phenotypic modulation of offspring born with ART refers to the increased risk of metabolic abnormalities (Heber, Ptak, 2020). Experiments on laboratory mice have provided evidence of an independent role of IVF in the formation of metabolic syndrome and obesity (Feuer et al., 2014), body composition (Sjöblom et al., 2005), and changes in carbohydrate homeostasis and predisposition to diabetes (Scott et al., 2010). However, these effects of IVF have not been confirmed in all studies and vary depending on the sex of the animals and conditions of embryo development outside the maternal body (Donjacour et al., 2014). At the same time, the question of the key factors that determine the manifestation of the metabolic syndrome in adult offspring obtained by *in vitro* fertilization remains out of sight.

Metabolic syndrome is a combination of hyperglycemia, abdominal obesity, dyslipidemia, and hypertension, and their manifestation is determined by eating behavior, physical activity, and food intake (Sousa, Norman, 2016). The most important factor leading to the development of metabolic syndrome is a change in the balance between energy expenditure, and its compensation by food calories. Moreover, fat accumulation is determined not only by the amount of food consumed, but also by its distribution in the daily cycle (Gill, Panda, 2015). At the same time, the analysis of the eating behavior of IVF mice is found in sporadic studies and is limited to the estimation of daily feed intake without analyzing circadian dynamics and without comparing it with the level of locomotor activity (Feuer et al., 2014).

Since the above-mentioned deviations of individual development are interrelated, it is of fundamental importance to investigate them comprehensively within a single experiment. But, as a rule, these works are limited to the study of individual phenotypic characteristics at different stages of individual development, which makes it difficult to analyze the cause-effect relationships between successive ontogenetic events.

In our work, we investigated the effect of ART on the formation of interdependencies of indicators of daily activity dynamics, feed intake, glucose homeostasis, and body composition associated with the risk of metabolic syndrome in naturally conceived and IVF-obtained sexually mature CD1 line mice progeny. The CD1 line mice do not have their own unique MHC haplotype (Marín et al., 2014) and this circumstance allows us to exclude the influence of the MHC haplotype differences between the embryos and the gestating mother on embryo development during pregnancy and, consequently, on the phenotype of adult progeny (Gerlinskaya, Evsikov, 2001; Rapacz-Leonard et al., 2014). We showed that IVF-obtained progeny of CD1 line are characterized by excess body weight, which is combined with an increase in the relative proportion of fat and with reduced tolerance to glucose load.

Table 1. Progeny studied

Group	Gender	Number of offspring			
		Body mass	Body composition	Feed intake, locomotor activity	Glucose tolerance test
IVF	Males	21	14	15	16
	Females	18	15	13	11
Control	Males	23	15	11	8
	Females	14	10	11	8

Materials and methods

The study was performed in the Center for Genetic Resources of Laboratory Animals of the Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, on outbred CD1 mice free from species-specific pathogens (SPF status). SPF status compliance was confirmed by pathogen analysis according to the European Laboratory Animal Health Association (FELASA) list.

The animals were kept in individually ventilated OptiMice cages (USA). The controlled environmental conditions had the following parameters: photoperiod 14C:10T, temperature 22–24 °C and humidity 40–50 %. Gradual switching off of the light began at 16:00 local time. Dusted birch pellets (Albion LLC, Novosibirsk) were used as bedding material. Food (SNIFF, Germany) and water were given without restrictions. Feed and bedding were given to animals after autoclaving (121 °C).

Experimental groups. IVF group – males and females obtained using *in vitro* fertilization (IVF), cultivation and embryo transfer; Control group – males and females obtained as a result of natural mating.

Offspring: *in vitro* fertilization, culture conditions and embryo transfers. IVF procedures, embryo culture conditions and embryo transfers were performed according to the technique described previously (Kontsevaya et al., 2021). IVF oocytes were obtained from female CD1 mice after superovulation by intraperitoneal injection of 5 IU of pregnant mare serum gonadotropin (PMSG) (Intervet International, Netherlands) followed by 5 IU of human chorionic gonadotropin (chorulon) (Intervet International, Netherlands) at 48 h intervals. Cumulus-oocyte complexes collected from the oocyte ampoule 17–18 h after hCG injection were placed in a drop (200 µl) of HTF medium (Human Tubal Fluid, Irvine Scientific, USA) for fertilization. To obtain spermatozoa, the caudal part of the epididymis was placed in HTF medium, and after incubation (1 h), a 3–5 µl drop containing spermatozoa was added to the oocytes and incubated for 4–5 h. Fertilized oocytes were washed in four drops of HTF medium and cultured for 72 h in 60 µl KSOM medium 8–12 embryos per drop under mineral oil (Sigma), at 37 °C and 5 % CO_2 until the blastocyst stage. The developmental efficiency of the embryos after IVF was 75.5 ± 2.86 % at 2 cells stage and 70.6 ± 4.64 % from 2 cells stage to blastocysts.

Blastocysts were transferred to CD1 pseudopregnant females on day 2.5 of pseudopregnancy (Kontsevaya et al., 2021). Pseudopregnancy was induced by mating females with vasectomized males of the same line. Male vasectomies were performed by thermal cauterization of the vas deferens at least 2 weeks before mating. Surgical procedures were performed under general anesthesia (Domitor, Orion Pharma, Finland – 15 µg/100 g body weight and Zoletil, Virbac, France – 3 mg/100 g body weight, intraperitoneally). The morning after mating, the females were examined and, in the presence of vaginal plugs, transferred to separate cages. On day 2.5 of pseudopregnancy, 17 females were surgically transferred 12 blastocysts each into the left oviduct under gas anesthesia (Aerrane, Baxter Healthcare Corp., USA). After embryonic transfers, the females were placed in separate cages until delivery. Since, as previously shown, the surgical procedures performed in embryonic transfers (narcotization and introduction of culture medium into the uterus) did not affect the course of pregnancy and hormonal background (Gerlinskaya, Evsikov, 2001), therefore, control males were obtained by natural mating.

After IVF procedures and embryo transfers, 13 females (76.5 %) gave birth and all newborn offspring were fed without losses. The offspring after maternal feeding (3 weeks) were weaned from the mothers and further kept in single-sex groups of 5 individuals in each cage. The average number of nursed offspring produced by IVF was 3.6 ± 0.24 and was significantly lower compared to 12.5 ± 0.58 in natural pregnancy. The decrease in litter size is caused by transplanting blastocysts into only one uterine horn, whereas in a natural pregnancy, fetuses develop in two horns. In turn, litter size affects maternal behavior and offspring development (Enes-Marques, Giusti-Paiva, 2018). To compensate for the effects of the number of nursed offspring, females in the control group had some of their newborns removed and the number of nursed offspring was reduced by 4–5 individuals per litter. Offspring from 13 IVF litters and 8 reduced control litters were examined (Table 1).

Body mass and body composition of the offspring were determined at 7–8, 10–11, and 19–20 weeks of age. Fat and lean mass were quantified using a low-field NMR (nuclear magnetic resonance) spectrometer EchoMRI (USA).

Locomotor activity and feed consumption were measured in males and females of the control and experimental groups at 11–12 weeks of age. Animals were housed one at a time in Phenomaster cages (TSE, Germany). After a 2-day period of habituation, we measured the distance traveled (spontaneous locomotor activity), water and feed consumption for 3 days. When analyzing circadian rhythms of locomotor and food activity, the values of the analyzed parameters were summed at 1-hour intervals. For intergroup comparisons of locomotor activity, feed and water consumption, we used the total values for the 3-day observation period.

Table 2. Age, gender and experimental group effects on body mass and body composition (two-factor analysis of variance)

Indicators	Age		Gender		Group	
	$F_{1.138}$	p	$F_{1.138}$	p	$F_{1.138}$	p
Body mass, g	34.97	<0.001	114.69	<0.001	16.96	<0.001
Fat mass, g	9.614	<0.001	0.733	=0.393	20.96	<0.001
% fat	1.10	=0.339	14.66	<0.001	13.96	<0.001
Lean mass, g	37.25	<0.001	290.79	<0.001	11.07	=0.001
% lean mass	2.14	=0.121	29.17	<0.001	6.73	=0.011
FLR	1.47	=0.23	20.28	<0.001	15.71	<0.001

Note. The effects of factor interactions were statistically unreliable and therefore are not included in the table.

Glucose tolerance test (GTT) was performed according to the standard technique on males and females of control and experimental groups at the age of 10–11 weeks. 16 hours before glucose injection, the feeder was removed from the mice maintenance cages. Glucose (PanEco, Russia) was injected intraperitoneally at the rate of 10 µl 20 % glucose per 1 g mouse weight. Blood was taken from the tip of the tail at five time points: 0 – baseline level before intraperitoneal injection of glucose solution, 1 – after 15 min, 2 – after 30 min, 3 – after 60 min, and 4 – after 120 min after intraperitoneal injection of glucose solution. Blood glucose levels were measured using a Contour TS glucose meter (Bayer, Switzerland). The area under the curve of increase from the baseline glucose concentration level (average under curve – AUC) was calculated by numerical integration as an integral index of GGT.

Statistical analysis was started by assessing the nature of the distribution of empirical data. According to the Kolmogorov–Smirnov criterion, all variation series of the analyzed traits corresponded to normal distribution. Therefore, two- or three-factor analysis of variance and analysis of variance with repeated measures were used to determine the effects of experimental group, age, and sex. Comparisons of the 2 mean values were performed using Student's test (*t*-test). Data are presented as means and errors (mean ± SE).

Results

Mass and body composition

Analysis of variance with the factors of experiment/control, sex, and age (7–8, 10–11, and 19–20 weeks) showed that the IVF procedure had a statistically significant effect on body weight and composition (Table 2, Fig. 1).

There was no interaction of factors for any of the analyzed traits, indicating a unidirectional effect of *in vitro* fertilization in different sex and age groups. Analysis of the effects of IVF performed separately for males and females revealed statistically significant differences in body weight between control and experiment only in males whose IVF progeny outweighed (43.7 ± 0.7 g) the control (38.0 ± 0.8 g) in all age groups (from 3 to 20 weeks). The statistical significance of IVF was confirmed by a two-factor analysis of variance with control/experiment and age factors ($F_{1,77} = 28.6$, $p < 0.001$). The same technique was applied to analyze the effects of IVF on body composition in males and females. Total fat and lean mass were higher in IVF males (6.0 ± 0.3 and 34.2 ± 0.4 g, respectively) than in

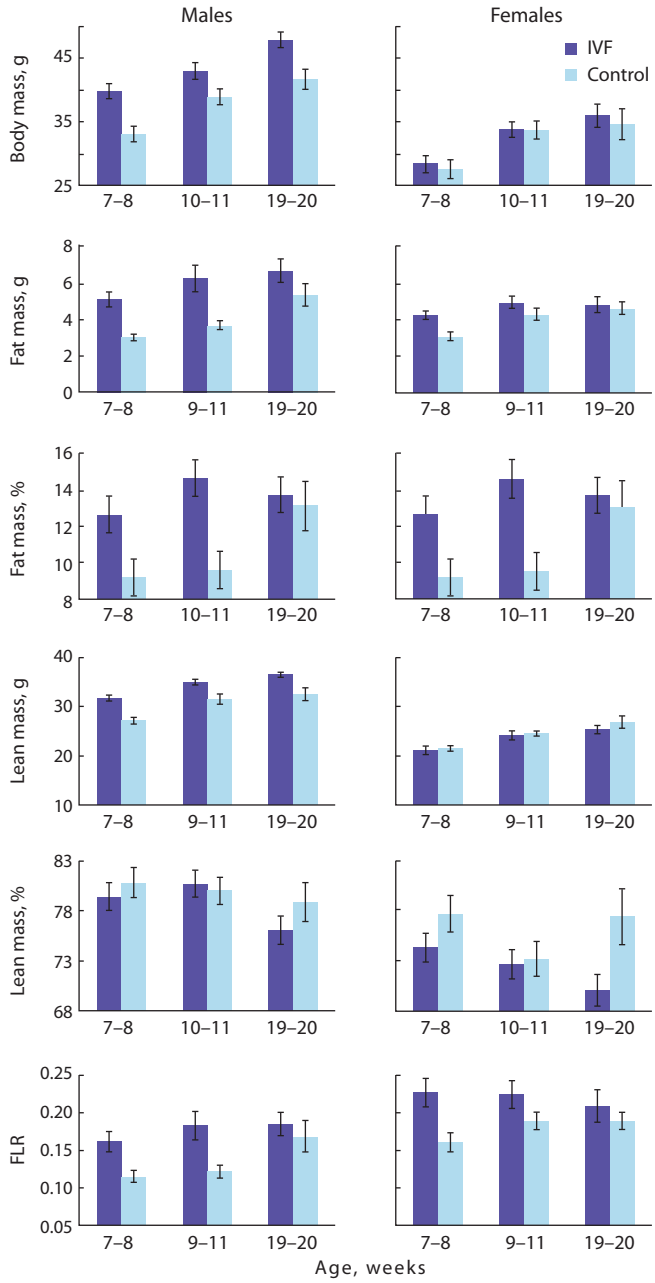


Fig. 1. Mass and body composition in IVF and control groups at different ages.

FLR is the ratio of fat to lean.

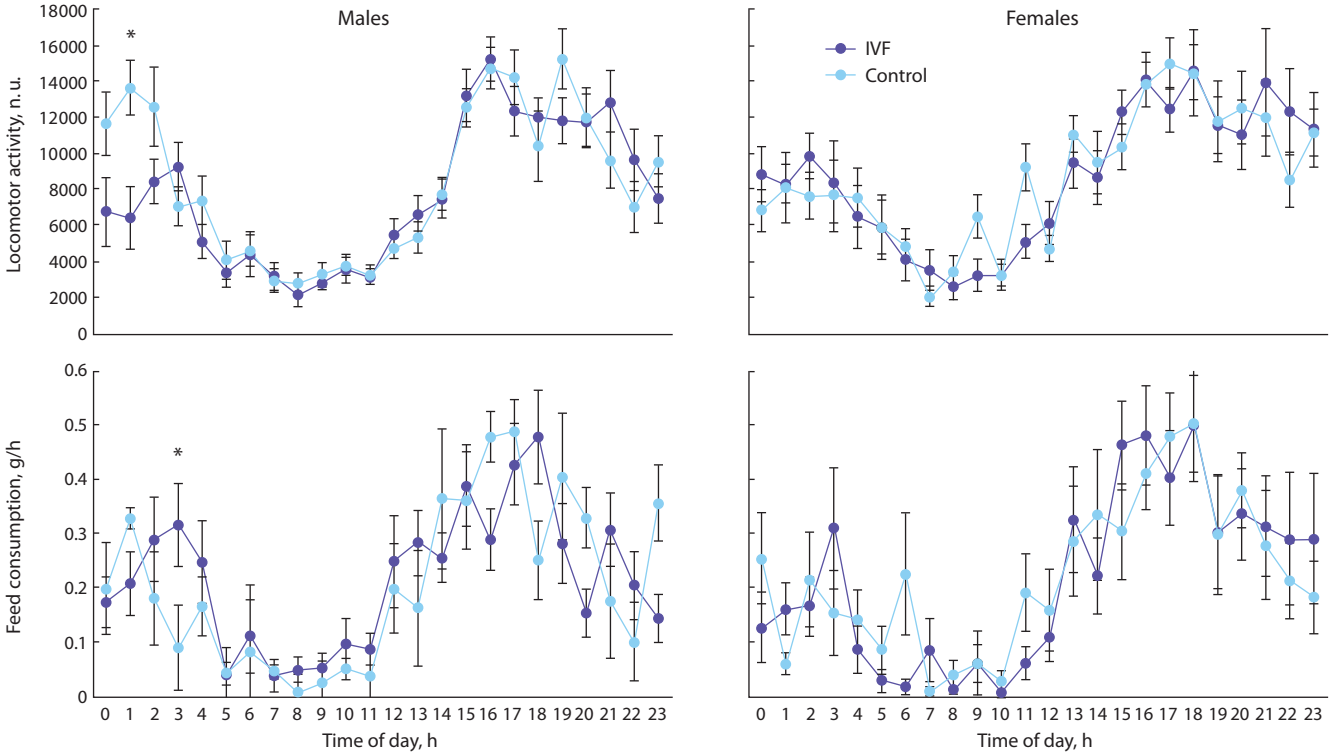


Fig. 2. Daily dynamics of locomotor activity and feed intake in progeny obtained by natural conception (Control) or by IVF.
* $p \le 0.05$, ANOVA with repeated measure: $F_{1.16} = 5.06$ for motor activity and $F_{1.16} = 5.16$ for feed intake.

Table 3. Effects of gender and experimental group on locomotor activity and feed consumption (two-factor analysis of variance)

Factors	*df*	Locomotor activity		Feed consumption/body mass		Feed consumption/activity	
		F	*p*	*F*	*p*	*F*	*p*
Gender	1	6.45	=0.014	1.37	=0.25	11.84	<0.010
Group	1	0.11	=0.74	5.16	=0.028	0.17	=0.68
Gender × Group	1	0.48	=0.49	0.03	=0.86	0.57	=0.45
Error	46						

controls (4.0±0.3 and 30.3±0.4 g; $F_{1,77} = 20.5$, $p < 0.001$ and $F_{1,77} = 40.0$, $p < 0.001$). In females, IVF had a statistically significant effect only on fat content: IVF, 5.1±0.2 g, control, 4.4±0.3 g; $F_{1,61} = 4.4$, $p = 0.04$.

The effects listed above are due in part to the effect of IVF on animal body weight. But intergroup differences (IVF *vs* control) persisted when analyzing relative body composition indices. Thus, the percentage of fat in IVF males (13.4±0.6 %) and females (15.7±0.5 %) exceeded that of controls (males, 10.6±0.7 %; $F_{1,77} = 10.5$, $p = 0.002$; females, 13.6±0.8 %; $F_{1,61} = 4.7$, $p = 0.03$). In contrast to fat, the proportion of lean mass was higher in controls, but the effect of IVF was statistically significant only in females: control, 76.2±1.3 %, IVF, 72.4±0.9 %; $F_{1,61} = 5.8$, $p = 0.02$.

FLR also depended on sex and experimental group. In this case, males and females obtained by *in vitro* fertilization surpassed control individuals in FLR: IVF males, 0.175±0.008 and control, 0.133±0.009; $F_{1,77} = 10.2$, $p = 0.002$; IVF females, 0.221 ± 0.009 and control, 0.180 ± 0.014; $F_{1,61} = 6.1$, $p = 0.016$.

Locomotor activity and feed consumption

Monitoring of locomotor activity and feed consumption showed typical circadian changes in the studied indices in mice (Fig. 2). Statistically significant differences between the control and experimental groups were found only in males during the second half of the active period, i. e., from 00:00 to 04:00 hours (local time). The control individuals showed an increase in activity, which was statistically significantly higher than that of the IVF progeny, at 01:00 h. In turn, IVF progeny showed higher feed intake than control individuals at 03:00 h.

Analysis of variance with experiment/control and gender factors of the results of 3-day monitoring of locomotor activity and feed consumption showed that spontaneous locomotor activity was independent of whether the animals belonged to the control or experimental (IVF) group (Table 3). For feed intake per 1 g body weight, a significant effect of IVF was found: control individuals consumed more feed (0.630±0.037 g/g, $n = 26$) than IVF-derived individuals (0.511±0.036 g/g, $n = 24$). Feed consumption per unit traveled was the same in the control and experimental groups. It should be noted that

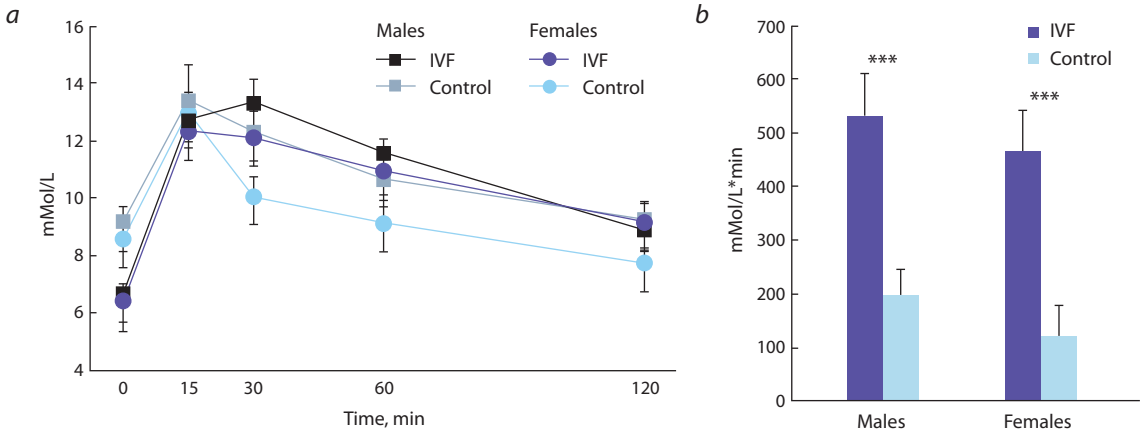


Fig. 3. Glucose tolerance test: *a*, glucose concentration; *b*, AUC.
*** $p < 0.001$ between experimental and control offspring groups (Student's *t*-test).

Table 4. Effects of gender and experimental group on basal and maximal glucose levels and on AUC in the glucose tolerance test (two-factor analysis of variance)

Factors	*df*	Basal level		Maximal level		AUC	
		F	*p*	*F*	*p*	*F*	*p*
Gender	1	0.78	=0.38	0.28	=0.60	0.66	=0.42
Group	1	21.079	<0.001	0.65	=0.43	14.35	<0.001
Gender × Group	1	0.06	=0.81	0.00	=0.99	0.004	=0.95
Error	34						

a statistically significant effect of sex was detected for the studied indicators. At the same time, females showed more spontaneous activity and higher feed consumption compared to males. But feed consumption per unit of the traversed way was 24 % less for them than for males. Statistical significance of sex differences is confirmed by analysis of variance: $F_{1,46} = 10.5$, $p = 0.0022$.

Glucose tolerance test (GTT)

The baseline glucose concentration (time 0 in Fig. 3) in males and females obtained by IVF was significantly lower than in controls (Table 4). The maximum values of glucose recorded 15 min after the injections were similar in individuals of different sexes and different experimental groups. The total deviations of glucose concentration (AUC) differed significantly depending on the affiliation with the experimental group (see Table 4). They were statistically significantly higher in IVF progeny than in controls (see Fig. 3).

Discussion

Despite the equalization of the size of the litter in the control and experimental groups, males obtained by *in vitro* fertilization surpass the body mass of naturally conceived individuals in all age groups. It should be noted that the positive effect of IVF on the growth rate of males is also noted by other authors (Van Montfoort et al., 2012; Donjacour et al., 2014; Narapareddy et al., 2021; Elhakeem et al., 2022). But this effect significantly depends on the conditions of embryo development outside the maternal body. When incubated in a medium with optimized amino acid composition, the body mass of sexually mature IVF males did not differ from controls (Donjacour et al., 2014; Duranthon, Chavatte-Palmer, 2018; Qin et al., 2021) and even exceeded that of females (Feuer et al., 2014).

In contrast to body mass, the effect of IVF on composition was statistically significant in both males and females. The total and relative (% of body mass) fat content was higher in IVF offspring of both sexes. In its turn, IVF males not only had higher body mass compared to control males, but the absolute values of lean mass were higher in them than in control individuals. At the same time, the relative lean mass in control males was superior to that in IVF progeny, at least in females. One metabolically relevant characteristic of body composition is the ratio of total fat mass to lean mass (Seo et al., 2020; Liu et al., 2021). The offspring of both sexes obtained by IVF were 31.6 % (males) and 22.8 % (females) higher in FLR than control individuals.

The most widespread cause of individual variations in fat accumulation is a change in the balance between energy substrate intake and expenditure, particularly for muscular work. Our study revealed no statistically significant differences in the level of spontaneous activity between control and IVF progeny. And feed consumption per unit body mass was lower

in IVF offspring than in naturally-raised individuals. Experimental and clinical studies indicate that fat accumulation increases when the main food intake shifts to the end of the active phase of the diurnal cycle (Gill, Panda, 2015; Panda, 2016; Wilkinson et al., 2020). A statistically significant excess of feed consumption 3 h before the end of the dark time of the day was noted in IVF males, which exceeded the control individuals in this indicator. In females of control and experimental groups, the dynamics of feed consumption was the same. But greater fat accumulation compared to the control was characteristic of IVF offspring of both sexes. Therefore, a change in the daily rhythm of feed intake occurring only in males cannot serve as a universal explanation for changes in body composition in IVF offspring.

Along with the balance of locomotor activity and feed intake, no less important for fat accumulation is the rate of utilization of energy substrates. The indicator reflecting the rate of utilization of energy substrates can be the drop in blood glucose concentration during a standard carbohydrate load. The area under curve (AUC), the value of which was 2.5 (males) and 3.2 (females) times higher in IVF progeny compared to the corresponding control, served as an integral criterion of the efficiency of the initial glucose level recovery.

Therefore, IVF-derived mice differ from the control mice in greater fat accumulation combined with lower feed consumption and reduced tolerance to glucose load. This body composition is in good agreement with the literature, according to which individuals with lower basal metabolic rates are more prone to diabetes mellitus (Maciak et al., 2020). The role of body composition may act as one of the significant factors of the observed metabolic changes. Here, the increase in the ratio of fat to lean mass detected in all sex and age groups of animals draws attention. It is known that adipose tissue makes a minimal contribution to total energy intake, which is largely determined by muscle (Seo et al., 2020; Liu et al., 2021), the main component of lean mass.

To summarize, the results demonstrate that a significant increase in the risk of metabolic syndrome in IVF offspring is independent of the amount of feed consumption and locomotor activity. In males, fat accumulation can be accounted for by impaired daily rhythm of feed consumption and decreased glucose utilization rate. In females, the main cause of fat accumulation and, as a consequence, the risk of metabolic syndrome may be associated with changes in the metabolic pathways that ensure the efficient utilization of energy substrates, and this cause is indicated by a significant decrease in glucose tolerance.

It should be noted that our findings indicate possible pathways for the development of the metabolic syndrome, but do not reveal its mechanisms, which may be due to many factors related to the specific effects at different stages of offspring ontogenesis when using the ART complex. In particular, the composition of the culture medium (Khosla et al., 2001; Sjöblom et al., 2005; Zandstra et al., 2018), the oxygen content of the gas medium and the pH of the culture medium (Kelley, Gardner, 2017; Ng et al., 2018), the duration of embryo culture (Johnson, 2019), and other factors associated with surgical embryo transfection procedures (Rozhkova et al., 2017), as well as maternal and fetal immunogenetic differences (Gerlinskaya et al., 2019) affect the phenotype of offspring. Nevertheless, it should be emphasized that these works are limited to the study of individual stages of individual development, which makes it difficult to analyze the cause-effect relationships between successive ontogenetic events. The period of ontogenesis including the first cell division and development of preimplantation embryos is critical and coincides with the global reprogramming of the epigenome and establishment of epigenetic modifications that persist into adulthood. It is likely that epigenetic modifications resulting from exposure to the procedures used in obtaining offspring by ART may play a central role in destabilizing prenatal development and, consequently, in increasing the risk of metabolic syndrome.

One of the criteria used to assess developmental destabilization is fluctuating asymmetry (FA) (Dongen, 2006). The feasibility of using this criterion as an indicator of developmental destabilization is supported by clinical observations showing that FA of fingerprints on the left and right hand with a high degree of reliability is associated with predisposition to diabetes (Morris et al., 2012, 2016; Yohannes et al., 2015).

Conclusion

Thus, the combination of our own and literature data allows us to outline a range of IVF-conditioned interrelated events that include developmental destabilization, a set of metabolic changes, and increased risk of diabetes. However, the mechanistic specification of the effects of IVF requires further research, including expanded studies of the relationships of epigenetic modifications, fluctuating asymmetry, and metabolic regulation. The relevance of such studies, judging by the data presented in the Norrman review (Norrman et al., 2020), is steadily increasing as the cohort of people born with assisted reproductive technology matures.

References

Cauldwell M., Patel R.R., Steer P.J., Swan L., Norman-Taylor J., Gatzoulis M., Johnson M.R. Managing subfertility in patients with heart disease: what are the choices? *Am. Heart. J.* 2017;187:29-36. DOI 10.1016/j.ahj.2017.02.007.

Dongen S.V. Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. *J. Evol. Biol*. 2006; 19(6):1727-1743. DOI 10.1111/j.1420-9101.2006.01175.x.

Donjacour A., Liu X., Lin W., Simbulan R., Rinaud P.F. In vitro fertilization affects growth and glucose metabolism in a sex-specific manner in an outbred mouse model. *Biol. Reprod*. 2014;90(4):80. DOI 10.1095/biolreprod.113.113134.

Duranthon V., Chavatte-Palmer P. Long term effects of ART: what do animals tell us? *Mol. Rep. Dev*. 2018;85(4):348-368. DOI 10.1002/mrd.22970.

Elhakeem A., Taylor A.E., Inskip H.M., Huang J., Tafflet M., Vrijkotte T.G.M., Nelson S.M., Andersen A.-M.N., Magnus M.C., Lawlor D.A. Association of assisted reproductive technology with offspring growth and adiposity from infancy to early adulthood. *JAMA Netw. Open.* 2022;5(7):e2222106. DOI 10.1001/jamanetworkopen.2022.22106.

Enes-Marques S., Giusti-Paiva A. Litter size reduction accentuates maternal care and alters behavioral and physiological phenotypes in rat adult offspring. *J. Physiol. Sci*. 2018;68(6):789-798. DOI 10.1007/s12576-018-0594-8.

Farquhar C.M., Bhattacharya S., Repping S., Mastenbroek S., Kamath M.S., Marjoribanks J., Boivin J. Female subfertility. *Nat. Rev. Dis. Primers*. 2019;5(1):7. DOI 10.1038/s41572-018-0058-8.

Feuer S.K., Liu X., Donjacour A., Lin W., Simbulan R.K., Giritharan G., Piane L.D., Kolahi K., Ameri K., Maltepe E., Rinaudo H.F. Use of a mouse in vitro fertilization model to understand the developmental origins of health and disease hypothesis. *Endocrinology*. 2014;155(5):1956-1969. DOI 10.1210/en.2013-2081.

Gerlinskaya L.A., Evsikov V.I. Influence of genetic dissimilarity of mother and fetus on progesterone concentrations in pregnant mice and adaptive features of offspring. *Reproduction*. 2001;121(3):409-417. DOI 10.1530/rep.0.1210409.

Gerlinskaya L.A., Litvinova E.A., Kontsevaya G.V., Feofanova N.A., Achasova K.M., Anisimova M.V., Maslennikova S.O., Zolotykh M.A., Moshkin Y.M., Moshkin M.P. Phenotypic variations in transferred progeny due to genotype of surrogate mother. *Mol. Hum. Reprod*. 2019;25(2):88-99. DOI 10.1093/molehr/gay052.

Gill S., Panda S. A smartphone app reveals erratic diurnal eating patterns in humans that can be modulated for health benefits. *Cell Metab*. 2015;3:789-798. DOI 10.1016/j.cmet.2015.09.005.

Halliday J., Lewis S., Kennedy J., Burgner D.P., Juonala M., Hammarberg K., Amor D.J., Doyle L.W., Saffery R., Ranganathan S., Welsh L., Cheung M., McBain J., Hearps S.J.C., McLachlan R. Health of adults aged 22 to 35 years conceived by assisted reproductive technology. *Fertil. Steril*. 2019;112(1):130-139. DOI 10.1016/j.fertnstert.2019.03.001.

Hart R., Norman R.J. The longer-term health outcomes for children born as a result of IVF treatment: Part I – General health outcomes. *Hum. Reprod. Update*. 2013;19(3):232-243. DOI 10.1093/humupd/dms062.

Heber M.F., Ptak G.V. The effects of assisted reproduction technologies on metabolic health and disease. *Biol. Reprod*. 2021;104(4):734-744. DOI 10.1093/biolre/ioaa224.

Hyrapetian M., Loucaides E.M., Sutcliffe A.G. Health and disease in children born after assistive reproductive therapies (ART). *J. Reprod. Immunol*. 2014;106:21-26. DOI 10.1016/j.jri.2014.08.001.

Johnson M.H. A short history of in vitro fertilization (IVF). *Int. J. Dev. Biol*. 2019;63(3-4-5):83-92. DOI 10.1387/ijdb.180364mj.

Kelley R.L., Gardner D.K. In vitro culture of individual mouse preimplantation embryos: the role of embryo density, microwells, oxygen, timing and conditioned media. *Reprod. Biomed. Online.* 2017;34(5):441-454. DOI 10.1016/j.rbmo.2017.02.001.

Khosla S., Dean W., Brown D., Reik W., Feil R. Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. *Biol. Reprod*. 2001;64(3):918-926. DOI 10.1095/biolreprod64.3.918.

Kontsevaya G.V., Gerlinskaya L.A., Moshkin Y.M., Anisimova M.V., Stanova A.K., Babochkina T.I., Moshkin M.P. The effects of sperm and seminal fluid of immunized male mice on in vitro fertilization and surrogate mother–embryo interaction. *Int. J. Mol. Sci*. 2021;22(19):10650. DOI org/10.3390/ijms221910650.

La Rovere M., Franzago M., Stuppia L. Epigenetics and neurological disorders in ART. *Int. J. Mol. Sci*. 2019;20(17):E4169. DOI 10.3390/ijms20174169.

Liu D., Zhong J., Ruan Y., Zhang Z., Sun J., Chen H. The association between fat-to-muscle ratio and metabolic disorders in type 2 diabetes. *Diabetol. Metab. Syndr*. 2021;13:129. DOI 10.1186/s13098-021-00748-y.

Maciak S., Sawicka D., Sadowska A., Prokopiuk S., Buczyńska S., Bartoszewicz M., Niklińska G., Konarzewski M., Car H. Low basal metabolic rate as a risk factor for development of insulin resistance and type 2 diabetes. *BMJ Open Diab. Res. Care*. 2020;8:e001381. DOI 10.1136/bmjdrc-2020-001381.

Marín N., Mecha M., Espejo C., Mestre L., Eixarch H., Montalban X., Álvarez-Cermeño J.C., Guaza C., Villar L.M. Regulatory lymphocytes are key factors in MHC-independent resistance to EAE. *J. Immunol. Res*. 2014;2014:156380. DOI 10.1155/2014/156380.

Morris M.R., Ludwar B.C., Swingle E., Mamo M.N., Shubrook J.H. A new method to assess asymmetry in fingerprints could be used as an early indicator of type 2 diabetes mellitus. *J. Diabetes Sci. Technol*. 2016;10(4):864-871. DOI 10.1177/1932296816629984.

Morris M.R., Rios-Cardenas O., Lyons S., Tudor M.S., Bono L. Fluctuating asymmetry indicates optimization of growth rate over developmental stability. *Funct. Ecol*. 2012;26(3):723-731. DOI 10.1111/j.1365-2435.2012.01983.x.

Narapareddy L., Rhon-Calderon E.A., Vrooman L.A., Baeza J., Nguyen D.K., Mesaros C., Lan Y., Garcia B.A., Schultz R.M., Bartolomei M.S. Sex-specific effects of in vitro fertilization on adult metabolic outcomes and hepatic transcriptome and proteome in mouse. *FASEB J.* 2021;35(4):e21523. DOI 10.1096/fj.202002744R.

Ng K.Y.B., Mingels R., Morgan H., Macklon N., Cheong Y. *In vivo* oxygen, temperature and pH dynamics in the female reproductive tract and their importance in human conception: a systematic review. *Hum. Reprod. Update*. 2018;24(1):15-34. DOI 10.1093/humupd/dmx028.

Norrman E., Petzold M., Clausen T.D., Henningsen A.-K., Opdahl S., Pinborg A., Rosengren A., Bergh C., Wennerholm U. Type 1 diabetes in children born after assisted reproductive technology: a register-based national cohort study. *Hum. Reprod*. 2020;35(1):221-231. DOI 10.1093/humrep/dez227.

Panda S. Circadian physiology of metabolism. *Science*. 2016;354(6315):1008-1015. DOI 10.1126/science.aah4967.

Qin N., Zhou Z., Zhao W., Zou K., Shi W., Yu C., Huang H. Abnormal glucose metabolism in male mice offspring conceived by in vitro fertilization and frozen-thawed embryo transfer. *Front. Cell Dev. Biol*. 2021;9:637781. DOI 10.3389/fcell.2021.637781.

Rapacz-Leonard A., Dąbrowska M., Janowski T. Major histocompatibility complex I mediates immunological tolerance of the trophoblast during pregnancy and may mediate rejection during parturition. *Mediators Inflamm*. 2014;214:579279. DOI 10.1155/2014/579279.

Roy M.C., Dupras C., Ravitsky V. The epigenetic effects of assisted reproductive technologies: ethical considerations. *J. Dev. Orig. Health. Dis*. 2017;8(4):436-442. DOI 10.1017/S2040174417000344.

Rozhkova I.N., Igonina T.N., Ragaeva D.S., Petrova O.M., Brusentsev E.Y., Naprimerov V.A., Amstislavsky S.Y. Long-term effects of maternal exposure to surgical stress at the earliest stage of pregnancy on blood pressure and behavior in offspring of OXYS rats. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2017;21(8):937-942. DOI 10.18699/VJ17.316. (in Russian)

Russian Association of Human Reproduction. National Register of ART of 2019. https://www.rahr.ru/d_registr_otchet/Registr ART2019.pdf. (in Russian)

Scott K.A., Yamazaki Y., Yamamoto M., Lin Y., Melhorn S.J., Krause E.G., Woods S.C., Yanagimachi R., Sakai R.R., Tamashiro K. Glucose parameters are altered in mouse offspring produced by assisted reproductive technologies and somatic cell nuclear transfer. *Biol. Reprod*. 2010;83(2):220-227. DOI 10.1095/biolreprod.109.082826.

Seo Y.-G., Song H.J., Song Y.R. Fat-to-muscle ratio as a predictor of insulin resistance and metabolic syndrome in Korean adults. *J. Cachexia Sarcopenia Muscle*. 2020;11(3):710-725. DOI 10.1002/jcsm.12548.

Sjöblom C., Roberts C.T., Wikland M., Robertson S.A. Granulocyte-macrophage colony-stimulating factor alleviates adverse consequences of embryo culture on fetal growth trajectory and placental morphogenesis. *Endocrinology*. 2005;146(5):2142-2153. DOI 10.1210/en.2004-1260.

Sousa S.M., Norman R.J. Metabolic syndrome, diet and exercise. *Best Pract. Res. Clin. Obstet. Gynaecol*. 2016;37:140-151. DOI 10.1016/j.bpobgyn.2016.01.006.

van Montfoort A.P.A., Hanssen L.L.P., de Sutter P., Viville S., Geraedts J.P.M., de Boer P. Assisted reproduction treatment and epigenetic inheritance. *Hum. Reprod. Update.* 2012;18(2):171-197. DOI 10.1093/humupd/dmr047.

Wilkinson M.J., Manoogian E.N.C., Zadourian A., Lo H., Fakhouri S., Shoghi A., Wang X., Fleischer J.G., Navlakha S., Panda S., Taub P.R. Ten-hour time-restricted eating reduces weight, blood pressure, and atherogenic lipids in patients with metabolic syndrome. *Cell Metab*. 2020;31(1):92-104.e5. DOI 10.1016/j.cmet.2019.11.004.

Wyns C., De Geyter C., Calhaz-Jorge C., Kupka M.S., Wyns C., Mocanu E., Motrenko T., Scaravelli G., Smeenk J., Vidakovic S., Goossens V. ART in Europe, 2018: results generated from European registries by ESHRE. *Hum. Reprod. Open*. 2022;2022(3):hoac022. DOI 10.1093/hropen/hoac022.

Yohannes S., Alebie G., Assefa L. Dermatoglyphics in type 2 diabetes with implications on gene linkage or early developmental noise: past perspectives. *Current Trends Future Prospects*. 2015;3(1D):297-305.

Zandstra H., Brentjens L.B.P.M., Spauwen B., Touwslager R.N.H., Bons J.A.P., Mulder A.L., Smits L.J.M., van der Hoeven M.A.H.B.M., van Golde R.J.T., Evers J.L.H., Dumoulin J.C.M., van Montfoort A.P.A. Association of culture medium with growth, weight and cardiovascular development of IVF children at the age of 9 years. *Hum. Reprod.* 2018;33(09):1645-1656. DOI 10.1093/humrep/dey246.

ORCID ID

M.V. Anisimova orcid.org/0000-0003-1522-9433
A.V. Romashchenko orcid.org/0000-0002-0674-0574
L.A. Gerlinskaya orcid.org/000-0002-8118-1362
M.P. Moshkin orcid.org/0000-0002-5388-2946

Acknowledgements. The study was supported by the Russian Science Foundation Grant No. 23-14-00179. The resources of the Center for Genetic Resources of Laboratory Animals, which is supported by budgetary projects (AAAA-A17-17072710029-7 and 0259-2019-0004), were used in this work.

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

Received May 16, 2023. Revised June 5, 2023. Accepted June 9, 2023.

Вавиловский журнал генетики и селекции. 2023;27(4):366-372
DOI 10.18699/VJGB-23-44

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

Genetic methods in honey bee breeding

M.D. Kaskinova ✉, A.M. Salikhova, L.R. Gaifullina, E.S. Saltykova

Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Center of the Russian Academy of Sciences, Ufa, Russia
✉ kaskinovamilyausha@mail.ru

Abstract. The honey bee *Apis mellifera* is a rather difficult object for selection due to the peculiarities of its biology. Breeding activities in beekeeping are aimed at obtaining bee colonies with high rates of economically useful traits, such as productivity, resistance to low temperatures and diseases, hygienic behavior, oviposition of the queen, etc. With two apiaries specializing in the breeding of *A. m. mellifera* and *A. m. carnica* as examples, the application of genetic methods in the selection of honey bees is considered. The first stage of the work was subspecies identification based on the analysis of the polymorphism of the intergenic mtDNA locus *tRNAleu-COII* (or *COI-COII*) and microsatellite nuclear DNA loci Ap243, 4a110, A24, A8, A43, A113, A88, Ap049, A28. This analysis confirmed that the studied colonies correspond to the declared subspecies. In the apiary with *A. m. mellifera*, hybrid colonies have been identified. A method based on the analysis of polymorphisms of the *tRNAleu-COII* locus and microsatellite nuclear DNA loci has been developed to identify the dark forest bee *A. m. mellifera* and does not allow one to differentiate subspecies from C (*A. m. carnica* and *A. m. ligustica*) and O (*A. m. caucasica*) evolutionary lineages from each other. The second stage was the assessment of the allelic diversity of the *csd* gene. In the apiary containing colonies of *A. m. mellifera* (*N* = 15), 20 *csd* alleles were identified. In the apiary containing colonies of *A. m. carnica* (*N* = 44), 41 alleles were identified. Six alleles are shared by both apiaries. DNA diagnostics of bee diseases showed that the studied colonies are healthy. Based on the data obtained, a scheme was developed for obtaining primary material for honey bee breeding, which can subsequently be subjected to selection according to economically useful traits. In addition, the annual assessment of the allelic diversity of the *csd* gene will shed light on the frequency of formation of new allelic variants and other issues related to the evolution of this gene.
Key words: *Apis mellifera*; *tRNAleu-COII* locus; microsatellites; *csd* gene; bee diseases.

For citation: Kaskinova M.D., Salikhova A.M., Gaifullina L.R., Saltykova E.S. Genetic methods in honey bee breeding. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):366-372. DOI 10.18699/VJGB-23-44

Генетические методы в селекции медоносной пчелы

М.Д. Каскинова ✉, А.М. Салихова, Л.Р. Гайфуллина, Е.С. Салтыкова

Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
✉ kaskinovamilyausha@mail.ru

Аннотация. Медоносная пчела *Apis mellifera* является довольно сложным объектом для селекции в силу особенностей ее биологии. Селекционные мероприятия в пчеловодстве нацелены на получение семей пчел с высокими показателями хозяйственно полезных признаков, таких как продуктивность, устойчивость к низким температурам и заболеваниям, гигиеническое поведение, яйценоскость матки и др. На примере двух пасек, специализирующихся на разведении *A. m. mellifera* и *A. m. carnica*, рассмотрено применение генетических методов в селекции медоносной пчелы. Первым этапом работы было установление подвидовой принадлежности на основе оценки полиморфизма межгенного локуса мтДНК *tRNAleu-COII* и микросателлитных локусов ядерной ДНК Ap243, 4a110, A24, A8, A43, A113, A88, Ap049 и A28. Подтверждено, что исследуемые семьи соответствуют заявленным подвидам. На пасеке с *A. m. mellifera* выявлены гибридные семьи. Метод, основанный на анализе полиморфизма локуса *tRNAleu-COII* (или *COI-COII*) и микросателлитных локусов ядерной ДНК, был разработан для идентификации темной лесной пчелы *A. m. mellifera* и не позволяет дифференцировать представителей эволюционных ветвей C (*A. m. carnica* и *A. m. ligustica*) и O (*A. m. caucasica*). На втором этапе оценивалось аллельное разнообразие гена *csd*. На пасеке, содержащей семьи *A. m. mellifera* (*N* = 15), выявлено 20 аллелей *csd*, на пасеке, содержащей семьи *A. m. carnica* (*N* = 44), – 41 аллель. Шесть аллелей были общими для двух пасек. ДНК-диагностика заболеваний пчелы показала, что исследуемые семьи являются здоровыми. По результатам исследований разработана схема получения первичного материала для селекции медоносной пчелы, на который впоследствии может быть наложен отбор по хозяйственно полезным признакам. Кроме того, ежегодная оценка аллельного разнообразия гена *csd* позволит пролить свет на частоту формирования новых аллельных вариантов и другие вопросы, связанные с эволюцией этого гена.
Ключевые слова: медоносная пчела; *Apis mellifera*; локус *tRNAleu-COII*; микросателлиты; ген *csd*; болезни пчел.

Introduction

In 1996, in the Laboratory of Biochemistry of Insect Adaptability of the Institute of Biochemistry and Genetics of the Ufa Scientific Center of the Russian Academy of Sciences, under the guidance of Professor Aleksey Gennadievich Nikolenko, studies of the honey bee were initiated (Nikonorov et al., 1998; Nikolenko, Poskryakov, 2002). This object was not chosen by chance. The most extensive habitat of the dark forest bee *Apis mellifera mellifera* has remained on the territory of the Republic of Bashkortostan (RB). This subspecies is native to the territory of Russia and belongs to the M evolutionary lineage.

In total, four evolutionary lineages for the honey bee were identified based on morphometric (Ruttner, 1988) and genetic (De La Rúa et al., 2009; Meixner et al., 2013; Cridland et al., 2017) data. These are evolutionary lineages M (bee subspecies of the western Mediterranean and northwestern Europe *A. m. mellifera* and *A. m. iberiensis*), A (subspecies from Africa *A. m. scutellata*, *A. m. sahariensis*, etc.), C (subspecies from southeastern Europe *A. m. ligustica*, *A. m. carnica*, etc.) and O (subspecies from the Middle East and Western Asia *A. m. caucasica*, *A. m. anatolica*, etc.). The import of bee colonies of other subspecies (particularly *A. m. caucasica* and *A. m. carnica*) has led to mass hybridization and the loss of the local population of honey bees in most of Russia. Therefore, the goal was to identify honey bee subspecies in order to preserve parts of scattered local populations of *A. m. mellifera*. The method developed by Garnery et al. (1998) was adopted. This method is based on the analysis of the polymorphism of the *tRNAleu-COII* intergenic mitochondrial DNA locus and makes it possible to determine the maternal line of bees. However, using this method, it was impossible to assess the influence of the drone background and identify hybrid colonies. Therefore, the search for new informative genetic markers was continued.

To establish the level of hybridization of colonies, the method proposed by Solignac et al. (2003), which is based on the analysis of polymorphism of microsatellite loci, was adapted. Out of 36 microsatellite loci, those that showed the greatest differentiating ability for reference samples of *A. m. mellifera*, *A. m. caucasica* and *A. m. carnica* (selected in their natural habitats in Russia) were selected (Il'yasov et al., 2007; Kaskinova et al., 2022). A set of 9 microsatellite loci made it possible to differentiate the subspecies *A. m. mellifera* from *A. m. caucasica* and *A. m. carnica*, but not the last two subspecies from each other. Using this method, surviving populations of *A. m. mellifera* were found on the territory of the RB and partially in other regions of Russia (Il'yasov et al., 2007; Kaskinova et al., 2022).

At the same time, work on DNA diagnostics of honey bee diseases was underway. In 2015, studies were launched to assess the allelic diversity of the *csd* (*complementary sex determiner*) gene of the honey bee (Kaskinova et al., 2019). These studies were initiated in response to the problem of irregular brood pattern (uneven distribution of brood cells in combs) faced by beekeepers that used instrumental insemination and closed type of bee breeding. An irregular brood pattern can be caused by brood diseases and inbreeding. The first reason was excluded by us based on the results of PCR diagnostics of bee diseases. To assess the effect of inbreeding, we analyzed the allelic diversity of the *csd* gene. The low allelic diversity of the *csd* gene results in a large number of diploid drones, which are killed by worker bees. This leads to such a phenomenon as genetic irregular brood pattern (Beye et al., 2003; Zareba et al., 2017; Mroczek et al., 2022). Genetic irregular brood pattern is more common in apiaries with a closed type of breeding (isolated apiary). We analyzed the allelic diversity of the *csd* gene in paternal apiaries and determined which colonies are best used for instrumental insemination of queens.

All these methods have found application in practical beekeeping and in combination can be used for selection of honey bees. Based on the results of more than 20 years of research, we have developed a scheme for obtaining primary material for honey bee breeding. This scheme includes the following items: (1) honey bee subspecies identification; (2) assessment of the allelic diversity of the *csd* gene; (3) assessment of the health status of the bee colony based on DNA diagnostics of bee diseases. On the example of two apiaries specializing in breeding *A. m. mellifera* and *A. m. carnica*, we will consider the application of genetic methods in honey bee breeding.

Materials and methods

Sample collection. Two breeding apiaries were selected for the study. At the first apiary (hereinafter, apiary No. 1) from the Iglinsky district of the RB, which specializes in breeding *A. m. mellifera*, worker bees and drones from 15 colonies were selected. In this apiary, the beekeeper uses instrumental insemination and paternal colonies of different origins. At the second apiary (No. 2) from the Chishminsky district of the RB, worker bees and drones belonging to the subspecies *A. m. carnica* were selected from 44 colonies. There is no breeding control in this apiary, but new queens are imported every year to maintain the genetic diversity of colonies. Worker bees (3 worker bees per colony) were used for subspecies identification, since they provide more complete information about the colony genotype. All bees were collected inside the hive from brood frames.

***Apis mellifera* subspecies identification.** DNA was isolated from the thorax muscles of worker bees using the DNA-Extran-2 kit (Syntol, Moscow). The analysis of the mtDNA *tRNAleu-COII* intergenic locus and 9 microsatellite loci (Ap243, 4a110, A24, A8, A43, A113, A88, Ap049, A28) was used for subspecies identification. Primer sequences are presented in Table 1. The PCR mixture for 10 samples with a total volume of 150 µl included 120 µl of distilled water, 15 µl of magnesium buffer, 3 µl of DNTP mixture (concentration 10 µmol each), 5 µl of F- and R-primer (concentration of 10 pmol/µl) and 3 µl of Taq polymerase (components of the PCR mixture produced by Sintol, Moscow). PCR mode: 3 min at 94 °C, then 30 cycles with denaturation for 30 s at 94 °C, annealing for 30 s at 49 °C (for the *tRNAleu-COII* locus) and 55 °C (for microsatellite loci), elongation for 60 s at 72 °C and final elongation for 3 min at 72 °C. PCR products were visualized using 8 % polyacrylamide gel electrophoresis (PAAG) followed by detection in a Gel Doc™ XR+ photosystem (BioRad, USA).

Samples of *A. m. mellifera* from the Burzyansky district of the RB and Perm Krai ($N = 136$) were used as a reference group of the evolutionary lineage M. Samples from the Republic of Adygea, Krasnodar Krai and Uzbekistan ($N = 120$) were used as representatives of the C/O evolutionary lineages.

Table 1. Primer sequences for *Apis mellifera* subspecies identification

No.	Locus	Primer sequence
1	*tRNAleu-COII*	F-GGCAGAATAAGTGCATTG
		R-CAATATCATTGATGACC
2	4a110	F-CGCTCGCGGTGGATTTCATTT
		R-GGCAAAAGTGGCGGAGAAAGA
3	A43	F-CACCGAAACAAGATGCAAG
		R-CCGCTCATTAAGATATCCG
4	A113	F-CTCGAATCGTGGCGTCC
		R-CCTGTATTTTGCAACCTCGC
5	A88	F-CGAATTAACCGATTTGTCG
		R-GATCGCAATTATTGAAGGAG
6	Ap049	F-CCAATAGCGGCGAGTGTG
		R-GGGCTTCGTACGTCCACC
7	A28	F-GAAGAGCGTTGGTTGCAGG
		R-GCCGTTCATGGTTACCACG
8	Ap243	F-AATGTCCGCGAGCATCTG
		R-TGTTTACGAGAATTCGACGGG
9	A24	F-CACAAGTTCCAACAATGC
		R-CACATTGAGGATGAGCG
10	A8	F-CGAAGGTAAGGTAAATGGAAC
		R-GGCGGTTAAAGTTCTGG

Cluster analysis was carried out using Structure 2.3.4 software. The Admixture model was used with Burnin Period and MCMC equal to 10,000 and 100,000, respectively. The number of clusters was set from 1 to 10. The estimated number of clusters was calculated using the Structure Harvester online service (Earl, von Holdt, 2012). The results obtained in Structure were processed in CLUMPP 1.1.2 using the FullSearch algorithm.

Assessment of allelic diversity of the *csd* gene. DNA was isolated from the thorax muscles of drones using the DNA-Extran-2 kit (Syntol, Moscow). Primers F-GGGAGAGAAG TTGCAGTAGAG and R-TTGATGCGTAGGTCCAAATCC flanking exons 6–8 of the *csd* gene were used (Kaskinova et al., 2019). Sequencing of PCR products was carried out at Syntol (Moscow) with forward and reverse primers used for PCR. The resulting nucleotide sequences were edited in Chromas v. 2.22 and aligned with ClustalW in MEGA v6.0. Alignment was performed on the reference sequence (NCBI Reference Sequence: NC_007072.3). Exons were identified using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The nucleotide sequences have been uploaded to Genbank and are available under numbers KY502199–KY502249 and MK531891–MK531969.

DNA diagnostics of honey bee diseases. We performed DNA diagnostics of the most common honey bee pathogens (Table 2): the fungus *Ascosphaera apis*, which causes ascospherosis (AscoA); microsporidium *Nosema/Vairimorpha apis* (NosA) and *Nosema/Vairimorpha cerana* (NosC), causing type A and C nosematosis, respectively; bacteria *Paenibacillus larvae* (AFB) and *Melissococcus plutonius* (EFB) causing foulbrood diseases. PCR diagnostics of bees for the presence of viral diseases was also performed: sacbrood virus (SBV), black queen cell virus (BQCV), wing deformation virus (DWV), Kashmir bee virus (KBV), acute bee paralysis virus (ABPV), chronic bee paralysis virus (CBPV) (see primer list in Table 2). DNA samples from diseased bees and synthetic DNA fragments that completely matched the expected PCR product were used as a positive control.

DNA isolation was performed from the midgut of worker bees (10 worker bees from each colony) using the DNA-Extran-2 kit. The PCR mixture for 10 samples with a total volume of 150 µl included 120 µl of distilled water, 15 µl of magnesium buffer, 3 µl of DNTP mixture (concentration 10 µmol each), 5 µl of F- and R-primer (concentration of 10 pmol/µl) and 3 µl of Taq polymerase. PCR mode: 5 min at 94 °C, then 30 cycles with 30 s denaturation at 94 °C, 30 s annealing at 50 °C, 60 s elongation at 72 °C and final elongation 7 min at 72 °C. Amplification products were visualized in 8 % PAAG.

To diagnose honey bee viruses, RNA was isolated from the thorax muscles of live bees (from one colony of 20 bees frozen in liquid nitrogen, taking into account two repetitions) using Trizol (Thermo FS). For cDNA synthesis, an RT-PCR kit (Syntol, Moscow) was used. The resulting cDNA was used for further PCR.

Results and discussion

Apis mellifera subspecies identification

Using the polymorphism analysis of the mtDNA *tRNAleu-COII* intergenic locus and 9 microsatellite loci, we analyzed the genetic structure of the studied samples.

Analysis of the mtDNA *tRNAleu-COII* intergenic locus is one of the simple and reliable methods for differentiating M and C/O evolutionary lineages. Allelic variants $P(Q)_{1–n}$ are markers of the origin of bees from *A. m. mellifera*, allelic variant Q – from subspecies from the evolutionary lineages C and O on the maternal line. P and Q are conventional designations for non-coding repeats located between the *tRNAleu* and *COII* genes. At the same time, the subspecies from the evolutionary branches C and O lack the P repeat (Bertrand et al., 2015). This analysis confirmed that the studied colonies correspond to the declared evolutionary lineages. In apiary No. 1, 14 out of 15 colonies had a PQQ allelic variant, and in one colony an allelic variant PQQQ was detected. All colonies from apiary No. 2 had the allelic variant Q.

Figure 1 shows the results of cluster analysis. Analysis of Structure output data in Structure Harvester showed that the total study sample consists of two main clusters (K = 2, delta K = 2554.6). The first cluster is represented by a dark forest bee; the second cluster includes subspecies from the evolutionary lineages C (*A. m. carnica*) and O (*A. m. caucasica*). The colonies from apiary No. 1 entered the same

Table 2. Primers for DNA diagnostics of honey bee diseases

Primer	Sequence, 5′→3′	Size, b.p.	Reference
F-AFB	GCAAGTCGAGCGGACCTTGT	237	Bakonyi et al., 2003
R-AFB	GCATCGTCGCCTTGGTAAGC		
F-EFB	GAAGAGGAGTTAAAAGGCGC	831	Govan et al., 1998
R-EFB	TTATCTCTAAGGCGTTCAAAGG		
F-NosA	GCATGTCTTTGACGTACTATGTA	312	Espregueira Themudo et al., 2020
R-NosA	CGTTTAAAATGTGAAACAACTATG		
F-NosC	CGACGATGTGATATGAAAATATTAA	218	
R-NosC	TCATTCTCAAACAAAAAACCG		
F-AscoA	GCACTCCCACCCTTGTCTA	550	James, Skinner, 2005
R-AscoA	CCCACTAGAAGTAAATGATGGTTA		
F-SBV	ACCAACCGATTCCTCAGTAG	468–469	Berényi et al., 2006
R-SBV	CCTTGGAACTCTGCTGTGTA		
F-BQCV	AGTAGTTGCGATGTACTTCC	471	
R-BQCV	CTTAGTCTTACTCGCCACTT		
F-DWV	ATTGTGCAAGATTGGACTAC	433	
R-DWV	AGATGCAATGGAGGATACAG		
F-KBV	GATGAACGTCGACCTATTGA	414	
R-KBV	TGTGGGTTGGCTATGAGTCA		
F-ABPV	GTGCTATCTTGGAATACTAC	618	
R-ABPV	AAGGYTTAGGTTCTACTACT		
F-CBPV	TGTCGAACTGAGGATCTTAC	315	
R-CBPV	GACCTGATTAACGACGTTAG		

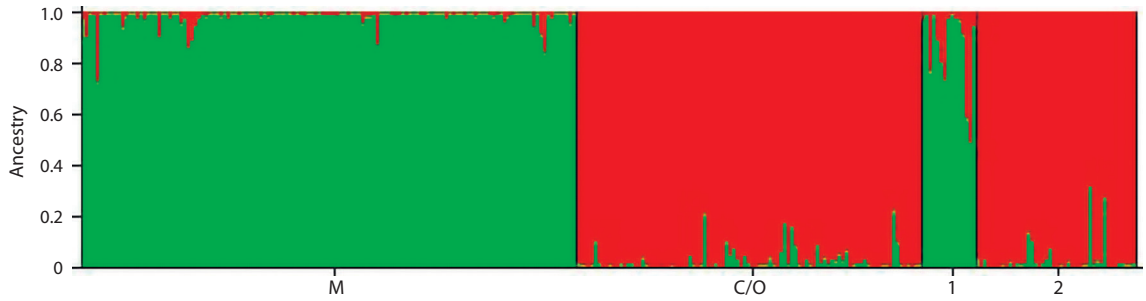


Fig. 1. Genetic structure of the studied samples at K=2, where M is a reference sample from the M evolutionary lineage; C/O is a reference sample from the C/O evolutionary lineages; 1 is a sample from apiary No. 1; 2 is a sample from apiary No. 2.

cluster with bees from the reference sample M (i. e., subspecies *A. m. mellifera* from the evolutionary lineage M). At the same time, introgression of the C/O gene pool is noted in two colonies. Colonies from apiary No. 2 form a common cluster with a C/O reference sample.

This method was developed for the identification of the dark forest bee *A. m. mellifera* and does not allow one to differentiate representatives of the C and O evolutionary lineages from each other. Therefore, we can say that the colonies in apiary No. 1 belong to the subspecies *A. m. mellifera*. Some of the colonies turned out to be hybrid and it was recommended to take measures to replace the queens. In apiary No. 2, hybrid colonies were also identified (with the share of cluster M > 15 %).

The problem of subspecies identification of the honey bee arose in connection with the uncontrolled transportation of bee colonies and the development of package beekeeping. If the differentiation of subspecies from the M and C/O evolutionary lineages is a resolved issue and is used to search for native populations (Nikolenko, Poskryakov, 2002; Il'yasov et al., 2007), then the differentiation of subspecies belonging to the same evolutionary lineage remains relevant.

Table 3. Summary table of *csd* alleles in the studied samples

Apiary	Number of colonies	Number of *csd* alleles	Number of rare *csd* alleles	Genbank ID
No. 1	15	20	13	KY502199–KY502248
No. 2	44	41	22	KY502249, MK531891–MK531969

Assessment of allelic diversity of the *csd* gene

Determining the allelic diversity of a sex-determining gene is one of the main tasks of a modern bee breeding program (Hyink et al., 2013). Keeping honey bees under closed breeding conditions results in the accumulation of a large number of common *csd* alleles, and consequently leads to the production of a large number of diploid drones. Diploid drones are killed by worker bees and this leads to a decrease in the strength of the colony.

In apiary No. 1, 20 *csd* alleles described by us earlier (Kaskinova et al., 2019) were detected, in apiary No. 2 – 41 alleles. Six alleles are shared by two apiaries. Almost half of the alleles are rare, i. e. meet once (Table 3).

In apiary No. 1, common alleles were identified in seven colonies (in colonies 2 and 7; 5 and 9; 5, 8 and 14, see Suppl. Material 1)[1]. The remaining alleles are rare. To avoid the appearance of diploid drones, it was recommended not to use together the sperm of drones from colonies with the same *csd* alleles during artificial insemination of queens. In apiary No. 2, almost half of the *csd* alleles are rare. Allele 3 *csd* (Suppl. Material 2) was found in 9 colonies, allele 6 – in 8 colonies. Common alleles have been identified in sister colonies descending from one queen.

Thus, we have established that nine colonies from apiary No. 1 belong to the *A. m. mellifera* subspecies and have a high allelic diversity of the *csd* gene, and therefore are recommended for further use as paternal colonies (see Suppl. Material 1). In apiary No. 2, in contrast to apiary No. 1, artificial insemination of queens is not used. Since the colonies of this apiary have a homogeneous genetic structure, colonies with the same alleles can be used as finisher colonies or a queen replacement can be carried out in them (see Suppl. Material 2). It is also possible to use these colonies as producers of honey products, preventing drones from breeding by installing drone cages.

Assessment of the health of bee colonies

PCR analysis showed that the studied colonies are healthy. Unfortunately, such healthy apiaries are the exception rather than the rule. In recent times, there has been a sad trend of a large number of apiaries suffering from nosematosis (our data, unpublished). The honey bee *A. mellifera* has been targeted by the dangerous invasive microsporidia species *Nosema/Vairimorpha ceranae* (Martín-Hernández et al., 2007; Tokarev et al., 2020), whose original host is the Asian bee *Apis cerana*. If earlier it was very rare for us to detect spores and DNA of this pathogen, now it is detected in almost every second apiary. This disease has no pronounced symptoms, and beekeepers seek help only when colonies begin to die or have already died. Over the entire period of the laboratory's work, we have identified colonies affected by nosematosis, ascospherosis, European foulbrood, as well as such viral diseases as sacbrood virus (SBV), wing deformation virus (DWV) and black queen cell virus (BQCV). As a rule, most of these colonies were imported from other countries and belong to the evolutionary branches C or O.

Scheme for obtaining primary material for bee breeding

Honey bee breeding is aimed at obtaining bee colonies with high rates of economic useful traits (EUT), such as productivity, resistance to low temperatures and diseases, hygienic behavior, queen egg production, etc. (Ruttner, 1988). The honey bee is a rather difficult object for selection due to the peculiarities of its biology. Complementary sex determination, in which sex depends on the allelic combination of the *csd* gene, enhances the consequences of inbreeding, and queen polyandry complicates the selection of parental pairs and reproduction control. In addition, the selection is carried out not at the individual level, but at the level of the bee colony. In this regard, there was a need to develop a method that would combine the assessment of EUT and the genetic potential of bee colonies. The assessment of EUT is entirely up to the beekeepers, while the result of our study was the method of selecting primary material for selection for EUT.

So-called 'purebred breeding' is a fundamental principle of bee breeding, according to Ruttner (1988). Based on this principle, we have developed a scheme for obtaining primary material for honey bee selection.

The first stage of selection work is the selection of colonies from the initial population of bees belonging to one subspecies. In hybrid colonies, it is necessary to carry out measures to replace queens. In addition, selected colonies are encouraged to be checked for diseases.

The second stage is the selection of colonies by EUT. Selected colonies are recommended to be placed in an apiary remote from other colonies from the original population in order to avoid unwanted crossbreeding. It is recommended to create several such apiaries for their further crossing. This step is optional if the goal is to restore the *A. m. mellifera* population.

At the third stage, the analysis of the allelic diversity of the *csd* gene in the selected colonies is carried out. Based on the data on the allelic composition of the *csd* gene, paternal and maternal colonies are formed. Maternal and paternal colonies should have different *csd* alleles. As maternal it is recommended to use colonies that have frequently occurring *csd* alleles. Paternal colonies are best formed from among those colonies that have rare *csd* alleles.

At the fourth stage, the offspring are evaluated for EUT. Selection for the selected trait(s) must be carried out in each generation.

[1] Supplementary Materials 1 and 2 are available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Kaskinova_Engl_27_4.pdf

Such a selection scheme (Fig. 2) will make it possible to obtain lines of bees with high productivity rates. In addition, the annual assessment of the allelic diversity of the *csd* gene in one apiary will shed light on the frequency of formation of new allelic variants and other issues related to the evolution of this gene.

Analysis of the polymorphism of the sex-determining gene in the studied samples of *A. mellifera* generally showed that their allelic diversity corresponds to previously obtained data for other populations (Hyink et al., 2013; Zareba et al., 2017). Most bee colonies living in a breeding apiary have many common *csd* alleles. Over time, this can lead to large numbers of diploid drones. Therefore, it is recommended to take measures to prevent the use of drones with the same *csd* alleles. Moreover, it is more important that the *csd* alleles of the drones do not coincide with the alleles of the queen.

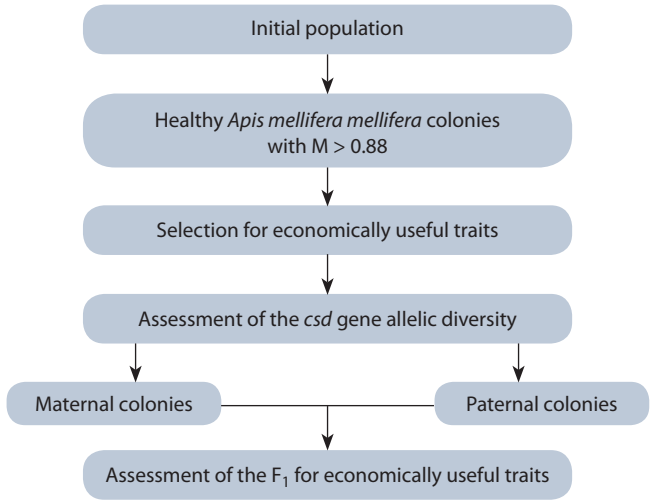


Fig. 2. Selection scheme for the *Apis mellifera mellifera* using information on subspecies and allelic diversity of the *csd* gene.

Conclusion

The purpose of this work was to present the results of more than 20 years of work aimed at finding and preserving the native population of the dark forest bee *A. m. mellifera*. During this time, methods have been developed for the subspecies identification of *A. m. mellifera*, DNA diagnostics of diseases and assessment of the allelic diversity of the *csd* gene. These methods were tested in two apiaries using different breeding subspecies and different methods of breeding control. With the help of genetic methods, it became possible to obtain primary material for selection of honey bees, which can subsequently be subjected to selection for economically useful traits.

The search for surviving native (or local) honey bee populations is only the first step in their conservation. Further actions are completely dependent on the beekeepers – whether they will use the data received or leave everything as it is. We received positive feedback from beekeepers from the Yanaulsky, Burzyansky and Iglinsky districts of the RB, as well as from beekeepers in Altai krai, Belgorodskaya oblast and some other regions of the Russian Federation. Thanks to the initiative of beekeepers, it was possible to improve the situation even in the moderately hybrid population of *A. m. mellifera* in the Iglinsky district of the RB (Kaskinova et al., 2022).

Without the interest of beekeepers in the conservation of native subspecies (be it *A. m. mellifera*, *A. m. carnica*, *A. m. caucasica* or any other subspecies), the results of monitoring the subspecies of bees will be important only for fundamental science. Thus, based on the geographical distribution of monitoring data for the Burzyan population of *A. m. mellifera*, the central, peripheral, and hybrid zones of the range were identified, as well as the main directions of introgression of subspecies from C/O evolutionary lineages (Nikolenko et al., 2010). The subsequent analysis of bees throughout the territory of the RB showed that a similar subdivision of the honey bee population is also observed in other surviving populations of *A. m. mellifera* (Tatyshlinsky and Yanaulsky populations).

The honey bee is a very valuable and fragile object of our ecosystem. Its numbers are decreasing all over the world due to the uncontrolled use of pesticides and insecticides, hybridization, and the spread of diseases (Neumann, Carreck, 2010; Espregueira Themudo et al., 2020). Human intervention in the process of its natural settlement has practically nullified the adaptive potential that was developed over millions of years of evolution. Therefore, in our opinion, it is necessary to preserve and, if possible, restore those populations of honey bees that still remain.

Dedicated to the blessed memory of our colleagues Nikolenko Alexei Gennadievich and Poskryakov Alexander Vitalievich.

References

Bakonyi T., Derakhshifar I., Grabensteiner E., Nowotny N. Development and evaluation of PCR assays for the detection of *Paenibacillus larvae* in honey samples: comparison with isolation and biochemical characterization. *Appl. Environ. Microbiol*. 2003;69(3):1504-1510. DOI 10.1128/AEM.69.3.1504-1510.2003.

Berényi O., Bakonyi T., Derakhshifar I., Köglberger H., Nowotny N. Occurrence of six honeybee viruses in diseased Austrian apiaries. *Appl. Environ. Microbiol*. 2006;72(4):2414-2420. DOI 10.1128/AEM.72.4.2414-2420.2006.

Bertrand B., Alburaki M., Legout H., Moulin S., Mougel F., Garnery L. MtDNA COI-COII marker and drone congregation area: an efficient method to establish and monitor honeybee (*Apis mellifera* L.) conservation centers. *Mol. Ecol. Resour*. 2015;15(3):673-683. DOI 10.1111/1755-0998.12339.

Beye M., Hasselmann M., Fondrk M., Page R.E., Omholt S.W. The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell*. 2003;114(4):419-429. DOI 10.1016/S0092-8674(03)00606-8.

Cridland J.M., Tsutsui N.D., Ramírez S.R. The complex demographic history and evolutionary origin of the western honey bee, *Apis mellifera*. *Genome Biol. Evol*. 2017;9(2):457-472. DOI 10.1093/gbe/evx009.

De La Rúa P., Jaffé R., Dall'olio R., Muñoz I., Serrano J. Biodiversity, conservation and current threats to European honeybees. *Apidologie*. 2009;40:263-284. DOI 10.1051/apido/2009027.

Earl D.A., von Holdt B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour*. 2012;4(2):359-361. DOI 10.1007/s12686-011-9548-7.

Espregueira Themudo G., Rey-Iglesia A., Robles Tascon L., Jensen A.B., da Fonseca R.R., Campos P.F. Declining genetic diversity of European honeybees along the twentieth century. *Sci. Rep*. 2020;10(1):10520. DOI 10.1038/s41598-020-67370-2.

Garnery L., Franck P., Baudry E., Vautrin D., Cornuet J.-M., Solignac M. Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA. *Genet. Sel. Evol*. 1998;30(Suppl.1):S31. DOI 10.1186/1297-9686-30-S1-S31.

Govan V.A., Brozel V., Allsopp M.H., Davison S. A PCR detection method for rapid identification of *Melissococcus pluton* in honeybee larvae. *Appl. Environ. Microb*. 1998;64(5):1983-1985. DOI 10.1128/AEM.64.5.1983-1985.1998.

Hyink O., Laas F., Dearden P. Genetic tests for alleles of *complementary-sex-determiner* to support honeybee breeding programmes. *Apidologie*. 2013;44(3):306-313. DOI 10.1007/s13592-012-0181-6.

Il'yasov R.A., Petukhov A.V., Poskryakov A.V., Nikolenko A.G. Local honeybee (*Apis mellifera mellifera* L.) populations in the Urals. *Russ. J. Genet*. 2007;43(6):709-711. DOI 10.1134/S1022795407060166.

James R.R., Skinner J.S. PCR diagnostic methods for *Ascosphaera* infections in bees. *J. Invertebr. Pathol.* 2005;90(2):98-103. DOI 10.1016/j.jip.2005.08.004.

Kaskinova M.D., Gaifullina L.R., Saltykova E.S., Poskryakov A.V., Nikolenko A.G. Dynamics of the genetic structure of *Apis mellifera* populations in the Southern Urals. *Russ. J. Genet*. 2022;58(1):36-41. DOI 10.1134/S1022795422010045.

Kaskinova M.D., Gataullin A.R., Saltykova E.S., Gaifullina L.R., Poskryakov A.V., Nikolenko A.G. Polymorphism of the hypervariable region of the *csd* gene in the *Apis mellifera* L. population in Southern Urals. *Russ. J. Genet.* 2019;55(2):267-270. DOI 10.1134/S102279541902008X.

Martín-Hernández R., Meana A., Prieto L., Salvador A.M., Garrido-Bailón E., Higes M. Outcome of colonization of *Apis mellifera* by *Nosema ceranae. Appl. Environ. Microbiol*. 2007;73(20):6331-6338. DOI 10.1128/AEM.00270-07.

Meixner M.D., Pinto M.A., Bouga M., Kryger P., Ivanova E., Fuchs S. Standard methods for characterising subspecies and ecotypes of *Apis mellifera*. *J. Apic. Res.* 2013;52(4):1-28. DOI 10.3896/IBRA.1.52.4.05.

Mroczek R., Laszkiewicz A., Blazej P., Adamczyk-Weglarzy K., Niedbalska-Tarnowska J., Cebrat M. New insights into the criteria of functional heterozygosity of the *Apis mellifera* complementary sex determining gene – discovery of a functional allele pair differing by a single amino acid. *PLoS One*. 2022;17(8):e0271922. DOI 10.1371/journal.pone.0271922.

Neumann P., Carreck N.L. Honey bee colony losses. *J. Apic. Res.* 2010;49(1):1-6. DOI 10.3896/IBRA.1.49.1.01.

Nikolenko A.G., Fakhretdinova S.A., Ilyasov R.A., Poskryakov A.V. The geographic range of the Burzyan population of the European dark honeybee *Apis mellifera mellifera* L. (Hymenoptera: Apidae). *Trudy Russkogo Entomologicheskogo Obshchestva = Proceedings of the Russian Entomological Society*. 2010;81(2):202-208. (in Russian)

Nikolenko A.G., Poskryakov A.V. Polymorphism of locus *COI-COII* of mitochondrial DNA in the honey bee *Apis mellifera* L. from the Southern Ural region. *Russ. J. Genet*. 2002;38(4):364-368. DOI 10.1023/A:1015289900666.

Nikonorov I.M., Ben'kovskaya G.V., Poskryakov A.V., Nikolenko A.G., Vakhitov V.A. The use of the PCR technique for control of pure-breeding of honeybee (*Apis mellifera mellifera* L.) colonies from the Southern Urals. *Russ. J. Genet*. 1998;34(11):1344-1347.

Ruttner F. Breeding Techniques and Selection for Breeding of the Honeybee: an Introduction to the Rearing of Queens, the Conduct of Selection Procedures and the Operation of Mating Stations. Condor, Derby: British Isles Bee Breeders Assoc., 1988.

Solignac M., Vautrin D., Loiseau A., Mougel F., Baudry E., Estoup A., Garnery L., Haberl M., Cornuet J.-M. Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera* L.) genome. *Mol. Ecol. Notes*. 2003;3(2):307-311. DOI 10.1046/j.1471-8286.2003.00436.x.

Tokarev Y.S., Huang W.F., Solter L.F., Malysh J.M., Becnel J.J., Vossbrinck C.R. A formal redefinition of the genera *Nosema* and *Vairimorpha* (Microsporidia: Nosematidae) and reassignment of species based on molecular phylogenetics. *J. Invertebr. Pathol*. 2020;169:107279. DOI 10.1016/j.jip.2019.107279.

Zareba J., Blazej P., Laszkiewicz A., Sniezewski L., Majkowski M., Janik S., Cebrat M. Uneven distribution of complementary sex determiner (*csd*) alleles in *Apis mellifera* population. *Sci. Rep*. 2017;7:2317. DOI 10.1038/s41598-017-02629-9.

ORCID ID

M.D. Kaskinova orcid.org/0000-0003-4960-6559
L.R. Gaifullina orcid.org/0000-0002-3285-118X
E.S. Saltykova orcid.org/0000-0003-0123-7037

Acknowledgements. The study was supported by the grant of the Russian Science Foundation No. 22-74-00004 (https://rscf.ru/project/22-74-00004/) using the resources of the Center for Collective Use of the UFRC RAS.

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

Received October 24, 2022. Revised November 1, 2022. Accepted November 1, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):373-382
DOI 10.18699/VJGB-23-45

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

The role of Beringia in human adaptation to Arctic conditions based on results of genomic studies of modern and ancient populations

B.A. Malyarchuk[1, 2]

[1] Institute of Biological Problems of the North, Far-East Branch of the Russian Academy of Sciences, Magadan, Russia
[2] N.A. Shilo North-East Interdisciplinary Scientific Research Institute, Far-East Branch of the Russian Academy of Sciences, Magadan, Russia
malyarchuk@ibpn.ru

Abstract. The results of studies in Quaternary geology, archeology, paleoanthropology and human genetics demonstrate that the ancestors of Native Americans arrived in mid-latitude North America mainly along the Pacific Northwest Coast, but had previously inhabited the Arctic and during the last glacial maximum were in a refugium in Beringia, a land bridge connecting Eurasia and North America. The gene pool of Native Americans is represented by unique haplogroups of mitochondrial DNA and the Y chromosome, the evolutionary age of which ranges from 13 to 22 thousand years. The results of a paleogenomic analysis also show that during the last glacial maximum Beringia was populated by human groups that had arisen as a result of interaction between the most ancient Upper Paleolithic populations of Northern Eurasia and newcomer groups from East Asia. Approximately 20 thousand years ago the Beringian populations began to form, and the duration of their existence in relative isolation is estimated at about 5 thousand years. Thus, the adaptation of the Beringians to the Arctic conditions could have taken several millennia. The adaptation of Amerindian ancestors to high latitudes and cold climates is supported by genomic data showing that adaptive genetic variants in Native Americans are associated with various metabolic pathways: melanin production processes in the skin, hair and eyes, the functioning of the cardiovascular system, energy metabolism and immune response characteristics. Meanwhile, the analysis of the existing hypotheses about the selection of some genetic variants in the Beringian ancestors of the Amerindians in connection with adaptation to the Arctic conditions (for example, in the *FADS*, *ACTN3*, *EDAR* genes) shows the ambiguity of the testing results, which may be due to the loss of some traces of the "Beringian" adaptation in the gene pools of modern Native Americans. The most optimal strategy for further research seems to be the search for adaptive variants using the analysis of paleogenomic data from the territory of Beringia, but such genetic data are still very scarce.
Key words: genomics; paleogenomics; mitochondrial DNA; Y chromosome; adaptation; Beringia; peopling of America.

For citation: Malyarchuk B.A. The role of Beringia in human adaptation to Arctic conditions based on results of genomic studies of modern and ancient populations. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):373-382. DOI 10.18699/VJGB-23-45

Роль Берингии в адаптации человека к условиям Арктики по результатам геномных исследований современного и древнего населения

Б.А. Малярчук[1, 2]

[1] Институт биологических проблем Севера Дальневосточного отделения Российской академии наук, Магадан, Россия
[2] Северо-Восточный комплексный научно-исследовательский институт им. Н.А. Шило Дальневосточного отделения Российской академии наук, Магадан, Россия
malyarchuk@ibpn.ru

Аннотация. Результаты исследований в области четвертичной геологии, археологии, палеоантропологии и генетики человека показывают, что предки американских индейцев прибыли в средние широты Северной Америки главным образом по тихоокеанскому побережью. Но до этого они населяли Арктику и во время последнего ледникового максимума укрывались в рефугиуме на территории Берингии, соединявшей Евразию и Северную Америку. Генофонд американских индейцев представлен уникальными гаплогруппами митохондриальной ДНК и Y-хромосомы, эволюционный возраст которых составляет от 13 до 22 тыс. лет. Результаты анализа палеогеномных данных также свидетельствуют о том, что во время последнего ледникового максимума на территории Берингии сохранялись группы населения, возникшие в результате взаимодействия древнейшего верхнепалеолитического населения Северной Евразии и пришлых групп населения

из Восточной Азии. Примерно 20 тыс. лет назад началось формирование популяций Берингии, а время их существования в относительной изоляции оценивается примерно в 5 тыс. лет. Таким образом, период адаптации берингийцев к условиям Арктики мог занять несколько тысячелетий. Адаптация предков американдов к высоким широтам и холодному климату подтверждается геномными данными, показавшими, что адаптивные генетические варианты у американских индейцев ассоциируются с различными метаболическими путями: процессами продукции меланина в коже, волосах и глазах, функционированием сердечно-сосудистой системы, энергетическим обменом и особенностями иммунного ответа. Между тем анализ существующих гипотез об отборе некоторых генетических вариантов у берингийских предков американдов в связи с адаптацией к условиям Арктики (например, в генах *FADS*, *ACTN3*, *EDAR*) демонстрирует неоднозначность результатов проверки, что может быть связано с утратой некоторых следов «берингийской» адаптации в генофондах современных американских индейцев. Наиболее оптимальной стратегией дальнейших исследований представляется поиск адаптивных вариантов с помощью анализа палеогеномных данных с территории Берингии, однако таких генетических данных пока очень мало.

Ключевые слова: геномика; палеогеномика; митохондриальная ДНК; Y-хромосома; адаптация; Берингия; заселение Америки.

Introduction

Interdisciplinary research is important in examining issues related to the role of Beringia as an Arctic land bridge in the settlement of the Western Hemisphere. Unfortunately, the flooding of the lowlands of Beringia by rising sea levels 12,000 years ago led to the loss of a large number of archaeological sites, including possible human settlements. As a result, the main hypotheses associated with the history of the peopling of the Americas are based primarily on a synthesis of archaeological, geological, and biological data from the more accessible regions of Asia and the Americas (Potter et al., 2018; Batchelor et al., 2019; Waters, 2019). According to one model of American settlement, ancient people from Northeast Asia migrated to the Americas via Beringia along an ice-free corridor in what is now Western Canada, which was created by melting glaciers at the end of the Pleistocene. According to another scenario, Native American ancestors actively developed the coastal zone and migrated to northwest North America along the southern coast of Beringia (Waters, 2019; Willerslev, Meltzer, 2021).

However, both of these scenarios appear to underestimate the more complex role that Beringia played in the settlement of the Americas by human populations (Potter et al., 2018). Results of new geochronological studies suggest that the "Canadian Corridor" was blocked by ice during presumed human migrations ~15–16 thousand years ago and was only freed ~13.8 thousand years ago (Clark et al., 2022). Therefore, it seems more likely that the migration route is the northwest coast of America (Lesnek et al., 2018). Since the melting of the glaciers in these territories could have occurred as early as 18,000 years ago, the coastal migration route could have been developed by the first Americans ~17,000 years ago (Lesnek et al., 2018).

At the same time, the results of genomic studies indicate that the ancestors of the Native Americans appear to have lived in the Beringian Arctic for quite a long time and survived the last glacial maximum there, which terminated ~19,000 years ago (Tamm et al., 2007; Amorim et al., 2017; Moreno-Mayar et al., 2018a; Niedbalski, Long, 2022). It is assumed that these populations adapted to the harsh conditions of the Far North while living in Beringia, and as the climate warmed, Beringian descendants migrated to Northwest America.

Population genomics of indigenous populations of Siberia: genetic origins of Native Americans

Due to the rapid development of DNA sequencing techniques, the genomic approach is widely used in modern research in archaeology and paleoanthropology. Genetic studies of the indigenous populations of Siberia have shown that the most ancient genetic components, which were present as early as 30–40 thousand years ago in the most ancient inhabitants of both southern and northern Siberia, were later lost as a result of genetic drift and population replacement periods. Thus, genomic analysis of the ancient populations of Siberia (in the time interval from 31 to 0.6 thousand years ago) has shown that in Siberia there were several periods of almost complete population replacements (Sikora et al., 2019). The most ancient Late Paleolithic population of Siberia was replaced 20–11 thousand years ago by newcomers from East Asia, resulting in the formation of Siberian populations that gave rise to the ancestors of both Paleosiberians and Native Americans. In the time interval from 11 to 4 thousand years ago, this population was partially replaced by other groups of East Asian origin, which gave rise to the modern populations of Siberia. Additional gene flows in both directions along the Bering Strait were also reported at this time. These flows affected the genetic structure of Eskimo and Na-Dene populations (Flegontov et al., 2019; Sikora et al., 2019).

The results of paleogenomic studies have shown that the formation of ancestral populations to ancient peoples of northern Siberia and America could have occurred in a wide area from the Trans-Baikal region to Beringia (Sikora et al., 2019; Yu et al., 2020; Kilinç et al., 2021). Paleogenomic data indicate that during the Holocene the populations of Northeast Asia (from the Cis-Baikal region to modern Yakutia) were very dynamic in contrast to the Trans-Baikal populations, which demonstrate genetic continuity from the Mesolithic to the Bronze Age (Kilinç et al., 2021). Evidence was also obtained that bearers of the Belkachin archaeological culture in Yakutia can be considered ancestors of the Paleo-Eskimos, who migrated to the Americas about 5–6 thousand years ago (Kilinç et al., 2021).

Population studies of polymorphism of mitochondrial DNA (mtDNA) and Y-chromosome, which are inherited through the maternal and paternal lineages, respectively, have developed

greatly over the past 40 years. The results of molecular dating of the evolutionary age of the mtDNA and Y-chromosome haplogroups distributed among the modern indigenous populations of Siberia demonstrate that the most ancient genetic lineages in the south of Siberia belong to the postglacial period (younger than 21,000 years), while in the northeast of Siberia their evolutionary age does not exceed 10,000 years (Derenko et al., 2010, 2014; Malyarchuk et al., 2011; Duggan et al., 2013).

Genetic data for the present-day populations of Siberia are thus inconsistent with archaeological and paleoanthropological data showing the presence of *Homo sapiens* both in the south and in the north of Siberia as far back as 40–45 thousand years ago (Vasil'ev et al., 2002; Pitulko et al., 2004), but suggest that due to genetic drift, the most ancient genetic lineages were lost in the modern Siberian population (Derenko et al., 2010, 2014; Raghavan et al., 2014). This assumption was completely confirmed by the results of the paleogenomic study of the Siberians, which showed, as noted above, that in Siberia there have been several periods of almost complete replacements of the earlier inhabitants (Sikora et al., 2019). The genetic data obtained thus do not contradict the hypothesis that, during the last glacial maximum, the Arctic part of Asia and America (in particular, Beringia) could have been inhabited by humans.

Based on the mtDNA variability data, it has been established that the divergence between the ancestors of the Siberian and Native American populations occurred ~25,000 years ago, the evolutionary age of Native American founding mtDNA lineages is approximately 18.5 thousand years, and their spread across Americas began ~16,000 years ago (Llamas et al., 2016). The mitochondrial gene pool of Native Americans is represented by only seven mtDNA haplogroups (A2, B2, C1, D1, D4h3a, C4c, and X2a), the evolutionary age of which ranges from 13 to 20 thousand years ago (Perego et al., 2009; Hooshiar Kashani et al., 2012). All of the Native American mtDNA haplogroups are unique, and for six of them (except D4h3a) there are still no analogues found anywhere else in the world, which means that these haplogroups were formed in deep isolation and, apparently, while living in Beringia.

For the haplogroup D4h3a, a related D4h3b-haplotype was found in modern Chinese, which diverged from the common D4h3 ancestor by ~18.3 thousand years, and the evolutionary age of D4h3a itself is ~13 thousand years (Behar et al., 2012). It is also noteworthy that in all cases (except X2a) the Native American mtDNA haplogroups are derived from East Asian ancestors within haplogroups A, B, C, and D, predominantly distributed in East Asia. Meanwhile, the area of haplogroup X is located in Western Eurasia, and therefore the appearance of haplogroup X2a in the Beringians (and thus in the ancestors of the Amerindians) could have happened on the basis of the gene pool of their Upper Paleolithic ancestors, related to the population of Western Eurasia. This follows from paleogenomic data showing that the Upper Paleolithic populations of Siberia and Eastern Europe (in the time range from 18 to 45.5 thousand years ago) were characterized by approximately the same set of mtDNA haplogroups of West Eurasian origin (e. g., haplogroups N*, U*, U2, U8c, R1b) (Raghavan et al., 2014; Sikora et al., 2017, 2019; Yu et al., 2020). Thus, it is not excluded that the mtDNA X-haplotypes, which emerged on the basis of haplogroup X2a, were also common in the gene pool of this ancient population. The age of haplogroup X2a is approximately 13,000 years (Behar et al., 2012), and its range is limited to northwestern North America only (Hooshiar Kashani et al., 2012).

Haplogroup C4c has a similar geographic distribution and evolutionary age (~14 thousand years) (Hooshiar Kashani et al., 2012; Achilli et al., 2013). This appears to indicate that the carriers of haplogroups X2a and C4c were among the last Beringians who populated North America. Around the same time, but along the Pacific coast, haplogroup D4h3a presumably spread to the south of America (Perego et al., 2009). Thus, the results of mtDNA polymorphism studies indicate that both pathways (both the "Canadian Corridor" and the Pacific Coast of America) may have been used in the settlement of the Americas.

The results of studies of Y-chromosome variability also showed that the male gene pool of Native Americans is characterized by a small set of genetic lineages – these are haplogroups Q-M3, Q-M848 and Q-Z780, widespread among Native Americans, as well as the more rare haplogroups C-P39 and Q-Y4276, typical of Native North Americans, and haplogroup C-MPB373, found in Native South Americans (Pinotti et al., 2019; Colombo et al., 2022). The evolutionary age of these Y-haplogroups ranges from 15.0 to 21.6 thousand years, indicating that they originated in Beringia. This hypothesis is supported by the discovery of the rare haplogroup Q-L804, related to the Native American haplogroup Q-M3, in populations of Northwestern Europe (Norwegians, Swedes, and Britons). Since the divergence time between these haplogroups is ~17.3 thousand years, it is assumed that migrations from Beringia took place not only in the direction of America, but also in the opposite direction to Europe (Pinotti et al., 2019).

In addition, it should be noted that according to mtDNA studies, very rare haplogroups C1e and C1f, related to Native American C1 haplogroups (C1b, C1c, and C1d), were found in Icelanders as well as in the Mesolithic population of Karelia (Yuzhnyi Oleniy Island) (Ebenesersdóttir et al., 2011; Der Sarkissian et al., 2014). Thus, these data allow us to believe that in Beringia there was not a brief stop of human populations that migrated from the south, but indeed the Beringian populations, with their own genetic characteristics and unique mtDNA and Y-chromosome haplogroups, were formed there over a long time.

The question about the duration of the Beringian Standstill is also of great interest and is regularly studied. One of the first genetic studies of this problem concluded that the Beringian Standstill can be estimated at ~15,000 years, if we start from the first discoveries of *Homo sapiens* remains in the Asian part of the Arctic ~30,000 years ago and until the moment when the Beringian expansion to America took place, i. e. ~15,000 years ago according to genetic data (Tamm et al., 2007). M. Raghavan et al. (2015) showed based on the results of a paleogenomic study that the upper boundary of divergence between the ancestors of the Native Americans and East Asians should be considered as ~23 thousand years.

In that case, the isolation period in Beringia (before leaving for America 15,000 years ago) was ~8,000 years. Estimates of the duration of the Beringian Standstill obtained by using the analysis of ancient mitogenomes demonstrated that, given various kinds of uncertainties, the isolation period of the Beringian population could have ranged from 2.4 to 9.0 thousand years (Llamas et al., 2016).

Such a long stay in Beringia suggests that its population was quite large in number and subdivided within Beringia (Hoffecker et al., 2016; Moreno-Mayar et al., 2018a). It is assumed that the western part of Beringia (Chukotka) was inhabited by Paleosiberian populations, while the eastern part of Beringia (Alaska and Yukon) was inhabited in the south by the Native American ancestors and in the north by a separate group of ancient Beringians who disappeared or were assimilated by northern Native American tribes (Moreno-Mayar et al., 2018a; Sikora et al., 2019; Willerslev, Meltzer, 2021). Both of these groups of the ancient Eastern Beringians were very closely related, since they showed the C1b and B2 mitochondrial haplogroups typical of the Native Americans (Tackney et al., 2015; Moreno-Mayar et al., 2018b). Moreover, at present these haplogroups are more characteristic of the southern Native American populations, but in the past they were present in the north of Eastern Beringia, thus confirming the Beringian Standstill hypothesis (Tackney et al., 2015; Moreno-Mayar et al., 2018a).

It is likely that during the glacial maximum, the southern part of Beringia was a refugium, characterized by a fairly mild climate, rich biota, and therefore quite suitable for settlement of ancient people (Hoffecker et al., 2016; Sikora et al., 2019). According to the results of paleogenomic data modeling, the divergence between the Paleosiberians and the population that gave rise to the Beringians occurred approximately 24 thousand years ago, but the Upper Paleolithic genetic component was obtained by both populations approximately 20 thousand years ago (in the ancient Beringians its fraction was ~18 %) (Sikora et al., 2019). Thus, it is assumed that the most ancient Upper Paleolithic peoples were actually present in Beringia during the last glacial maximum, and that the formation of Beringians, including Native American ancestors, began through interaction with the newcomer East Asia-related peoples ~20 thousand years ago (Sikora et al., 2019). In this scenario, the time of existence of the relatively isolated Beringian population is estimated to be ~5,000 years.

A study of Y-chromosome polymorphism in Native American populations also showed that the duration of the Beringian Standstill may not have exceeded 4,600 years (Pinotti et al., 2019). Thus, the results of genomic studies suggest that the adaptation to the Arctic conditions of the ancient populations that took refuge in Beringia may have taken several millennia.

On adaptive selection in Native American ancestors

According to some genetic studies, it is proposed that the genomes of the Native Americans may contain signals of adaptation to the Arctic conditions, which emerged during the Beringian stage of the Native American ancestors formation, which, as noted above, may have taken a fairly long period of time (Ruiz-Pesini et al., 2004; Amorim et al., 2017). A recent study of genomic polymorphism in various contemporary continental groups showed that Native Americans harbor at least 20,424 genetic variants suspected to have originated in Beringia (Niedbalski, Long, 2022). This is comparable to the number of genetic variants that distinguish Africans from non-Africans. However, there are tens of times less specific genetic variants in Eurasia. At the same time, group-specific polymorphisms have not been found at all in Europeans, thus indicating rather intensive inter-ethnic contacts within Eurasia (Niedbalski, Long, 2022).

A study of the influence of adaptive selection on the genetic profile of Native Americans has shown that the appearance of American-specific genetic variants is associated with various metabolic pathways, but the most important ones are associated with the production of melanin in the skin, hair and eyes, as well as cardiovascular function (Niedbalski, Long, 2022).

It should be noted, however, that the most distinct traces of adaptation to the severe conditions of the Far North have been revealed in the gene pools of the Eskimo and Paleo-Asiatic peoples rather than in those of the Native Americans. Eskimo and Paleo-Asiatic populations also originated in the Asian and North American Arctic during the last 5–6 thousand years (Flegontov et al., 2019; Willerslev, Meltzer, 2021) and thus, their periods of adaptation are quite comparable in time to the Beringian population. However, in the next 15 thousand years after their isolation, the Beringians began to occupy more southern American territories, and therefore it is unknown to what extent the traces of their ancestors' adaptation to the Arctic are preserved in the gene pools of modern Native Americans (Adhikari et al., 2019).

A number of studies of modern and ancient populations of the Far North of Asia and America have found evidence of genetic and physiological adaptation of Eskimo and Paleo-Asiatic ancestors to low temperatures and an "Arctic" diet based primarily on the consumption of seafood rich in polyunsaturated fatty acids (PUFAs) (Cardona et al., 2014; Clemente et al., 2014; Fumagalli et al., 2015).

One of the most striking examples of diet-related genetic markers is the "Arctic" variant of the *CPT1A* gene (rs80356779-A) (Clemente et al., 2014; Malyarchuk, 2020). It is known that the "Arctic" variant is widely spread in modern populations of the Eskimos, Chukchi, Koryaks, and other peoples of the Sea of Okhotsk region, whose economic style is connected with sea hunting (Malyarchuk et al., 2016). According to paleogenomic data, the earliest records of the "Arctic" variant of the *CPT1A* gene were found among the Greenland and Canadian Paleo-Eskimos (4,000 years ago), representatives of the Tokarev culture of the North of the Sea of Okhotsk region (3,000 years ago) and carriers of the Late Jomon culture of Hokkaido Island (3,500–3,800 years ago). The appearance of the "Arctic" variant of the *CPT1A* gene is most likely associated with the adaptive response of the Far North indigenous peoples to the "Arctic" diet, which is characterized by a marked excess of lipids and proteins and a deficit of carbohydrates.

In addition, genetic variants associated with carbohydrate metabolism are very frequent in indigenous populations of

Northeast Asia. For example, the maximum frequency (52 %) of deletion of the pancreatic amylase gene *AMY2A*, necessary for starch digestion, was found in the Eskimo, Chukchi, and Koryak populations; ~30 % of the indigenous peoples of Northeast Asia lack this gene (Inchley et al., 2016). The high frequency of *AMY2A* gene deletion in Northeast Asia may be explained by the deficiency of starch and disaccharides in the traditional diet of indigenous peoples in the past. A deficiency in oligosaccharides can also explain the high incidence of the inactive sucrose-isomaltase gene (gene *SI*) among the indigenous populations of the Far North of Asia and America (Malyarchuk et al., 2017; Pedersen et al., 2017). In Greenlandic Eskimos, glucose homeostasis-related mutations in the *TBC1D4* and *ADCY3* genes have also been identified, which increase the risk of obesity and type 2 diabetes (Moltke et al., 2014; Grarup et al., 2018).

Thus, genetic changes that occurred in the past due to adaptation to extreme environments, in the current conditions (when there is no carbohydrate deficiency in the diets of the Far North indigenous populations) become harmful, because they increase the risk of metabolic and other diseases.

Interestingly, in a number of cases, adaptive changes in the gene pools of Eskimo and Paleo-Asiatic populations are caused by nonsynonymous substitutions with high pathogenicity indices (in the *CPT1A*, *CRAT* genes), nonsense mutations leading to stop codons or splicing disruption (in the *SI*, *TBC1D4* and *ADCY3* genes), and whole gene deletions (as in the case of the *AMY2A* gene) (Malyarchuk, Derenko, 2017; Pedersen et al., 2017; Malyarchuk, 2018). It seems likely that this kind of fairly profound genetic changes could have been preserved in the gene pools of Native Americans if they had adapted to the Arctic environment in a similar way, involving a very specific "Arctic" diet.

There is a suggestion that since the ancient inhabitants of Beringia migrated first to the sub-Arctic Pacific coast and then on to Northwest America, they may have practiced a maritime type of economy at the sub-Arctic stage (Hoffecker et al., 2016; Pedersen et al., 2016; McLaren et al., 2018). But, apparently, this stage of migration history was not reflected in the gene pools of the Native Americans, or their ancestors did not hunt sea mammals, but hunted mainly megafauna: mammoths, bison, horses and other terrestrial animals of the Great Arctic Plain (Pitulko et al., 2017; Lindgren et al., 2018; Harris et al., 2019). In that case, it is assumed that Amerindian ancestors had to adapt primarily to vitamin D deficiency, because at high latitudes UV radiation levels are insufficient for annual synthesis of cholecalciferol D3 in the skin, and the intake of ergocalciferol D2 depends on diet.

Genetic adaptation to vitamin D deficiency suggests that the dietary sources of this substance that are available to Arctic peoples (e. g., sea mammals, fish, eggs) were scarce or absent in the Beringian steppe-tundra biome. Vitamin D deficiency affects the health of children to a greater extent, and so L.J. Hlusko et al. (2018) suggested that Beringians adapted to UV deficiency through a V370A substitution in the ectodyslasin receptor encoded by the *EDAR* gene. This mutation is widespread in populations of East and North Asia, as well as in Native Americans. One manifestation of this mutation is associated with an increase in the branching density of mammary gland ducts, which leads to higher amounts of vitamin D in breast milk (Hlusko et al., 2018).

A recent study of Latin American populations also showed that genes related to energy metabolism played a crucial role in population adaptation to the environment of America (Mendoza-Revilla et al., 2022). The great importance of immune response genes, primarily *HLA* genes, in adaptive processes has also been highlighted. Moreover, positive selection in these genes is recorded in Native Americans at different periods of their history, both after the European colonization of America, which introduced new pathogens, and during the initial settlement of America at the first contact with endemic pathogens (Lindo et al., 2016; Mendoza-Revilla et al., 2022).

FADS gene cluster

The hypothesis suggesting the influence of high-latitude UV deficiency on the genetic features of the ancient Beringians is based on an additional argument involving the adaptive selection of the *FADS* genes variants in Native Americans (Hlusko et al., 2018). The fatty acid desaturase genes *FADS1* and *FADS2* encode enzymes involved in PUFA biosynthesis from shorter precursors (Nakamura, Nara, 2004). It is assumed that the ancestral A-variants of the *FADS* genes gave advantages to human populations that consumed food rich in lipids and proteins (for example, the Upper Paleolithic population of Eurasia), but later in the Neolithic, after the emergence of agricultural technology, the D-variants became more common, allowing for a higher rate of synthesis of PUFAs from plant lipids (Ameur et al., 2012; Mathieson S., Mathieson I., 2018).

Studies have shown that *FADS* genes have been affected by positive selection for a long time – for example, in Europe there has been an increase in the frequency of haplotype D from less than 10 % 10,000 years ago to 60–75 % at present (Mathieson, 2020). However, the exact reasons for the increased frequencies of the *FADS* genes variants in certain regions of the world have not yet been established. This is due to the fact that *FADS* genes are highly pleiotropic. Associations were found for *FADS1* and *FADS2* genetic variants with blood lipids and other metabolites, various blood cell phenotypes, suggesting a link with cardiovascular function (Draisma et al., 2015). Besides, metabolic balance with respect to PUFAs and other lipids is very important for brain function, in which the *FADS1* gene is actively expressed (Mathieson et al., 2020).

It has been suggested that the emergence of the D haplotype may have been due not only to the need to increase the efficiency of PUFA synthesis when there is a deficiency in the diet (for example, in Africa, where the frequency of the D haplotype is still high), but also contributed to an increase in the brain size (Ameur et al., 2012). In addition, it was found that in Greenlandic Eskimos, *FADS* genes are involved in the processes of adaptation to the cold, as there are links between genetic variants and the distribution of body fat and height; and the participation of *FADS* genes in the regulation of growth hormones is also suggested (Fumagalli et al., 2015).

Population genetic studies have shown that the ancestral haplotype A is much more common among the indigenous populations of Siberia and America (Amorim et al., 2017;

Hsieh et al., 2017; Ye et al., 2017; Malyarchuk, Derenko, 2018). It has also been found that in Greenlandic Eskimos and Native Americans, the *FADS1* and *FADS2* genes are under positive selection, which contributed to an almost complete fixation of haplotype A in these populations (Fumagalli et al., 2015; Amorim et al., 2017). Meanwhile, a number of other studies on Arctic populations of Siberia (Eskimos, Chukchi, Koryaks, Nganasans, Yakuts) and North America (Eskimos) have not revealed the effect of selection on the *FADS1* and *FADS2* genes (Cardona et al., 2014; Hsieh et al., 2017; Reynolds et al., 2019).

Nevertheless, there is a hypothesis that the original positive selection signal in the *FADS1* and *FADS2* genes arose in the ancient population of Beringia due to the need to adapt to the cold and limited food resources provided by the Arctic (Amorim et al., 2017). It is assumed that these adaptive genetic variants have been preserved in the gene pools of the Native Americans and Eskimos. However, different combinations of the *FADS* genes variants are responsible for the effect of selection in Arctic and Native American aborigines and in Europeans (Fumagalli et al., 2015; Amorim et al., 2017; Ye et al., 2017; Mathieson, 2020). This can be due to different reasons for the selection of the *FADS* genes variants in populations. In Europeans it was a shift towards a plant-based diet in the Neolithic, due to which the D-haplotypes with increased desaturase activity were more favorable, and in Eskimos ancestors it was the emergence of a specific "Arctic" diet with very high PUFA content, due to which maintaining the maximum frequency of A-haplotypes was more preferable. It is possible that these *FADS* variants in Eskimos possess even lower desaturase activity (Mathieson, 2020).

Meanwhile, a recent analysis of the *FADS* genes polymorphism in modern and ancient populations has shown that the widespread distribution of haplotype A occurred under the influence of selection already in the Upper Paleolithic populations of Eurasia, and therefore in the gene pool of the ancient Beringians this haplotype could be fixed quite accidentally under the effects of genetic drift (Harris et al., 2019; Mathieson, 2020). Thus, the adaptive changes of *FADS* genes in the Native American ancestors are questioned, and the hypothesis that the *FADS* haplotypes inherited from the Upper Paleolithic Eurasians are preserved in the gene pools of Native Americans looks more likely.

ACTN3 gene

According to the published data, another example of selective advantages obtained by Native Americans from the ancient inhabitants of Beringia may be related to adaptive changes in the *ACTN3* gene (Amorim et al., 2015). This gene encodes the protein α-actinin-3, which is expressed exclusively in fast-twitch skeletal muscle fibers. Of the polymorphic loci of the *ACTN3* gene, the most studied is rs1815739, a mutation in which leads to termination of protein synthesis at amino acid position 577 of exon 16 (R577X substitution) (North et al., 1999). This causes a deficiency of α-actinin-3 in fast-twitch skeletal muscle fibers, which, in turn, can result in a decrease in the speed and power indicators of human physical performance (Alfred et al., 2011). Genetic studies have shown that the frequency of the 577X allele increases with distance from Africa to the north of Eurasia and reaches a maximum in the Native American populations (Amorim et al., 2015). It is suggested that the advantages of the 577X allele, such as increased endurance and improved protection against the cold, have contributed to the spread of this mutation (Bramble, Lieberman, 2004; Wyckelsma et al., 2021).

Recent physiological and metabolomic studies have shown that carriers of the 577XX genotype tolerate cold much better than 577RR individuals (Wyckelsma et al., 2021), which may support the hypothesis of selection of the X allele at the initial stage of human expansion into Eurasia (Amorim et al., 2015). According to this hypothesis, the elevated frequency of the 577X allele in some human populations could be related to selection to improve metabolic efficiency and promote dynamic activities (e. g., long hunting and, in sports, marathon running, swimming, cycling).

The high frequency of the 577X allele in Native Americans suggests that this allele was also widespread in ancient Beringians, ancestral to Native Americans. However, the results of analysis of the distribution of the rs1815739 polymorphic variants in the indigenous populations of northeastern Siberia (in Chukchi, Koryaks, and Evens) have shown that the frequencies of the 577X allele and 577XX genotype were not the highest in Eurasia (Malyarchuk et al., 2018). While the frequency of the 577X allele in Northeast Siberia is approximately 36 % (Malyarchuk et al., 2018), the maximum frequencies of this genetic variant are registered in South Asian and Native American populations – over 60 % according to dbSNP database (www.ncbi.nlm.nih.gov/snp).

A recent analysis of geographic distribution of rs1815739 variants in modern and ancient populations has shown that population frequencies of the 577X allele correlate neither with geographic latitude nor with temperature (Mörseburg et al., 2022). None of the statistical tests used in this study found evidence of the effect of positive selection for the 577X allele at high latitudes. Thus, the high frequency of this genetic variant in Native Americans is most consistent with the effect of genetic drift, possibly occurring in Beringia at the stage of low size of ancestral population. It is possible, however, that some advantages of this genetic variant, namely increased endurance and resistance to cold, contributed to the increased frequency of the 577X allele in Native American ancestors.

Conclusion

The results of genetic studies of modern and ancient populations of Eurasia and America have demonstrated quite convincingly that the role of Beringia in the settlement of the Americas is very high. To date, sufficiently reliable estimates of divergence times between the ancestral genetic lineages that led to the formation of the various populations of Northeast Asia and the Americas have been obtained. The most probable scenario seems to be that during the last glacial maximum, the population that gave rise to the Native Americans ancestors persisted in Beringia for several millennia (20,000–15,000 years ago). It is assumed that this was a relatively small (from several hundred to several thousand individuals (Fagundes et al., 2018)) group of people well adapted to high latitudes and cold

environments, which is confirmed by the results of genomic studies of modern Native American populations.

Possible traces of adaptation to Arctic environments at the genomic level are associated with various metabolic pathways – with melanin synthesis processes, cardiovascular system functioning, energy metabolism, and immune response genes (Mendoza-Revilla et al., 2022; Niedbalski, Long, 2022). It has also been suggested that the relaxation of negative selection in a number of protein-coding genes observed in Native Americans is also associated with the Beringian stage of population adaptation (Niedbalski, Long, 2022). However, it is still unclear to what extent such adaptive signals could have survived, given the approximately 15-thousand-year period of American colonization following the Beringian Standstill (Adhikari et al., 2019).

Obviously, answers to such questions can be obtained by using paleogenomic data, but very few ancient human remains have been found from the territory of ancient Beringia. So far, we can get some insight into the genetic features of the Beringians from only one genome of the Eastern Beringian descendant (11.5 thousand years ago, Upward Sun River site, Alaska) (Moreno-Mayar et al., 2018a). Thus, investigations into the role of Beringia in the origins of the Native American ancestors remain relevant and ongoing.

References

Achilli A., Perego U.A., Lancioni H., Olivieri A., Gandini F., Hooshiar Kashani B., Battaglia V., Grugni V., Angerhofer N., Rogers M.P., Herrera R.J., Woodward S.R., Labuda D., Smith D.G., Cybulski J.S., Semino O., Malhi R.S., Torroni A. Reconciling migration models to the Americas with the variation of North American native mitogenomes. *Proc. Natl. Acad. Sci. USA*. 2013;110(35):14308-14313. DOI 10.1073/pnas.1306290110.

Adhikari K., Mendoza-Revilla J., Sohail A., Fuentes-Guajardo M., Lampert J., Chacón-Duque J.C., Hurtado M., Villegas V., Granja V., Acuña-Alonzo V., Jaramillo C., Arias W., Lozano R.B., Everardo P., Gómez-Valdés J., Villamil-Ramírez H., Silva de Cerqueira C.C., Hunemeier T., Ramallo V., Schuler-Faccini L., Salzano F.M., Gonzalez-José R., Bortolini M.C., Canizales-Quinteros S., Gallo C., Poletti G., Bedoya G., Rothhammer F., Tobin D.J., Fumagalli M., Balding D., Ruiz-Linares A. A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. *Nat. Commun.* 2019;10(1):358. DOI 10.1038/s41467-018-08147-0.

Alfred T., Ben-Shlomo Y., Cooper R., Hardy R., Cooper C., Deary I.J., Gunnell D., Harris S.E., Kumari M., Martin R.M., Moran C.N., Pitsiladis Y.P., Ring S.M., Sayer A.A., Smith G.D., Starr J.M., Kuh D., Day I.N., HALCyon study team. *ACTN3* genotype, athletic status, and life course physical capability: meta-analysis of the published literature and findings from nine studies. *Hum. Mutat*. 2011; 32(9):1008-1018.

Ameur A., Enroth S., Johansson A., Zaboli G., Igl W., Johansson A.C., Rivas M.A., Daly M.J., Schmitz G., Hicks A.A., Meitinger T., Feuk L., van Duijn C., Oostra B., Pramstaller P.P., Rudan I., Wright A.F., Wilson J.F., Campbell H., Gyllensten U. Genetic adaptation of fatty-acid metabolism: a human specific haplotype increasing the biosynthesis of long-chain omega-3 and omega-6 fatty acids. *Am. J. Hum. Genet*. 2012;90(5):809-820. DOI 10.1016/j.ajhg.2012.03.014.

Amorim C.E., Acuña-Alonzo V., Salzano F.M., Bortolini M.C., Hünemeier T. Differing evolutionary histories of the *ACTN3*R577X* polymorphism among the major human geographic groups. *PLoS One*. 2015;10(2):e0115449. DOI 10.1371/journal.pone.0115449.

Amorim C.E., Nunes K., Meyer D., Comas D., Bortolini M.C., Salzano F.M., Hünemeier T. Genetic signature of natural selection in first Americans. *Proc. Natl. Acad. Sci. USA*. 2017;114(9):2195-2199. DOI 10.1073/pnas.1620541114.

Batchelor C.L., Margold M., Krapp M., Murton D.K., Dalton A.S., Gibbard P.L., Stokes C.R., Murton J.B., Manica A. The configuration of Northern Hemisphere ice sheets through the Quaternary. *Nat. Commun*. 2019;10(1):3713. DOI 10.1038/s41467-019-11601-2.

Behar D.M., van Oven M., Rosset S., Metspalu M., Loogväli E.L., Silva N.M., Kivisild T., Torroni A., Villems R. A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet*. 2012;90(4):675-684. DOI 10.1016/j.ajhg.2012.03.002.

Bramble D.M., Lieberman D.E. Endurance running and the evolution of *Homo*. *Nature*. 2004;432(7015):345-352. DOI 10.1038/nature03052.

Cardona A., Pagani L., Antao T., Lawson D.J., Eichstaedt C.A., Yngvadottir B., Shwe M.T.T., Wee J., Romero I.G., Raj S., Metspalu M., Villems R., Willerslev E., Tyler-Smith C., Malyarchuk B.A., Derenko M.V., Kivisild T. Genome-wide analysis of cold adaptation in indigenous Siberian populations. *PLoS One*. 2014;9(5):e98076. DOI 10.1371/journal.pone.0098076.

Clark J., Carlson A.E., Reyes A.V., Carlson E.C.B., Guillaume L., Milne G.A., Tarasov L., Caffee M., Wilcken K., Rood D.H. The age of the opening of the Ice-Free Corridor and implications for the peopling of the Americas. *Proc. Natl. Acad. Sci. USA*. 2022;119(14): e2118558119. DOI 10.1073/pnas.2118558119.

Clemente F.J., Cardona A., Inchley C.E., Peter B.M., Jacobs G., Pagani L., Lawson D.J., Antão T., Vicente M., Mitt M., DeGiorgio M., Faltyskova Z., Xue Y., Ayub Q., Szpak M., Mägi R., Eriksson A., Manica A., Raghavan M., Rasmussen M., Rasmussen S., Willerslev E., Vidal-Puig A., Tyler-Smith C., Villems R., Nielsen R., Metspalu M., Malyarchuk B., Derenko M., Kivisild T. A selective sweep on a deleterious mutation in *CPT1A* in Arctic populations. *Am. J. Hum. Genet*. 2014;95(5):584-589. DOI 10.1016/j.ajhg.2014.09.016.

Colombo G., Traverso L., Mazzocchi L., Grugni V., Rambaldi Migliore N., Capodiferro M.R., Lombardo G., Flores R., Karmin M., Rootsi S., Ferretti L., Olivieri A., Torroni A., Martiniano R., Achilli A., Raveane A., Semino O. Overview of the Americas' first peopling from a patrilineal perspective: new evidence from the Southern continent. *Genes.* 2022;13(2):220. DOI 10.3390/genes13020220.

Der Sarkissian C., Brotherton P., Balanovsky O., Templeton J.E.L., Llamas B., Soubrier J., Moiseyev V., Khartanovich V., Cooper A., Haak W., The Genographic Consortium. Mitochondrial genome sequencing in Mesolithic North East Europe unearths a new sub-clade within the broadly distributed human haplogroup C1. *PLoS One*. 2014;9(2):e87612. DOI 10.1371/journal.pone.0087612.

Derenko M., Malyarchuk B., Denisova G., Perkova M., Litvinov A., Grzybowski T., Dambueva I., Skonieczna K., Rogalla U., Tsybovsky I., Zakharov I. Western Eurasian ancestry in modern Siberians based on mitogenomic data. *BMC Evol. Biol*. 2014;14:217. DOI 10.1186/s12862-014-0217-9.

Derenko M., Malyarchuk B., Grzybowski T., Denisova G., Rogalla U., Perkova M., Dambueva I., Zakharov I. Origin and post-glacial dispersal of mitochondrial DNA haplogroups C and D in northern Asia. *PLoS One*. 2010;5(12):e15214. DOI 10.1371/journal.pone.0015214.

Draisma H.H.M., Pool R., Kobl M., Jansen R., Petersen A.K., Vaarhorst A.A.M., Yet I., Haller T., Demirkan A., Esko T., Zhu G., Böhringer S., Beekman M., van Klinken J.B., Römisch-Margl W., Prehn C., Adamski J., de Craen A.J.M., van Leeuwen E.M., Amin N., Dharuri H., Westra H.J., Franke L., de Geus E.J.C., Hottenga J.J., Willemsen G., Henders A.K., Montgomery G.W., Nyholt D.R., Whitfield J.B., Penninx B.W., Spector T.D., Metspalu A., Slagboom P.E., van Dijk K.W., 't Hoen P.A.C., Strauch K., Martin N.G., van Ommen G.B., Illig T., Bell J.T., Mangino M., Suhre K., McCarthy M.I., Gieger C., Isaacs A., van Duijn C.M., Boomsma D.I. Genome-wide

association study identifies novel genetic variants contributing to variation in blood metabolite levels. *Nat. Commun*. 2015;6:7208. DOI 10.1038/ncomms8208.

Duggan A.T., Whitten M., Wiebe V., Crawford M., Butthof A., Spitsyn V., Makarov S., Novgorodov I., Osakovsky V., Pakendorf B. Investigating the prehistory of Tungusic peoples of Siberia and the Amur-Ussuri region with complete mtDNA genome sequences and Y-chromosomal markers. *PLoS One*. 2013;8(12):e83570. DOI 10.1371/journal.pone.0083570.

Ebenesersdóttir S.S., Sigurðsson A., Sánchez-Quinto F., Lalueza-Fox C., Stefánsson K., Helgason A. A new subclade of mtDNA haplogroup C1 found in Icelanders: evidence of pre-Columbian contact? *Am. J. Phys. Anthropol.* 2011;144(1):92-99. DOI 10.1002/ajpa.21419.

Fagundes N.J.R., Tagliani-Ribeiro A., Rubicz R., Tarskaia L., Crawford M.H., Salzano F.M., Bonatto S.L. How strong was the bottleneck associated to the peopling of the Americas? New insights from multilocus sequence data. *Genet. Mol. Biol*. 2018;41(Suppl. 1):206-214. DOI 10.1590/1678-4685-GMB-2017-0087.

Flegontov P., Altınışık N.E., Changmai P., Rohland N., Mallick S., Adamski N., Bolnick D.A., Broomandkhoshbacht N., Candilio F., Culleton B.J., Flegontova O., Friesen T.M., Jeong C., Harper T.K., Keating D., Kennett D.J., Kim A.M., Lamnidis T.C., Lawson A.M., Olalde I., Oppenheimer J., Potter B.A., Raff J., Sattler R.A., Skoglund P., Stewardson K., Vajda E.J., Vasilyev S., Veselovskaya E., Hayes M.G., O'Rourke D.H., Krause J., Pinhasi R., Reich D., Schiffels S. Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America. *Nature*. 2019;570(7760):236-240. DOI 10.1038/s41586-019-1251-y.

Fumagalli M., Moltke I., Grarup N., Racimo F., Bjerregaard P., Jørgensen M.E., Korneliussen T.S., Gerbault P., Skotte L., Linneberg A., Christensen C., Brandslund I., Jørgensen T., Huerta-Sánchez E., Schmidt E.B., Pedersen O., Hansen T., Albrechtsen A., Nielsen R. Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science*. 2015;349(6254):1343-1347. DOI 10.1126/science.aab2319.

Grarup N., Moltke I., Andersen M.K., Dalby M., Vitting-Seerup K., Kern T., Mahendran Y., Jørsboe E., Larsen C.V.L., Dahl-Petersen I.K., Gilly A., Suveges D., Dedoussis G., Zeggini E., Pedersen O., Andersson R., Bjerregaard P., Jørgensen M.E., Albrechtsen A., Hansen T. Loss-of-function variants in *ADCY3* increase risk of obesity and type 2 diabetes. *Nat. Genet*. 2018;50(2):172-174. DOI 10.1038/s41588-017-0022-7.

Harris D.N., Ruczinski I., Yanek L.R., Becker L.C., Becker D.M., Guio H., Cui T., Chilton F.H., Mathias R.A., O'Connor T.D. Evolution of hominin polyunsaturated fatty acid metabolism: from Africa to the New World. *Genome Biol. Evol*. 2019;11(5):1417-1430. DOI 10.1093/gbe/evz071.

Hlusko L.J., Carlson J.P., Chaplin G., Elias S.A., Hoffecker J.F., Huffman M., Jablonski N.G., Monson T.A., O'Rourke D.H., Pilloud M.A., Scott G.R. Environmental selection during the last ice age on the mother-to-infant transmission of vitamin D and fatty acids through breast milk. *Proc. Natl. Acad. Sci. USA*. 2018;115(19):E4426-E4432. DOI 10.1073/pnas.1711788115.

Hoffecker J.F., Elias S.A., O'Rourke D.H., Scott G.R., Bigelow N.H. Beringia and the global dispersal of modern humans. *Evol. Anthropol.* 2016;25(2):64-78. DOI 10.1002/evan.21478.

Hooshiar Kashani B., Perego U.A., Olivieri A., Angerhofer N., Gandini F., Carossa V., Lancioni H., Semino O., Woodward S.R., Achilli A., Torroni A. Mitochondrial haplogroup C4c: a rare lineage entering America through the ice-free corridor? *Am. J. Phys. Anthropol*. 2012;147(1):35-39. DOI 10.1002/ajpa.21614.

Hsieh P., Hallmark B., Watkins J., Karafet T.M., Osipova L.P., Gutenkunst R.N., Hammer M.F. Exome sequencing provides evidence of polygenic adaptation to a fat-rich animal diet in indigenous Siberian populations. *Mol. Biol. Evol*. 2017;34(11):2913-2926. DOI 10.1093/molbev/msx226.

Inchley C.E., Larbey C.D., Shwan N.A., Pagani L., Saag L., Antão T., Jacobs G., Hudjashov G., Metspalu E., Mitt M., Eichstaedt C.A., Malyarchuk B., Derenko M., Wee J., Abdullah S., Ricaut F.X., Mormina M., Mägi R., Villems R., Metspalu M., Jones M.K., Armour J.A., Kivisild T. Selective sweep on human amylase genes postdates the split with Neanderthals. *Sci. Rep.* 2016;6:37198. DOI 10.1038/srep37198.

Kilinç G.M., Kashuba N., Koptekin D., Bergfeldt N., Dönertaş H.M., Rodríguez-Varela R., Shergin D., Ivanov G., Kichigin D., Pestereva K., Volkov D., Mandryka P., Kharinskii A., Tishkin A., Ineshin E., Kovychev E., Stepanov A., Dalén L., Günther T., Kırdök E., Jakobsson M., Somel M., Krzewińska M., Storå J., Götherström A. Human population dynamics and *Yersinia pestis* in ancient northeast Asia. *Sci. Adv*. 2021;7(2):eabc4587. DOI 10.1126/sciadv.abc4587.

Lesnek A.J., Briner J.P., Lindqvist C., Baichtal J.F., Heaton T.H. Deglaciation of the Pacific coastal corridor directly preceded the human colonization of the Americas. *Sci. Adv*. 2018;4(5):aar5040. DOI 10.1126/sciadv.aar5040.

Lindgren A., Hugelius G., Kuhry P. Extensive loss of past permafrost carbon but a net accumulation into present-day soils. *Nature*. 2018;560(7717):219-222. DOI 10.1038/s41586-018-0371-0.

Lindo J., Huerta-Sánchez E., Nakagome S., Rasmussen M., Petzelt B., Mitchell J., Cybulski J.S., Willerslev E., DeGiorgio M., Malhi R.S. A time transect of exomes from a Native American population before and after European contact. *Nat. Commun*. 2016;7:13175. DOI 10.1038/ncomms13175.

Llamas B., Fehren-Schmitz L., Valverde G., Soubrier J., Mallick S., Rohland N., Nordenfelt S., Valdiosera C., Richards S.M., Rohrlach A., Romero M.I., Espinoza I.F., Cagigao E.T., Jiménez L.W., Makowski K., Reyna I.S., Lory J.M., Torrez J.A., Rivera M.A., Burger R.L., Ceruti M.C., Reinhard J., Wells R.S., Politis G., Santoro C.M., Standen V.G., Smith C., Reich D., Ho S.Y., Cooper A., Haak W. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* 2016;2(4):e1501385. DOI 10.1126/sciadv.1501385.

Malyarchuk B.A. Long-term gene-environment interactions and genetics of metabolic disorders in aboriginal populations of Northeast Asia. *Ekologicheskaya Genetika = Ecological Genetics*. 2018;16(2):30-35. DOI 10.17816/ecogen16230-35. (in Russian)

Malyarchuk B.A. Genetic markers on the distribution of ancient marine hunters in Priokhotye. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2020;24(5):539-544. DOI 10.18699/VJ20.646. (in Russian)

Malyarchuk B.A., Derenko M.V. Polymorphism of the genes encoding for the carnitine acyltransferases in native populations of Siberia. *Ekologicheskaya Genetika = Ecological Genetics*. 2017;15(4):13-18. DOI 10.17816/ecogen15413-18. (in Russian)

Malyarchuk B.A., Derenko M.V. Polymorphism of genes of polyunsaturated fatty acids metabolism (*FADS1* and *FADS2*) in aboriginal populations of Siberia. *Vestnik Severo-Vostochnogo Nauchnogo Tsentra DVO RAN = The Bulletin of the North-East Scientific Center*. 2018;3:106-111. (in Russian)

Malyarchuk B.A., Derenko M.V., Denisova G.A. The frequency of inactive sucrase-isomaltase variant in indigenous populations of Northeast Asia. *Russ. J. Genet.* 2017;53(9):1052-1054. DOI 10.1134/S1022795417090095.

Malyarchuk B.A., Derenko M.V., Denisova G.A. R577X polymorphism of alpha-actinin-3 in the human populations of Northeastern Asia. *Russ. J. Genet. Appl. Res*. 2018;8(1):59-64. DOI 10.1134/S2079059718010094.

Malyarchuk B.A., Derenko M.V., Denisova G.A., Litvinov A.N. Distribution of the arctic variant of the *CPT1A* gene in indigenous populations of Siberia. *Vavilovskii Zhurnal Genetiki i Selektsii =*

Vavilov Journal of Genetics and Breeding. 2016;20(5):571-575. DOI 10.18699/VJ16.130. (in Russian)

Malyarchuk B., Derenko M., Denisova G., Maksimov A., Wozniak M., Grzybowski T., Dambueva I., Zakharov I. Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. *J. Hum. Genet*. 2011;56(8):583-588. DOI 10.1038/jhg.2011.64.

Mathieson I. Limited evidence for selection at the *FADS* locus in Native American populations. *Mol. Biol. Evol*. 2020;37(7):2029-2033. DOI 10.1093/molbev/msaa064.

Mathieson I., Day F.R., Barban N., Tropf F.C., Brazel D.M., eQTLGen Consortium, BIOS Consortium, Vaez A., van Zuydam N., Bitarello B.D., Snieder H., den Hoed M., Ong K.K., Mills M.C., Perry J.R.B. Genome-wide analysis identifies genetic effects on reproductive success and ongoing natural selection at the *FADS* locus. *bioRxiv*. 2020. DOI 10.1101/2020.05.19.104455.

Mathieson S., Mathieson I. *FADS1* and the timing of human adaptation to agriculture. *Mol. Biol. Evol*. 2018;35(12):2957-2970. DOI 10.1093/molbev/msy180.

McLaren D., Fedje D., Dyck A., Mackie Q., Gauvreau A., Cohen J. Terminal Pleistocene epoch human footprints from the Pacific coast of Canada. *PLoS One*. 2018;13(3):e0193522. DOI 10.1371/journal.pone.0193522.

Mendoza-Revilla J., Chacón-Duque J.C., Fuentes-Guajardo M., Ormond L., Wang K., … Balding D., Fumagalli M., Adhikari K., Ruiz-Linares A., Hellenthal G. Disentangling signatures of selection before and after European colonization in Latin Americans. *Mol. Biol. Evol*. 2022;39(4):msac076. DOI 10.1093/molbev/msac076.

Moltke I., Grarup N., Jorgensen M.E., Bjerregaard P., Treebak J.T., Fumagalli M., Korneliussen T.S., Andersen M.A., Nielsen T.S., Krarup N.T., Gjesing A.P., Zierath J.R., Linneberg A., Wu X., Sun G., Jin X., Al-Aama J., Wang J., Borch-Johnsen K., Pedersen O., Nielsen R., Albrechtsen A., Hansen T. A common Greenlandic *TBC1D4* variant confers muscle insulin resistance and type 2 diabetes. *Nature*. 2014;512(7513):190-193. DOI 10.1038/nature13425.

Moreno-Mayar J.V., Potter B.A., Vinner L., Steinrücken M., Rasmussen S., Terhorst J., Kamm J.A., Albrechtsen A., Malaspinas A.S., Sikora M., Reuther J.D., Irish J.D., Malhi R.S., Orlando L., Song Y.S., Nielsen R., Meltzer D.J., Willerslev E. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature*. 2018a;553(7687):203-207. DOI 10.1038/nature25173.

Moreno-Mayar J.V., Vinner L., de Barros Damgaard P., de la Fuente C., Chan J., … Sikora M., Nielsen R., Song Y.S., Meltzer D.J., Willerslev E. Early human dispersals within the Americas. *Science*. 2018b;362(6419):eaav2621. DOI 10.1126/science.aav2621.

Mörseburg A., Pagani L., Malyarchuk B., Derenko M., Kivisild T. Response to Wyckelsma et al.: Loss of α-actinin-3 during human evolution provides superior cold resilience and muscle heat generation. *Am. J. Hum. Genet*. 2022;109(5):967-972. DOI 10.1016/j.ajhg.2022.03.014.

Nakamura M.T., Nara T.Y. Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. *Annu. Rev. Nutr*. 2004;24:345-376. DOI 10.1146/annurev.nutr.24.121803.063211.

Niedbalski S.D., Long J.C. Novel alleles gained during the Beringian isolation period. *Sci. Rep*. 2022;12(1):4289. DOI 10.1038/s41598-022-08212-1.

North K.N., Yang N., Wattanasirichaigoon D., Mills M., Easteal S., Beggs A.H. A common nonsense mutation results in alpha-actinin-3 deficiency in the general population. *Nat. Genet*. 1999;21(4):353-354. DOI 10.1038/7675.

Pedersen C.T., Lohmueller K.E., Grarup N., Bjerregaard P., Hansen T., Siegismund H.R., Moltke I., Albrechtsen A. The effect of an extreme and prolonged population bottleneck on patterns of deleterious variation: insights from the Greenlandic Inuit. *Genetics*. 2017;205(2):787-801. DOI 10.1534/genetics.116.193821.

Pedersen M.W., Ruter A., Schweger C., Friebe H., Staff R.A., Kjeldsen K.K., Mendoza M.L., Beaudoin A.B., Zutter C., Larsen N.K., Potter B.A., Nielsen R., Rainville R.A., Orlando L., Meltzer D.J., Kjær K.H., Willerslev E. Postglacial viability and colonization in North America's ice-free corridor. *Nature*. 2016;537(7618):45-49. DOI 10.1038/nature19085.

Perego U.A., Achilli A., Angerhofer N., Accetturo M., Pala M., Olivieri A., Hooshiar Kashani B., Ritchie K.H., Scozzari R., Kong Q.P., Myres N.M., Salas A., Semino O., Bandelt H.J., Woodward S.R., Torroni A. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr. Biol*. 2009;19(1):1-8. DOI 10.1016/j.cub.2008.11.058.

Pinotti T., Bergström A., Geppert M., Bawn M., Ohasi D., Shi W., Lacerda D.R., Solli A., Norstedt J., Reed K., Dawtry K., González-Andrade F., Paz-Y-Miño C., Revollo S., Cuellar C., Jota M.S., Santos J.E.Jr., Ayub Q., Kivisild T., Sandoval J.R., Fujita R., Xue Y., Roewer L., Santos F.R., Tyler-Smith C. Y chromosome sequences reveal a short Beringian standstill, rapid expansion, and early population structure of Native American founders. *Curr. Biol.* 2019;29(1):149-157.e3. DOI 10.1016/j.cub.2018.11.029.

Pitulko V.V., Nikolsky P.A., Girya E.Y., Basilyan A.E., Tumskoy V.E., Koulakov S.A., Astakhov S.N., Pavlova E.Y., Anisimov M.A. The Yana RHS site: humans in the Arctic before the last glacial maximum. *Science*. 2004;303(5654):52-56. DOI 10.1126/science.1085219.

Pitulko V., Pavlova E., Nikolskiy P. Revising the archaeological record of the Upper Pleistocene Arctic Siberia: Human dispersal and adaptations in MIS 3 and 2. *Quat. Sci. Rev*. 2017;165:127-148. DOI 10.1016/j.quascirev.2017.04.004.

Potter B.A., Baichtal J.F., Beaudoin A.B., Fehren-Schmitz L., Haynes C.V., Holliday V.T., Holmes C.E., Ives J.W., Kelly R.L., Llamas B., Malhi R.S., Miller D.S., Reich D., Reuther J.D., Schiffels S., Surovell T.A. Current evidence allows multiple models for the peopling of the Americas. *Sci. Adv*. 2018;4(8):eaat5473. DOI 10.1126/sciadv.aat5473.

Raghavan M., Skoglund P., Graf K.E., Metspalu M., Albrechtsen A., … Nielsen R., Jakobsson M., Willerslev E. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*. 2014;505(7481):87-91. DOI 10.1038/nature12736.

Raghavan M., Steinrücken M., Harris K., Schiffels S., Rasmussen S., … Song Y.S., Nielsen R., Willerslev E. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*. 2015;349(6250):aab3884. DOI 10.1126/science.aab3884.

Reynolds A.W., Mata-Míguez J., Miró-Herrans A., Briggs-Cloud M., Sylestine A., Barajas-Olmos F., Garcia-Ortiz H., Rzhetskaya M., Orozco L., Raff J.A., Hayes M.G., Bolnick D.A. Comparing signals of natural selection between three Indigenous North American populations. *Proc. Natl. Acad. Sci. USA*. 2019;116(19):9312-9317. DOI 10.1073/pnas.1819467116.

Ruiz-Pesini E., Mishmar D., Brandon M., Procaccio V., Wallace D.C. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*. 2004;303(5655):223-226. DOI 10.1126/science.1088434.

Sikora M., Pitulko V., Sousa V.C., Allentoft M.E., Vinner L., … Meltzer D., Excoffier L., Willerslev E. The population history of northeastern Siberia since the Pleistocene. *Nature*. 2019;570(7760):182-188. DOI 10.1038/s41586-019-1279-z.

Sikora M., Seguin-Orlando A., Sousa V.C., Albrechtsen A., Korneliussen T., Ko A., Rasmussen S., Dupanloup I., Nigst P.R., Bosch M.D., Renaud G., Allentoft M.E., Margaryan A., Vasilyev S.V., Veselovskaya E.V., Borutskaya S.B., Deviese T., Comeskey D., Higham T., Manica A., Foley R., Meltzer D.J., Nielsen R., Excoffier L., Mirazon Lahr M., Orlando L., Willerslev E. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science.* 2017;358(6363):659-662. DOI 10.1126/science.aao1807.

Tackney J.C., Potter B.A., Raff J., Powers M., Watkins W.S., Warner D., Reuther J.D., Irish J.D., O'Rourke D.H. Two contempora-

neous mitogenomes from terminal Pleistocene burials in eastern Beringia. *Proc. Natl. Acad. Sci. USA*. 2015;112(45):13833-13838. DOI 10.1073/pnas.1511903112.

Tamm E., Kivisild T., Reidla M., Metspalu M., Smith D.G., Mulligan C.J., Bravi C.M., Rickards O., Martinez-Labarga C., Khusnutdinova E.K., Fedorova S.A., Golubenko M.V., Stepanov V.A., Gubina M.A., Zhadanov S.I., Ossipova L.P., Damba L., Voevoda M.I., Dipierri J.E., Villems R., Malhi R.S. Beringian standstill and spread of Native American founders. *PLoS One*. 2007;2(9):e829. DOI 10.1371/journal.pone.0000829.

Vasil'ev S.A., Kuzmin Y.V., Orlova L.A., Dementiev V.N. Radiocarbon-based chronology of the Paleolithic in Siberia and its relevance to the peopling of the New World. *Radiocarbon*. 2002;44(2):503-530. DOI 10.1017/S0033822200031878.

Waters M.R. Late Pleistocene exploration and settlement of the Americas by modern humans. *Science*. 2019;365(138):eaat5447. DOI 10.1126/science.aat5447.

Willerslev E., Meltzer D.J. Peopling of the Americas as inferred from ancient genomics. *Nature*. 2021;594(7863):356-364. DOI 10.1038/s41586-021-03499-y.

Wyckelsma V.L., Venckunas T., Houweling P.J., Schlittler M., Lauschke V.M., Tiong C.F., Wood H.D., Ivarsson N., Paulauskas H., Eimantas N., Andersson D.C., North K.N., Brazaitis M., Westerblad H. Loss of α-actinin-3 during human evolution provides superior cold resilience and muscle heat generation. *Am. J. Hum. Genet*. 2021;108(3):446-457. DOI 10.1016/j.ajhg.2021.01.013.

Ye K., Gao F., Wang D., Bar-Yosef O., Keinan A. Dietary adaptation of *FADS* genes in Europe varied across time and geography. *Nat. Ecol. Evol*. 2017;1:0167. DOI 10.1038/s41559-017-0167.

Yu H., Spyrou M.A., Karapetian M., Shnaider S., Radzevičiūtė R., Nägele K., Neumann G.U., Penske S., Zech J., Lucas M., LeRoux P., Roberts P., Pavlenok G., Buzhilova A., Posth C., Jeong C., Krause J. Paleolithic to Bronze Age Siberians reveal connections with First Americans and across Eurasia. *Cell*. 2020;181(6):1232-1245.e20. DOI 10.1016/j.cell.2020.04.037.

ORCID ID

B.A. Malyarchuk orcid.org/0000-0002-0304-0652

Acknowledgements. This study was funded by the Russian Science Foundation grant No. 22-18-00319 (https://rscf.ru/project/22-18-00319/).

Conflict of interest. The author declares no conflict of interest.

Received August 3, 2022. Revised October 15, 2022. Accepted October 15, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):383-392
DOI 10.18699/VJGB-23-46

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

Clinical, genetic aspects and molecular pathogenesis of osteopetrosis

D.D. Nadyrshina[1] ✉, R.I. Khusainova[1, 2]

[1] Ufa University of Science and Technology, Ufa, Russia
[2] Saint Petersburg State University, St. Petersburg, Russia
✉ nadyrshinadina@gmail.com

Abstract. Osteopetrosis ("marble bone", ICD-10-78.2) includes a group of hereditary bone disorders distinguished by clinical variability and genetic heterogeneity. The name "osteopetrosis" comes from the Greek language: '*osteo*' means 'bone' and '*petrosis*' means 'stone', which characterizes the main feature of the disease: increased bone density caused by imbalances in bone formation and remodeling, leading to structural changes in bone tissue, predisposition to fractures, skeletal deformities. These defects, in turn, affect other important organs and tissues, especially bone marrow and the nervous system. The disease can be autosomal recessive, autosomal dominant, X-linked or sporadic. Autosomal dominant osteopetrosis has an incidence of 1 in 20,000 newborns and autosomal recessive one has 1 in 250,000. To date, 23 genes have been described, structural changes in which lead to the development of osteopetrosis. Clinical symptoms in osteopetrosis vary greatly in their presentation and severity. The mildest skeletal abnormalities are observed in adulthood and occur in the autosomal dominant form of osteopetrosis. Severe forms, being autosomal recessive and manifesting in early childhood, are characterized by fractures, mental retardation, skin lesions, immune system disorders, renal tubular acidosis. Clinical examination and review of radiographs, bone biopsy and genetic testing provide the bases for clinical diagnosis. The early and accurate detection and treatment of the disease are important to prevent hematologic abnormalities and disease progression to irreversible neurologic consequences. Most patients die within the first decade due to secondary infections, bone marrow suppression and/or bleeding. This article summarizes the current state of the art in this field, including clinical and genetic aspects, and the molecular pathogenesis of the osteopetrosis.
Key words: osteopetrosis; classification; connective tissue.

For citation: Nadyrshina D.D., Khusainova R.I. Clinical, genetic aspects and molecular pathogenesis of osteopetrosis. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):383-392. DOI 10.18699/VJGB-23-46

Клинико-генетические аспекты и молекулярный патогенез остеопетроза

Д.Д. Надыршина[1] ✉, Р.И. Хусаинова[1, 2]

[1] Уфимский университет науки и технологий, Уфа, Россия
[2] Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
✉ nadyrshinadina@gmail.com

Аннотация. Остеопетроз («мраморная кость», МКБ-10-78.2) включает группу наследственных нарушений костной системы, отличающихся клинической вариабельностью и генетической гетерогенностью. Название «остеопетроз» происходит от греческих слов: ὀστέον (остео) – кость, πέτρα (петросис) – камень, что характеризует основной признак заболевания – повышенную плотность костей, обусловленную нарушениями равновесия формирования и ремоделирования кости, приводящими к структурным изменениям в костной ткани, предрасположенности к переломам, деформациям скелета. Эти дефекты, в свою очередь, влияют на другие важные органы и ткани, особенно на костный мозг и нервную систему. Заболевание наследуется по аутосомно-рецессивному, аутосомно-доминантному типам, встречаются Х-сцепленные формы заболевания, а также спорадические случаи. Частота аутосомно-доминантного остеопетроза составляет 1 на 20 тыс., а аутосомно-рецессивного – 1 на 250 тыс. новорожденных. На сегодняшний день описано 23 гена, структурные изменения в которых приводят к развитию остеопетроза. Клинические симптомы при остеопетрозных состояниях сильно различаются по проявлению и степени тяжести. Наиболее легкие скелетные нарушения наблюдаются во взрослом возрасте и встречаются при аутосомно-доминантной форме остеопетроза. Тяжелые формы, характеризующиеся переломами, умственной отсталостью, поражениями кожи, нарушениями иммунной системы, ацидозом почечных канальцев, наблюдаются при аутосомно-рецессивном типе остеопетроза, проявляющемся в раннем детском возрасте. Диагноз «остеопетроз» ставится на основании клини-

ческой и рентгенологической оценки, подтвержденной биопсией костей и генетическим тестированием. Для аутосомно-рецессивных форм остеопетроза важны ранняя диагностика и лечение заболевания с целью установления гематологических нарушений, а также для предотвращения прогрессирования заболевания до возникновения необратимых неврологических последствий. Большинство пациентов умирают в течение первого десятилетия из-за вторичных инфекций и/или кровотечений, а также угнетения функции костного мозга. В настоящей работе представлен обзор современного состояния изучения остеопетроза, клинико-генетических аспектов, молекулярного патогенеза заболевания.
Ключевые слова: остеопетроз; классификация; соединительная ткань.

Introduction

Bone is a dynamic tissue that undergoes constant self-renewal; bone tissue homeostasis depends on the functional balance between three cell types: osteoclasts necessary for bone resorption; osteoblasts responsible for bone matrix formation, and osteocytes involved in the reception and transduction of mechanical stimuli and in the regulation of osteoclast/osteoblast differentiation and function. The balance between bone synthesis and resorption is finely tuned and any perturbations of this balance in adults trigger bone disease (Coudert et al., 2015).

Osteopetrosis is a group of inherited metabolic bone diseases characterized by increased bone mass due to defects in osteoclast function or formation, leading to fractures, generalized osteosclerosis, pancytopenia, and in severe cases, cranial neuropathies and hepatosplenomegaly. Abnormalities in the structural organization of multiple genes are responsible for the development of the disease, leading to marked clinical heterogeneity.

Classification and clinical features of osteopetrotic conditions

In 2006, the Nosology Group of the International Skeletal Dysplasia Society presented a classification of increased bone density conditions into several distinct entities based on clinical features, mode of inheritance and underlying molecular and pathogenetic mechanisms (Stark, Savarirayan, 2009). 13 clinical forms of osteopetrosis were identified: severe neonatal or infantile forms of osteopetrosis, the intermediate form of osteopetrosis with renal tubular acidosis, the late form (Albers-Schönberg disease), osteopetrosis with ectodermal dysplasia and immune defect (OLEDAID), leukocyte adhesion deficiency syndrome (LAD-III) and osteopetrosis, pycnodisostosis, osteopoikilosis, melorheostosis with osteopoikilosis, dysosteosclerosis, osteomesopiknosis, congenital striated osteopathy with cranial stenosis, Stanescu-type osteosclerosis (Stark, Savarirayan, 2009).

Recently, the advent of next-generation sequencing technology has continued to identify new molecular causes of the disease, leading to an expansion of the classification. We have systematized all currently known osteopetrosis conditions with a description of the genetic defect and the main clinical characteristics, and showed it in Table.

Autosomal recessive, intermediate forms of osteopetrosis, and autosomal dominant and X-linked type of osteopetrosis have been described.

Autosomal dominant osteopetrosis (Albers-Schönberg disease, ADO) is commonly referred to as benign osteopetrosis, with an incidence of 1 : 20,000 newborns. Clinical and radiologic signs of ADO most often appear late in childhood or adolescence. The major complications are skeletal, including fractures, scoliosis, osteoarthritis of the hip joint, and osteomyelitis, especially affecting the mandible in combination with dental abscesses or dental caries. About 70 % of patients with ADO have mutations in the *CLCN7* gene. In the remaining ~30 % of cases, no mutations in the *CLCN7* gene sequences were found, suggesting involvement of additional genes in the pathogenesis of this form of osteopetrosis (Coudert et al., 2015).

The X-linked form of osteopetrosis results from mutations in the *IKBKG* gene. Patients with this type have a specific phenotype with ectodermal dysplasia and an immune defect.

Intermediate forms of osteopetrosis are caused by defects in the *PLEKHM1* and *SNX10* genes. Clinical manifestations vary in patients with this type.

The most severe malignant disease states are **autosomal recessive forms of osteopetrosis (ARO)**, which are caused by defects in various genes with products that are involved in the formation, function and differentiation of osteoclasts. Clinically, patients with ARO are characterized by severe disorders of the musculoskeletal system, central nervous system (CNS), manifesting in the first few months of life. Patients are treated in pediatric or hematologic departments. Sick children have recurrent infections. They also suffer from frequent bleeding secondary to medullary hyperplasia caused by bone invasion of the medullary space. Cranial nerve compression can lead to blindness and deafness. Neurologic defects may also occur in some patients regardless of nerve compression. Radiological examination reveals dense bones characterized by extreme fragility.

Currently, there is no universal therapy regimen for osteopetrosis and strategies and tactics in the treatment of the disease are determined by its molecular pathogenesis, so it is necessary to identify the genetic cause of the disease in each individual case. Hematopoietic stem cell transplantation (HSCT) is recognized as the most effective treatment, which allows restoration of bone resorption by cells of donor origin. This therapy is suitable for patients with mutations in the *TCIRG1*, *TNFRSF11A* (*RANK*), *SNX10*, *CAII*, *IKBKG*, *FERMT3*, *CalDAG-GEF1* genes. Patients with severe neurological disorders caused by mutations in the *TNFSF11* and *OSTM1* genes are not indicated for HSCT

Clinical and genetic classification of osteopetrosis conditions

Condition	Inheritance	Gene	Protein	Clinical features
Osteopetrosis, severe neonatal or infantile forms	AR*	*TCIRG1*	Subunit of V-ATPase pump	Reminiscent of ADO type 2, rickets, osteomalacia Delayed development of the central nervous system
	AR	*CLCN7*	Chloride channel	Fractures, short stature, base sclerosis with optic nerve compression, facial nerve palsy and hearing loss, absence of medullary cavity with severe anemia and thrombocytopenia, dental anomalies, mandibular osteomyelitis, hypocalcemia and secondary hyperparathyroidism
	AR	*OSTM1*	Osteopetrosis associated transmembrane protein	Central nervous system damage. Abnormalities of the brain
	AR	*TNFSF11* (*RANKL*)	TNF superfamily member 11	T-cell defect, hypogammaglobulinemia, similar to variant immunodeficiency
	AR	*TNFRSF11A* (*RANK*)	TNF receptor superfamily member 11	Visual impairments, neurological defects, fractures, physical and mental retardation
	AR	*SNX10*	Sorting nexin 10	Stunting, hypocalcemia, hydrocephalus, severe visual and hematopoietic disorders
	AR	*PLEKHM1*	Pleckstrin homology domain containing family M, member 1	Frequent fractures, deformities of different parts of the hip, bone pain
Osteopetrosis with renal tubular acidosis	AR	*CAII*	Carbonic anhydrase II	Cerebral calcification and renal tubular acidosis
Osteopetrosis, late-onset form (Albers-Schönberg disease)	AD	*CLCN7*	Chloride channel	Fractures, scoliosis, osteoarthritis of the hip joint and osteomyelitis of the mandible or septic osteitis or osteoarthritis elsewhere. Cranial nerve compression (rare)
Osteopetrosis with ectodermal dysplasia and immune defect (OLEDAID)	XL	*IKBKG* (*NEMO*)	Inhibitor of kappa light polypeptide gene enhancer, kinase of	Hypohydrotic ectodermal dysplasia, congenital dental, hair and extracellular glandular diseases, immunodeficiency, increased susceptibility to fungal infections
Leukocyte adhesion deficiency syndrome (LAD-III) and osteopetrosis	AR	*FERMT3*	Kindlin-3	Recurrent infections and hemorrhagic diathesis regardless of platelet or leukocyte count, high bone density
	AR	*CalDAG-GEF1*	Calcium and diacylglycerol-regulated guanine nucleotide exchange factor 1	Frequent bleeding and recurrent infections
Pycnodysostosis	AR	*CTSK*	Cathepsin K	Short stature, typical facial appearance (convex back of nose and small jaw with blunt mandibular angle), osteosclerosis with bone fragility, acroosteolysis of distal phalanges, delayed closure of cranial sutures, clavicle dysplasia, dental anomalies, joint hypermobility, cerebral demyelination and hepatosplenomegaly
Osteopoikilosis	AD	*LEMD3*	LEM domain-containing 3	Sclerotic darkening of the sciatic, pubic bones and epimetaphyseal areas of the short tubular bones
Melorheostosis with osteopoikilosis	AD	*LEMD3*	LEM domain-containing 3	Pain and stiffness in extremities, skeletal deformities

Table (end)

Condition	Inheritance	Gene	Protein	Clinical features
Dysosteosclerosis	AR	*SLC29A3*, *TNFRSF11A*	Transmembrane glycoprotein TNF receptor superfamily member 11a	Red-violet macular atrophy, platyspondylitis, metaphyseal osteosclerosis
Osteomesopyknosis	AD	Unknown	Unknown	Back pain, axial sclerosis
Osteopathia striata congenita with cranial stenosis	XL	*WTX* (*AMER1*)	Wilm's tumor gene on the X chromosome	Longitudinal stripping of metaphyses of long bones, macrocephaly, cleft palate and hearing loss, mental retardation
Osteosclerosis, Stanescu type	AD	Unknown	Unknown	Short stature, sclerosis of long bones, malformations of the skull bones
COMMAD	AR	*MITF*	Melanocyte inducing transcription factor	Coloboma, osteopetrosis, microphthalmia, macrocephaly, albinism, deafness
Poikiloderma with Neutropenia	AR	*C16orf57*	Phosphodiesterase	Osteopetrosis, neutropenia
Generalized osteopetrosis with severe cerebral malformation	AR	*CSF1R*	Colony-stimulating factor 1 receptor	Structural abnormalities of the brain and primary neurodegenerative phenotype, platyspondylosis
Severe combined immunodeficiency	AR	*RAG1*, *RAG2*, *TRAF6*	Recombination activating 1, recombination activating 2, TNF receptor associated factor 6	Hypocalcemia, optic atrophy, increased pelvic bone density, and rickety changes in knees and wrists
Osteosclerotic metaphyseal dysplasia	AR	*LRRK1*	Leucine rich repeat kinase 1	Osteopetrosis, mainly affecting metaphyses of long bones, vertebral endplates, rib ends, and edges of flat bones
Rela associated osteopetrosis	AR	*RELA*	NF-κB nuclear transcription factor subunit	Congenital osteopetrosis, osteosclerosis

* Mode of inheritance: AD – autosomal-dominant, AR – autosomal-recessive.

and only symptomatic treatment is available for these patients. Risky and invasive transplant procedures are also not indicated for mild forms of the disease, for example, with mutations in the *PLEKHM1*, *SLC29A3*, *CTSK*, or *CLCN7* genes. For recently diagnosed forms of the disease, there is currently insufficient knowledge to determine specific treatment tactics.

Thus, osteopetrosis is a clinically variable disease with a wide spectrum of clinical manifestations and symptoms of varying severity. It is necessary to understand the molecular pathogenesis of the disease in order to correctly diagnose and determine the treatment tactics of the disease.

Molecular pathogenesis of different forms of osteopetrosis

The disease is characterized by a complex molecular pathogenesis caused by mutations in 23 genes (see the Table) responsible for the development of corresponding clinical osteopetrosis conditions (*TCIRG1*, *CLCN7*, *OSTM1*, *PLEKHM1*, *SNX10*, *TNFSF11* (*RANKL*), *TNFRSF11A* (*RANK*), *IKBKG* (*NEMO*), *RAG1*, *RAG2*, *TRAF6*, *FERMT3*, *LRRK1*, *MITF*, *C16orf57*, *CSF1R*, *CAII*, *SLC29A3*, *CalDAG-GEF1*, *CTSK*, *WTX*, *LEMD3*, *RELA*).

Autosomal recessive forms of osteopetrosis arise from mutations in genes that are involved in osteoclast function (osteoclast-rich) or differentiation (osteoclast-poor forms of osteopetrosis).

Osteoclast-rich osteopetrosis is caused by mutations in genes responsible for lacunar acidification, resorption and pH regulation (*TCIRG1*, *CLCN7*, *OSTM1* and *CAII*), vesicular transport and sorting of protein complexes to the membrane (*SNX10* and *PLEKHM1*), lysosomal nucleoside transport (*SLC29A3*) cytoskeletal rearrangement for "corrugated edge" formation (*KINDLIN3*, integrin-β and *LRRK1*) and lysosomal proteolytic cleavage for bone remodeling and resorption (*CTSK*), for signal transduction and osteoclast function (*MITF*, *TRAF6*, *RELA* and *NEMO*) (De Cuyper et al., 2021; Penna et al., 2021).

In osteopetrosis with osteoclast deficiency, osteoclast differentiation is impaired due to mutations in the *TNFSF11* and *TNFRSF11A* genes encoding RANKL and its receptor RANK, respectively, or in the *CSF1R* gene encoding

M-CSF. As a consequence, osteoclast precursors are unable to fuse and differentiate into multinucleated resorbing osteoclasts.

About 50 % of patients with ARO have mutations in the *TCIRG1* (T-cell immunoregulator 1) gene. This gene encodes a subunit of a large protein complex known as vacuolar H+-ATPase (V-ATPase), mainly expressed by osteoclasts and gastric parietal cells on apical membrane. The protein complex acts as a pump to move protons across the membrane. The V-ATPase pump acidifies the resorption lacuna in the bone for the dissolution of the hydroxyapatite crystals that form the bone mineral fraction and the degradation of the matrix.

The a3 V-ATPase subunit is also involved in the interaction between the actin cytoskeleton and microtubules, necessary for the osteoclast ruffled border formation (corrugated paper). Accordingly, *TCIRG1*-mutated osteoclasts show defective ruffled border and markedly reduced resorptive activity. In addition, V-ATPase maintains low pH in the stomach for the dietary Ca^{2+} absorption, and because gastric acidification is also relevant for calcium uptake, this form of osteopetrosis is characterized by rickets or osteomalacia (Penna et al., 2019).

To date, more than 120 different mutations in the *TCIRG1* gene have been described, including missense mutations, nonsense mutations, small insertions/deletions, large genomic deletions, and splicing defects, demonstrating the high genetic heterogeneity of the *TCIRG1*-deficient ARO cohort (Palagano et al., 2018).

Mutations in the *CLCN7* (chloride potential-dependent channel 7) gene are responsible for 17 % of autosomal recessive osteopetrosis cases and for the majority of autosomal dominant osteopetrosis cases (70 %) (Penna et al., 2021). Bi-allelic mutations cause a very severe form of the disease in which bone defects and hematologic failure are combined in some patients with primary neurodegeneration resembling lysosomal accumulation disease, cerebral atrophy, spasticity, axial hypotonia, and peripheral hypertension. Conversely, single-allelic *CLCN7* mutations result in autosomal dominant osteopetrosis and are associated with milder symptoms and later onset.

The *CLCN7* gene encodes $2Cl^-/H^+$-antiporters, regulated by a potential-dependent mechanism, expressed on the "corrugated edges" of osteoclasts and on the membranes of late endosomes and lysosomes. CLC family proteins transport chlorine ions across cell membranes to maintain membrane potential, regulate transepithelial Cl^- transport, and control intravesical pH between different organelles.

The neuropathic form of autosomal recessive osteopetrosis is caused by mutations in the *OSTM1* or *CLCN7* genes. Mutations in *OSTM1* (transmembrane protein 1 associated with osteopetrosis) account for about 5 % of ARO cases and invariably cause osteopetrosis and severe primary neurodegeneration with a life expectancy of less than two years. OSTM1 acts as the auxiliary β-subunit of CLC-7 to support bone resorption and lysosomal function.

Virtually all of the identified mutations in this gene result in protein shortening. A secreted form of shortened OSTM1 has been shown to inhibit osteoclast formation *in vitro* through suppression of the BLIMP1-NFATc1 axis, thereby providing a putative additional pathogenetic mechanism of OSTM1-deficient ARO. Moreover, using a specially designed quantitative PCR strategy, two different homozygous microdeletions spanning ~110 and ~10 bp, respectively, and affecting the N-terminal part of the *OSTM1* gene were detected in two unrelated families of Arab and Indian origin consisting of five critically ill patients. Sequence analysis of the relevant genomic region identified AluSx-mediated recombination and nonrecurrent rearrangement followed by nonhomologous end joining as the respective underlying molecular mechanism (Palagano et al., 2018; Zhang et al., 2020).

Osteopetrosis with early onset neurodegeneration and iron accumulation in certain brain regions has been described in one patient, which is a very unusual finding. Full-exome sequencing revealed the presence of a novel c.783+5G>T mutation in the *OSTM1* gene, causing exon 4 skipping, and a frameshift variant c.446dup in the homozygous state in the *MANEAL* gene. This gene encodes an endo-alpha-like mannosidase protein, which probably localizes in the Golgi complex and is potentially involved in glycoprotein metabolism; indeed, increased mannose tetrasaccharide molecules have been found in the patient's urine and cerebrospinal fluid. How this might be related to iron accumulation in the brain and the contribution of a mutation in the *MANEAL* gene to the formation of the osteopetrosis phenotype requires further investigation.

Osteopetrosis with renal tubular acidosis and cerebral calcinosis is caused by mutations in the *CAII* gene. Carboanhydrase (CAII) is a zinc-containing metalloenzyme responsible for catalyzing the reversible conversion of carbon dioxide (CO_2) and water (H_2O) to bicarbonate ($HCO3^-$) and protons (H^+). Carboanhydrase helps in the maintenance of homeostasis in the body. The substrates and products of the reaction (CO_2, $HCO3^-$ and H^+) are necessary for the regulation of biological processes such as respiration, cerebrospinal fluid formation, and bone resorption (Sanyanga et al., 2019).

About 30 different mutations have been identified in the *CAII* gene: missense mutations, nonsense mutations, and splice site mutations. The majority of patients with this mutation are of Arabian origin.

Intermediate forms of ARO caused by mutations in the *PLEKHM1* (member 1 of the M family containing the pleckstrin homology domain) and *SNX10* (sorting nexin 10) genes have been described (Coudert et al., 2015). The *PLEKHM1* gene encodes a cytosolic protein involved in endosome transport pathways through interaction with small GTPases RAB7 and ARL8. In addition, PLEKHM1

is involved in the fusion of autophagosomes and lysosomes required for the clearance of a variety of protein aggregates. Accordingly, disruption of specific domains of this protein or its loss impairs vesicle distribution, secretion, and formation of corrugated wukras, thereby undermining the resorptive function of osteoclasts. PLEKHM1 is a large protein containing various functional domains: the RUN domain in which the c.296+1G>A mutation was localized, originally identified in two siblings with ARO; two plectrin homology (PH) domains separated by an LC3-interacting region (LIR); the Rubicon homology (RH) domain and the C1 zinc finger at the C-terminal.

Two different presumably dominant mutations in the *PLEKHM1* gene have been reported in two unrelated patients: c.2140C>T (p.Arg714Cys), clearly unrelated to osteopetrosis, was found in the second PH domain; and the recently discovered c.3051_3052delCA mutation, located in the RH domain, is predicted to eliminate the zinc finger motif. The RH domain is essential for the interaction of PLEKHM1 with RAB7, leading to reduced interaction of the mutant protein with RAB7, resulting in abnormal intracellular localization and increased autophagy.

Less than 5 % of ARO cases are caused by mutations in the *SNX10* (sorting nexin 10) gene, which encodes a protein family of cytoplasmic and membrane-bound proteins characterized by a phosphoinositide-binding domain called the PX domain (Zhou et al., 2017). SNX proteins take part in protein sorting and transport across membranes by establishing protein-protein and protein-lipid interactions. Specifically, SNX10 interacts with V-ATPase and regulates its intracellular transport; accordingly, this autosomal recessive form of osteopetrosis results from altered transport of V-ATPase to the "corrugated edges" of osteoclasts and, consequently, their defective function. It has been suggested that *SNX10* plays a role in the delivery and secretion of matrix metalloprotease 9, which is involved in the degradation of the extracellular matrix (Palagano et al., 2018).

Mutations in the *SNX10* gene cause Västerbottenian osteopetrosis (named after the Swedish county), where the c.212+1G>T mutation in the *SNX10* gene causing activation of a hidden splicing site in intron 4, leading to frameshift and stop codon formation (p.S66Nfs*15), occurs with a frequency of 1 : 93 in the population of this region. Genealogical studies and haplotype analysis have traced the origin of this mutation to a common ancestor in the early 19th century, and the age of the mutation is estimated to be approximately 950 years (Pangrazio et al., 2013; Stattin et al., 2017).

2 % of patients with ARO are deficient in the cytokine RANKL (receptor-activator nuclear kappa-B ligand) and 4.5 % are deficient in its receptor RANK. RANKL is encoded by the *TNFSF11* gene, and binding to its receptor RANK, encoded by the *TNFRSF11A* gene, determines the activation of a downstream pathway that controls osteoclast differentiation and activation. The RANK/RANKL signaling pathway regulates the formation of mature osteoclasts from their precursors as well as their activity in bone remodeling. Disruption of this pathway results in a complete absence of mature osteoclasts in bone biopsy specimens. Patients with RANKL deficiency show severe osteopetrosis with slower disease progression compared to classical ARO (Penna et al., 2021).

In bone, RANKL is produced mainly by the stromal compartment under physiological conditions, whereas other cell sources are more important in pathological processes. Recent evidence suggests that RANKL also plays an osteogenic role through an autocrine loop in mesenchymal stem cells and through reverse signaling from osteoclasts to osteoblasts. In addition, in patients, the absence of RANKL leads to a partial disruption of T-cell proliferation and cytokine production, while RANK deficiency impairs B-cell memory subpopulation and immunoglobulin production (Penna et al., 2021).

Importantly, unlike TNFSF11 deficiency, osteopetrosis in patients with TNFRSF11A deficiency can be rescued by hematopoietic stem cell transplantation.

X-linked osteopetrosis is caused by mutations in the *IKBKG* gene. The *IKBKG* gene encodes NEMO, a regulatory subunit of the IKK complex (inhibitor of κB kinases) fundamental for the activation of the NF-κB (nuclear factor κB) transcription factor for the induction of osteoclastogenesis. NF-κB signal transduction involves a number of molecules (mainly kinases and transcription factors) that play a crucial role in the regulation of gene expression in many organs and in physiopathological conditions. In bone, it is supported by the fact that hypomorphic mutations in the *IKBKG* gene encoding a component of the IκB kinase complex required for inhibition of IκB-α and subsequent nuclear translocation of the released p65/p50 heterodimer are responsible for X-linked osteopetrosis with ectodermal dysplasia and immunodeficiency. These mutations are mainly localized in the zinc finger protein domain and lead to osteopetrosis by altering the RANKL/RANK signaling pathway (Frost et al., 2019; Jimi, Katagari, 2022).

Recently, a case was described in a newborn infant who died suddenly of unknown causes and pathological examination revealed a pathological increase in bone density associated with increased osteoblast function caused by *de novo* (c.1534_1535delinsAG (p.Asp512Ser)) mutation in the *RELA* gene (11q13.1). This mutation has been shown to disrupt NF-κB signaling in patient fibroblasts, which supports the hypothesis of possible changes in various vital functions (Frederiksen et al., 2016).

Severe combined immunodeficiency (SCID) is caused by a large deletion on chromosome 11 spanning the *RAG1* and *RAG2* genes and the 5′-region of TRAF6 (Weisz Hubshman et al., 2017).

Among the various adaptor molecules recruited via RANKL/RANK binding, TRAF6 (TNF receptor-associated factor 6) appears to be the most important. TRAF6 also

acts downstream of the T- and B-cell receptor, leading to NF-κB activation.

Several years ago, inactivation of the *TRAF6* gene in mice was shown to cause severe osteopetrosis, and more recently, similar evidence was obtained in humans. In fact, a homozygous 2064 bp genomic deletion on chromosome 11 covering the 5′-region of the *TRAF6*, *RAG1* and *RAG2* genes (RAG proteins are necessary for B- and T-cell receptor recombination and for the survival and differentiation of these cells) was identified in two sibling patients with osteopetrosis and severe combined immune deficiency (SCID) by chromosomal microarray analysis. This genomic deletion covers the region above exon 1 and part of the non-coding sequences of exon 1. It is likely that these regions are regulatory; in fact, at the protein level, their deletion completely abolishes TRAF6 production. This mutation has been described in a single family, and the osteopetrosis was not generalized, but was pronounced in the pelvis and legs; because both patients, brother and sister, died at a very young age due to a severe immunological defect, it is currently difficult to predict the evolution of the disease in this particular case (Weisz Hubshman et al., 2017).

Mutations in the *FERMT3* and *CALDAGGEF1* genes cause osteopetrosis combined with leukocyte adhesion deficiency type III (LAD III).

The *CALDAGGEF1* gene lies at the distal edge of the region of chromosome 11q13.1, is activated through diacylglycerol and Ca^{2+} binding, and is a guanine replacement factor for Rap1, a GTPase that plays an essential role in integrin activation. The gene encodes two proteins by alternative splicing, a cytosolic form of 68-kDa and a form of 72-kDa localized in the membrane through an additional amino-terminal myristoylated and palmitoylated domain (Svensson et al., 2009).

The *FERMT3* gene (chromosome 11: 63.73–63.75 Mb) is located 0.5 Mb from *CALDAGGEF1* on chromosome 11q13.1. The *FERMT3* gene (representative of fermitin family 3) is expressed in hematopoietic cells and codes for kindlin-3, a member of the kindlin family that includes three different focal adhesion proteins involved in integrin activation. This process is necessary for cell adhesion, proliferation and migration, organization of the extracellular matrix, cell survival, proliferation, and differentiation.

Kindlin-3 is an intracellular protein bound to the actin cytoskeleton. It interacts with several classes of integrins and mediates their adhesive function and the transmission of signals from inside to outside, which is essential in bone for the resorptive activity of osteoclasts. Accordingly, kindlin-3 deficiency causes a major morphological change in osteoclasts and impairs their ability to attach to the bone surface. Mutations with a prematurely-terminating codon have been mainly described: nonsense mutations, splice defects, frameshifts, and, very rarely, missense mutations. Unfortunately, because the number of cases published in the literature is very limited, a gene-phenotypic correlation cannot be made at this time (Svensson et al., 2009).

Mutations in the *LRRK1* (leucine-rich repeat kinase 1) gene are responsible for osteosclerotic metaphyseal dysplasia. The *LRRK1* gene consists of 34 exons spanning about 150 bp on chromosome 15q26.3. *LRRK1* encodes a multidomain protein of 2015 amino acids that contains ankyrin repeats, leucine-rich repeats), a C-terminal Roc (COR) domain and a serine-threonine kinase domain, and seven tryptophan-aspartic acid (WD) 40-domain dipeptides at the C-terminal.

Mutations of the *LRRK1* gene have been described in only five patients; a homozygous seven-nucleotide deletion in the last exon of the gene (c.5938_5944delGAGTGGT, p.Glu1980Alafs*66) was recently identified in one of these patients. This mutation is predicted to cause frameshift and premature termination with loss of the seventh tryptophan-aspartic acid (WD) 40 domain. The WD40 domain, like other functional domains in the LRRK1 protein, mediates protein-protein interactions. In particular, it has been suggested that LRRK1 interacts with components of the c-Src signal transduction pathway to achieve cytoskeleton and "corrugated edge" rearrangement and podosome assembly. Accordingly, LRKK1-deficient osteoclasts are flat and large because they are unable to properly reorganize the cytoskeleton and resorb bone (Iida et al., 2016; Xing et al., 2017).

Another gene associated with osteopetrosis is *MITF* (microphthalmic-associated growth factor), which encodes a transcription factor that acts downstream of the RANK/RANKL pathway. *MITF* deficiency is responsible for the COMMAD syndrome (Coloboma, Osteopetrosis, Microphthalmia, Macrocephaly, Albinism, and Deafness).

Microphthalmia-associated transcription factor (MITF) is a major helix-loop-helix-zipper transcription factor that forms homo/heterodimers that regulate gene expression in various tissues, so a range of phenotypes can reasonably be expected when it is mutated. In bone, MITF is thought to act along the RANKL/RANK signaling pathway downstream of NFATc1 to enhance NFATc1-dependent osteoclastogenic signaling.

Complex heterozygous mutations in the *MITF* gene have most recently been found in two unrelated patients with COMMAD syndrome manifesting coloboma, osteopetrosis, microphthalmia, macrocephaly, albinism and deafness. The identified mutations (c.952_954delAGA (p.Arg318del) and c.921G>C (p.Lys307Asn) in proband I; c.952A>G (p.Arg318Gly) and c.938-1G>A (p.Leu312fs*) in proband II) do not alter MITF dimerization, but rather its nuclear migration and DNA binding properties. This finding broadens the spectrum of phenotypes defined by MITF; in fact, unlike recessive mutations, dominant mutations are associated with Waardenburg type 2A syndrome and Titz syndrome, which share the characteristics of deafness and pigmentation deficiency. Overall, these data support

an essential role of MITF in developmental processes as well as in cell differentiation and survival (George et al., 2016).

Poikiloderma with neutropenia is an autosomal recessive genodermatosis caused by mutations in the *C16orf57* gene located on chromosome 16q21. To date, 17 mutations (deletions, nonsense mutations, and splice site mutations) have been identified in 31 patients with poikiloderma. The *C16orf57* gene encoding phosphodiesterase is responsible for the modification and stabilization of small nuclear RNA U6 (USB1), which is an important element of the splicing mechanism (Colombo et al., 2012; Larizza et al., 2013).

Generalized osteopetrosis with severe cerebral malformation has been reported in consanguineous patients with mutations in the *CSF1R* gene who had osteopetrosis and cerebral malformations. The *CSF1R* gene encodes the M-CSF (macrophage colony-stimulating factor) receptor, which is a key transmembrane tyrosine kinase receptor that modulates microglial homeostasis, neurogenesis and neuronal survival in the CNS. CSF1R, which can be proteolytically cleaved into a soluble ectodomain and an intracellular protein fragment, supports myeloid cell survival when activated by two ligands, colony-stimulating factor 1 and interleukin 34 (Hu et al., 2021).

M-CSF is an important osteoclastogenic molecule, as well as RANKL, and it is well demonstrated in osteopetrotic mice with osteoclast deficiency lacking this cytokine. M-CSF receptor deficient mice (CSF1R) show a similar osteopetrotic phenotype; in addition, both models have defects in innate immunity, fertility and neurological function. Interestingly, dominant mutations in the *CSF1R* gene cause the adult form of encephalomyopathy, whereas just recently, a recessive mutation in this gene was thought to be responsible for the lethal complex phenotype in two siblings with generalized osteopetrosis and severe cerebral malformation. Exome sequencing in the blood parents of the deceased children revealed a heterozygous mutation (c.1620C>T (p.Tyr540*)) in the *CSF1R* gene that is predicted to result in a protein lacking the intracellular domain that is required for ligand-dependent dimerization and autophosphorylation. In the absence of a patient DNA sample, homozygosity for the *CSF1R* mutation has not been demonstrated in sick patients; therefore, these conclusions were not definitive. However, it would be interesting to analyze the gene in other patients with a similar phenotype trying to identify additional mutations as confirmation.

A rare form of osteopetrosis with low osteoclast content, called dysosteosclerosis, accompanied by red-purple macular atrophy, platyspondylitis, and metaphyseal osteosclerosis, is caused by mutations in the *SLC29A3* gene (member 3 of the 29 solute carrier family), which codes for a highly expressed lysosomal nucleoside carrier in myeloid cells. The described mutations c.607T>C (p.Ser203Pro), c.1157G>A (p.Arg386Gln), c.1346C>G (p.Thr449Arg), c.303_320dup (p.102_107dup) identified in the *SLC29A3* gene affect osteoclast function and differentiation, as suggested by reduced osteoclast numbers after *in vitro* differentiation from patient peripheral blood mononuclear cells and in patient bone biopsy samples (Palagano et al., 2018). More recently, a new splice site mutation in intron 6 of the *TNFRSF11A* gene has been described in one patient, indicating that *TNFRSF11A* is an additional gene responsible for dyosteosclerosis.

Osteopoikylosis, Buschke–Ollendorff syndrome, and melorheostosis are benign and more often asymptomatic conditions of osteopetrosis, diagnosed more often radiologically, and caused by mutations in the *LEMD3* gene. LEMD3 is an integral protein of the inner nuclear membrane. It contains a nucleoplasmic N- and C-terminal domain and two helical transmembrane segments. The N-terminal segment shares a conserved globular domain of approximately 40 amino acids with other inner nuclear membrane proteins such as lamina-associated polypeptide 2 (LAP2) and emerin. The coding protein functions to counteract transforming growth factor-beta signaling at the inner nuclear membrane (Hellemans et al., 2004).

The gene responsible for pycnodysostosis is *CTSK*, located on chromosome 1 (1q21), encoding cathepsin K, a papain superfamily cysteine peptidase used by osteoclasts to degrade bone matrix and endowed with the unique ability to cleave collagen molecules in multiple sites. In addition, cathepsin K has recently been shown to cleave and activate matrix metalloproteinase 9 *in vitro*, indicating the presence of a protease signaling network likely significant in various physiopathological conditions. More recently, cathepsin K has been shown to contribute to the regulation of bone modeling by downregulating periostin, a cortical compartment matricellular protein required for Wnt-β-catenin-mediated periosteal formation (Pangrazio et al., 2014; Amr et al., 2021).

To date, about 60 different mutations have been described in the literature in patients of different geographic origins. Missense variants are the most frequent mutations; frameshifts, nonsense mutations, and splicing defects have also been identified. Mutations mainly occur in the mature CTSK protein, where exons 5 and 6 are "hot spots". In addition, about 6 % of mutations are mapped to the pre-region, and 25 %, to the pro-region, which are short N-terminal domains necessary for proper protein localization, protein folding, and intracellular transport, respectively; the pro-region is also necessary to keep the enzyme in an inactive state and is detached at low pH. However, genotype-phenotype correlations, which probably also explain atypical manifestations, have not been specifically investigated (Pangrazio et al., 2014).

Striated osteopathy with skull sclerosis is caused by mutations in the *WTX* gene (*AMER1*). This gene is located on chromosome Xq11.2 and contains 2 exons. The protein encoded by this gene enhances the activation of transcription by Wilms' tumor protein and interacts with many

other proteins. The prevalence of this form of osteopetrosis is 0.1:1,000,000 people (Jeoung et al., 2015). More than one hundred patients with this syndrome worldwide have been described, of which about one-third of the patients described are sporadic. Cranial sclerosis, in particular, is a clinically heterogeneous condition, ranging from mild skeletal manifestations to multisystem organ damage even within the same family.

Conclusion

Osteopetrosis is a clinically and genetically heterogeneous group of disorders the diagnosis of which is complicated by the presence of different clinical forms and types of inheritance and the absence of a clear correlation between genotype and phenotype. Moreover, the mutations identified to date explain only 70 % of cases of osteopetrosis. The search for the molecular defects responsible for the remaining 30 % of the disease continues.

The study of osteopetrosis is necessary for DNA diagnosis, treatment prescription, and prognosis. The study of osteopetrosis has shed light on little-known aspects of bone tissue cell biology and identified new mechanisms of osteoclast differentiation and function.

References

Amr K.S., El-Bassyouni H.T., Abdel Hady S., Mostafa M.I., Mehrez M.I., Coviello D., El-Kamah G.Y. Genetic and molecular evaluation: Reporting three novel mutations and creating awareness of pycnodysostosis disease. *Genes*. 2021;12(10):1552. DOI 10.3390/genes12101552.

Colombo E.A., Bazan J.F., Negri G., Gervasini C., Elcioglu N.H., Yucelten D., Altunay I., Cetincelik U., Teti A., Del Fattore A., Luciani M., Sullivan S.K., Yan A.C., Volpi L., Larizza L. Novel *C16orf57* mutations in patients with Poikiloderma with Neutropenia: bioinformatic analysis of the protein and predicted effects of all reported mutations. *Orphanet J. Rare Dis*. 2012;7:7. DOI 10.1186/1750-1172-7-7.

Coudert A.E., de Vernejoul M.C., Muraca M., Del Fattore A. Osteopetrosis and its relevance for the discovery of new functions associated with the skeleton. *Int. J. Endocrinol*. 2015;2015:372156. DOI 10.1155/2015/372156.

De Cuyper E., De Cuyper C., Willems L., Casselman J., Dhooge I., Van Hoecke H. Hearing loss in malignant infantile osteopetrosis: A case-based review. *J. Int. Adv. Otol*. 2021;17(6):551-558. DOI 10.5152/iao.2021.21266.

Frederiksen A.L., Larsen M.J., Brusgaard K., Novack D.V., Knudsen P.J., Schrøder H.D., Qiu W., Eckhardt C., McAlister W.H., Kassem M., Mumm S., Frost M., Whyte M.P. Neonatal high bone mass with first mutation of the NF-κB complex: Heterozygous *de novo* missense (p.Asp512Ser) *RELA* (Rela/p65). *J. Bone Miner. Res*. 2016;31(1):163-172. DOI 10.1002/jbmr.2590.

Frost M., Tencerova M., Andreasen C.M., Andersen T.L., Ejersted C., Svaneby D., Qui W., Kassem M., Zarei A., McAlister W.H., Veis D.J., Whyte M.P., Frederiksen A.L. Absence of an osteopetrosis phenotype in *IKBKG* (NEMO) mutation-positive women: A case-control study. *Bone*. 2019;121:243-254. DOI 10.1016/j.bone.2019.01.014.

George A., Zand D.J., Hufnagel R.B., Sharma R., Sergeev Y.V., Legare J.M., Rice G.M., Scott Schwoerer J.A., Rius M., Tetri L., Gamm D.M., Bharti K., Brooks B.P. Biallelic mutations in *MITF* cause coloboma, osteopetrosis, microphthalmia, macrocephaly, albinism, and deafness. *Am. J. Hum. Genet*. 2016;99(6):1388-1394. DOI 10.1016/j.ajhg.2016.11.004.

Hellemans J., Preobrazhenska O., Willaert A., Debeer P., Verdonk P.C., Costa T., Janssens K., Menten B., Van Roy N., Vermeulen S.J., Savarirayan R., Van Hul W., Vanhoenacker F., Huylebroeck D., De Paepe A., Naeyaert J.M., Vandesompele J., Speleman F., Verschueren K., Coucke P.J., Mortier G.R. Loss-of-function mutations in *LEMD3* result in osteopoikilosis, Buschke–Ollendorff syndrome and melorheostosis. *Nat. Genet.* 2004;36(11):1213-1218. DOI 10.1038/ng1453.

Hu B., Duan S., Wang Z., Li X., Zhou Y., Zhang X., Zhang Y.W., Xu H., Zheng H. Insights into the role of CSF1R in the central nervous system and neurological disorders. *Front. Aging Neurosci.* 2021;13:789834. DOI 10.3389/fnagi.2021.789834.

Iida A., Xing W., Docx M.K., Nakashima T., Wang Z., Kimizuka M., Van Hul W., Rating D., Spranger J., Ohashi H., Miyake N., Matsumoto N., Mohan S., Nishimura G., Mortier G., Ikegawa S. Identification of biallelic *LRRK1* mutations in osteosclerotic metaphyseal dysplasia and evidence for locus heterogeneity. *J. Med. Genet*. 2016;53(8):568-574. DOI 10.1136/jmedgenet-2016-103756.

Jeoung B.N., Kim J.M., Kang G.E., Lim J.H., Kim E.H., Seo H.A. A first case of osteomesopyknosis in Korea. *J. Bone Metab*. 2015;22(2):83-86. DOI 10.11005/jbm.2015.22.2.83.

Jimi E., Katagiri T. Critical roles of NF-κB signaling molecules in bone metabolism revealed by genetic mutations in osteopetrosis. *Int. J. Mol. Sci*. 2022;23(14):7995. DOI 10.3390/ijms23147995.

Larizza L., Negri G., Colombo E.A., Volpi L., Sznajer Y. Clinical utility gene card for: poikiloderma with neutropenia. *Eur. J. Hum. Genet*. 2013;21(10). DOI 10.1038/ejhg.2012.298.

Palagano E., Menale C., Sobacchi C., Villa A. Genetics of osteopetrosis. *Curr. Osteoporos. Rep*. 2018;16(1):13-25. DOI 10.1007/s11914-018-0415-2.

Pangrazio A., Fasth A., Sbardellati A., Orchard P.J., Kasow K.A., Raza J., Albayrak C., Albayrak D., Vanakker O.M., De Moerloose B., Vellodi A., Notarangelo L.D., Schlack C., Strauss G., Kühl J.S., Caldana E., Lo Iacono N., Susani L., Kornak U., Schulz A., Vezzoni P., Villa A., Sobacchi C. *SNX10* mutations define a subgroup of human autosomal recessive osteopetrosis with variable clinical severity. *J. Bone Miner. Res*. 2013;28(5):1041-1049. DOI 10.1002/jbmr.1849.

Pangrazio A., Puddu A., Oppo M., Valentini M., Zammataro L., Vellodi A., Gener B., Llano-Rivas I., Raza J., Atta I., Vezzoni P., Supertifurga A., Villa A., Sobacchi C. Exome sequencing identifies *CTSK* mutations in patients originally diagnosed as intermediate osteopetrosis. *Bone*. 2014;59:122-126. DOI 10.1016/j.bone.2013.11.014.

Penna S., Capo V., Palagano E., Sobacchi C., Villa A. One disease, many genes: Implications for the treatment of osteopetroses. *Front. Endocrinol. (Lausanne)*. 2019;10:85. DOI 10.3389/fendo.2019.00085.

Penna S., Villa A., Capo V. Autosomal recessive osteopetrosis: mechanisms and treatments. *Dis. Model. Mech*. 2021;14(5):dmm048940. DOI 10.1242/dmm.048940.

Sanyanga T.A., Nizami B., Bishop Ö.T. Mechanism of action of non-synonymous single nucleotide variations associated with α-carbonic anhydrase II deficiency. *Molecules*. 2019;24(21):3987. DOI 10.3390/molecules24213987.

Stark Z., Savarirayan R. Osteopetrosis. *Orphanet J. Rare Dis*. 2009;4:5. DOI 10.1186/1750-1172-4-5.

Stattin E.L., Henning P., Klar J., McDermott E., Stecksen-Blicks C., Sandström P.E., Kellgren T.G., Ryden P., Hallmans G., Lönnerholm T., Ameur A., Helfrich M.H., Coxon F.P., Dahl N., Wikström J., Lerner U.H. *SNX10* gene mutation leading to osteopetrosis with dysfunctional osteoclasts. *Sci. Rep*. 2017;7(1):3012. DOI 10.1038/s41598-017-02533-2.

Svensson L., Howarth K., McDowall A., Patzak I., Evans R., Ussar S., Moser M., Metin A., Fried M., Tomlinson I., Hogg N. Leukocyte adhesion deficiency-III is caused by mutations in *KINDLIN3* affecting

integrin activation. *Nat. Med.* 2009;15(3):306-312. DOI 10.1038/nm.1931.

Weisz Hubshman M., Basel-Vanagaite L., Krauss A., Konen O., Levy Y., Garty B.Z., Smirin-Yosef P., Maya I., Lagovsky I., Taub E., Marom D., Gaash D., Shichrur K., Avigad S., Hayman-Manzur L., Villa A., Sobacchi C., Shohat M., Yaniv I., Stein J. Homozygous deletion of RAG1, RAG2 and 5′ region TRAF6 causes severe immune suppression and atypical osteopetrosis. *Clin. Genet*. 2017;91(6):902-907. DOI 10.1111/cge.12916.

Xing W., Goodluck H., Zeng C., Mohan S. Role and mechanism of action of leucine-rich repeat kinase 1 in bone. *Bone Res.* 2017;5:17003. DOI 10.1038/boneres.2017.3.

Zhang S., Liu Y., Zhang B., Zhou J., Li T., Liu Z., Li Y., Yang M. Molecular insights into the human CLC-7/Ostm1 transporter. *Sci. Adv*. 2020;12;6(33):eabb4747. DOI 10.1126/sciadv.abb4747.

Zhou C., Wang Y., Peng J., Li C., Liu P., Shen X. *SNX10* plays a critical role in MMP9 secretion via JNK-p38-ERK signaling pathway. *J. Cell. Biochem*. 2017;118(12):4664-4671. DOI 10.1002/jcb.26132.

ORCID ID

D.D. Nadyrshina orcid.org/0000-0001-8981-688X
R.I. Khusainova orcid.org/0000-0002-8643-850X

Acknowledgements. This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation (No. 122041400169-2), with partial support from the megagrant of the Government of the Russian Federation (agreement No. 075-15-2021-595), and the project of St. Petersburg University (No. 94034528).

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

Received November 8, 2022. Revised December 23, 2022. Accepted December 24, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):393-400
DOI 10.18699/VJGB-23-47

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

Genetic markers of children asthma: predisposition to disease course variants

M.V. Smolnikova ✉, Ed.W. Kasparov, M.A. Malinchik, K.V. Kopylova

Scientific Research Institute of Medical Problems of the North – a separate division of the Federal Research Center "Krasnoyarsk Science Center" of the Siberian Branch of the Russian Academy of Sciences, Krasnoyarsk, Russia
✉ smarinv@yandex.ru

Abstract. Asthma is a heterogeneous and often difficult to treat condition that results in a disproportionate cost to healthcare systems. Children with severe asthma are at increased risk for adverse outcomes including medication-related side effects, life-threatening exacerbations, and impaired quality of life. An important therapeutic focus is to achieve disease control, which is supposed to involve a personalized approach to treatment of asthma of any severity. Asthma is a multifactorial disease with a significant genetic determinant, however, the inheritance of asthma has not been fully elucidated. Polymorphic genes of inflammatory mediators, including cytokines, play an important role in developing various disease forms. In the current study, large-scale original data on the prevalence of cytokine gene genotypes (*IL2*, *IL4*, *IL5*, *IL6*, *IL10*, *IL12*, *IL13*, *IL17A*, *IL31*, *IL33*, *IFNG*, *TNFA*) among Russian children with asthma in Krasnoyarsk region have been obtained. Genotyping was carried out using real-time PCR. We identified markers predisposing to the development of different variants of the course of childhood asthma: the CT genotype and T allele of *IL4* rs2243250 are associated with asthma ($p < 0.05$), especially in mild asthma and in controlled asthma. The TT genotype and allele T of *IL13* rs1800925 are associated with severe and uncontrolled asthma ($p < 0.05$). The AA genotype of *IL17A* rs2275913, the TT genotype of *IFNG* rs2069705 and allelic A variants of *TNFA* rs1800629 are associated with mild asthma, and the TT genotype of *IFNG* rs2069705 is additionally associated with controlled asthma. The results obtained will supplement information on the prevalence of polymorphic variants of the cytokine genes in the Russian population and in asthma patients with different disease courses, which is likely to be used in order to shape a plan for Public Health Authority to prevent the development of severe uncontrolled asthma and to optimize personalized therapy.
Key words: asthma; cytokine; gene polymorphism; child; asthma severity; level of diseases control.

For citation: Smolnikova M.V., Kasparov Ed.W., Malinchik M.A., Kopylova K.V. Genetic markers of children asthma: predisposition to disease course variants. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):393-400. DOI 10.18699/VJGB-23-47

Генетические маркеры бронхиальной астмы у детей: предрасположенность к вариантам течения заболевания

М.В. Смольникова ✉, Э.В. Каспаров, М.А. Малинчик, К.В. Копылова

Научно-исследовательский институт медицинских проблем Севера – обособленное подразделение Федерального исследовательского центра «Красноярский научный центр Сибирского отделения Российской академии наук», Красноярск, Россия
✉ smarinv@yandex.ru

Аннотация. Астма – хроническое гетерогенное и часто трудно поддающееся лечению состояние, приводящее к несоразмерным расходам системы здравоохранения. Дети с тяжелой астмой подвержены повышенному риску неблагоприятных исходов, включая побочные эффекты, связанные с приемом лекарств, угрожающие жизни обострения и ухудшение качества жизни. Важным терапевтическим акцентом является достижение контроля над заболеванием, что подразумевает персонифицированный подход к лечению при любой степени тяжести астмы. Астма относится к мультифакториальным заболеваниям, имеющим значимую генетическую детерминанту, однако наследование астмы на сегодняшний день полностью не объяснено. В развитии разных форм заболевания особую роль играют полиморфные гены медиаторов воспаления, в том числе цитокинов. В настоящем исследовании впервые получены масштабные данные о распределении генотипов генов цитокинов (*IL2*, *IL4*, *IL5*, *IL6*, *IL10*, *IL12*, *IL13*, *IL17A*, *IL31*, *IL33*, *IFNG*, *TNFA*) среди больных астмой русских детей Красноярского края. Генотипирование осуществлено с использованием метода полимеразной цепной реакции в режиме реального времени (ПЦР-РВ). В ходе исследования нами выявлены маркеры, предрасполагающие к развитию различных вариантов течения астмы у детей: генотип CT и аллель T rs2243250 *IL4* ассоциированы с развитием астмы ($p < 0.05$), особенно при легкой форме и контролируемом течении. Генотип TT и аллель T rs1800925 *IL13* ассоциированы с астмой, в том числе тяжелой степени, и с неконтролируемой формой ($p < 0.05$). Установлено, что генотипы AA *IL17A* rs2275913, TT *IFNG* rs2069705 и аллельный вариант A *TNFA* rs1800629 ассоциированы с

легкой степенью астмы, генотип TT *IFNG* rs2069705 ассоциирован также с контролируемой формой. Полученные результаты дополнят данные о характеристике распределения полиморфных вариантов генов цитокинов в русской популяции и у больных астмой с различным течением заболевания, что можно будет использовать для формирования планов органов практического здравоохранения в отношении профилактики развития тяжелой неконтролируемой астмы и в целях оптимизации персонифицированной терапии.
Ключевые слова: бронхиальная астма; цитокин; полиморфизм генов; дети; степень тяжести астмы; уровень контроля.

Introduction

Asthma is one of the most common diseases of the lower respiratory tract; it is a heterogeneous disease characterized by airway inflammation and hyperactivity. Asthma most often begins in early childhood, has a variable course and an unstable phenotype progressing over time (Hancox et al., 2012). It significantly limits and worsens the quality of human life in case of uncontrolled and severe disease. According to WHO estimates, asthma annually leads to the loss of 26.2 million in the world as measured by DALYs (disability-adjusted life years – an indicator of healthy life lost due to disability), which is 1 % of the total global burden of disease (GBD 2015 Chronic Respiratory Disease Collaborators…, 2017). Today, asthma is a global health problem of great socio-economic importance, i. e. about 339 million people worldwide suffer from asthma. The increase in the prevalence and incidence of asthma worldwide is influenced by both genetic background and a large number of environmental factors included in the "modern lifestyle" concept. Moreover, asthma prevalence, severity, and mortality vary greatly by ethno-geographic origin.

Based on expert estimates, the number of asthma patients in Russia exceeds official figures, i. e. according to their calculations, 5.9 million instead of 1.3 million people suffer from asthma in our country. In addition, according to the reported data, since asthma is a disabling and dangerous disease, about 41 % of asthma patients receive a disability pension. The prevalence of the disease among adults is 6–7 %, among children and adolescents it is 8–10 %, exceeding the incidence rate of cardiovascular disease, breast cancer and HIV infection (Chuchalin et al., 2014). According to 2020 data, more than 42.5 thousand asthma individuals were recorded in Krasnoyarsk region, including both adults and children. The prevalence among adolescents has been noted to be steadily increasing.

There are a number of asthma phenotypes and endotypes. Asthma classification is to group patients based on observable combinations of clinical, biological and physiological characteristics into so-called phenotypes. Simply stated, phenotypes are defined as "observable characteristics resulting from a combination of hereditary and environmental influences" (Wenzel, 2012). It is important to emphasize that the phenotype of asthma can change over time, which is caused by environmental factors, allergens, seasonal changes, respiratory infections, iGCS (inhalant glucocorticosteroids) therapy, etc. Asthma is known to be classified according to the Global Initiative for Asthma (GINA) and both severity and level of control as well.

Asthma severity is associated with the intensity of the pathological process and it is possible to correctly determine its degree before treatment, since there is a decrease in symptoms with effective therapy. According to the recommendations of the GINA working group, the asthma severity can be distinguished as: intermittent, persistent-mild, persistent-moderate, and persistent-severe. Level of disease control is the degree to which symptoms and functional limitations are controlled, as well as the minimization of risks of asthma exacerbation, and the prevention of deterioration in lung function with medical treatment. Whatever the disease severity, the goal for patients is to have well-controlled asthma. According to the degree of control, it is classified into controlled and uncontrolled asthma.

An important characteristic of asthma is the multifactorial nature of the disease, with the pathogenesis of development combining both genetic and environmental factors. Extrinsic factors are sure to be numerous and responsible for the activation of asthma manifestation or cause its exacerbation. The internal characteristics of the individual are of greatest importance. Intrinsic (congenital) factors include genetic predisposition, gender and ethnic origin. The important role of heredity in asthma occurrence has been confirmed by family, twin, and genetic epidemiological studies (Thomsen, 2014).

The genetic component of the disease is provided by the combined action of various groups of genes. The same asthma phenotype in different individuals may result from the "breakdown" of various genes; the disease development might result from a mutation of several genes at once in every single individual. In addition, not only the possibility of developing the disease, but also its severity, response to therapy, etc. are determined by hereditary factors.

Airway inflammation underlying asthma is also caused by the so-called cytokine network, which is a self-regulating system; when its functionality is impaired, an excess or insufficient production of various cytokines takes place, which turns out to result in the development of pathological processes. More than 50 cytokines are known to be involved in the immune pathogenesis of this disease, and the role of each of those has not been fully elucidated. Significantly higher levels of cytokines such as GM-CSF, IL-4, IL-5, IL-10, IL-12, IL-13, IL-17A, IL-8, IL-18, TNF-α can be detected in serum.

Gene polymorphism is known to cause differences in the expression and level of protein production. Currently, a huge number of polymorphic regions have been identified in the genes of a number of cytokines and their receptors. Despite the progress made in the study of the immune-pathogenesis of asthma, there has been no agreement of opinion on the pathogenetic role of polymorphic variants for cytokine genes related to asthma development as well as its phenotypes, which is to be further studied. In addition, there have been some contradictions in the study results for different populations worldwide, as the frequency distribution of polymorphic variants of genes, including cytokine ones, has unique features depending on ethno-geographic characteristics (Puzyrev et al.,

2007). Therefore, a comparative analysis of genetic parameters in a single population in order to identify risk factors for asthma phenotype development is to be relevant.

Thus, asthma is the subject of research aimed at studying the disease process, the role of various mediators (including cytokines), treatment approaches, and the role of the genetic determinant as well.

The aim of the study was to identify markers for the development of various asthma phenotypes in Russian children of Krasnoyarsk region.

Materials and methods

Asthma patients (n = 317) and healthy children (control group) (n = 229) matched by sex, age and ethnicity were the object of the study. Criteria for patients to be involved in the study were the following: an established diagnosis of bronchial asthma; age from 8 to 18 years; more than one-year of asthma experience; both parents of the child being Russians. There are some criteria for exclusion from the study such as concomitant decompensated diseases as well as for inclusion in the control group: the absence of allergic pathologies and bronchopulmonary diseases; age from 6 to 18 years; both parents of the child being Russians.

Depending on the severity of the disease, determined in accordance with the recommendations of the GINA working group (Global Initiative for Asthma, updated, 2018 and 2021), the following groups were distinguished: intermittent, persistent-mild, persistent-moderate, and persistent-severe. In the course of the study, we grouped patients according to the following severity levels: intermittent asthma and persistent-mild asthma into the "mild" asthma group (n = 131), persistent-moderate and persistent-severe into the "severe" asthma group (n = 186) due to the small number of patients in some groups. Depending on the level of disease control, based on the results of the asthma control test in children (C-ACT, asthma control test), the following groups were distinguished: controlled asthma (n = 171) – 20 or more points, and uncontrolled asthma – less than 19 points (n = 146).

The work was performed in accordance with the principles stated in the Declaration of Helsinki on research in humans and animals. The studies were approved at a meeting of the local ethical committee of Scientific Research Institute of Medical Problems of the North (The Minutes No. 12 dated December 10, 2013). The examination protocol for patients and healthy children (control group) met ethical standards and was approved by the Biomedical Ethics Committee of Scientific Research Institute of Medical Problems of the North. The right to conduct an examination was legally secured by the informed written consent of the parent.

DNA extraction from blood was carried out using the DIAtom™ DNA Prep100 reagent kit (Isogene, Russia). Genotyping was conducted by the real-time PCR method using specific oligonucleotide primers and fluorescently labeled probes according to the manufacturer's protocol (DNA Synthesis, Russia) and the Rotor-Gene Q 6 plex instrument (QIAGEN, Germany).

Comparison of allele and genotype frequencies between groups was performed using an online calculator https://medstatistic.ru/. The χ^2 test was used to assess the association of a trait-genotype with the disease in groups of sick and almost healthy children. The threshold significance level was taken equal to 0.05. The odds ratio (OR) was used with a 95 % confidence interval (CI) for an assessment of the degree of association of genetic markers with traits.

Results

In order to identify genetic markers of asthma, a comparative analysis of the frequency of single nucleotide polymorphisms (SNPs) between patients and children in the control group was made. Comparative analysis of allele and genotype frequencies between the cohort of asthma patients and controls revealed statistically significant differences in the SNPs distribution in the promoter regions of *IL4* rs2243250 and *IL13* rs1800925 (Table 1).

The prevalence of the *IL4* rs2243250 T allele in the group of asthma patients relative to the control group was shown (28 % versus 23.5 %, p = 0.05), with the frequency of the heterozygous CT genotype of *IL4* rs2243250 being also statistically significantly higher in asthma patients compared to the control group (p = 0.006). The frequencies of the TT genotype and T allele of *IL13* rs1800925 are significantly higher in the group of patients relative to the control group ($p < 0.05$).

A comparison of the genotype and allele frequencies in the group of asthma patients depending on the severity and level of asthma control was made (Tables 2 and 3) to study in detail the association of allelic variants of cytokine genes with the characteristics of asthma development.

As a result of analysis of the *IL4* rs2243250 and *IL13* rs1800925 distribution depending on the severity of asthma, a high frequency of both the CT genotypes of *IL4* rs2243250 and TT genotypes of *IL13* rs1800925 in the group with severe asthma relative to the control group was noted, and for the CT genotype of *IL4* rs2243250, in the group of children with mild asthma ($p < 0.05$). Analysis of the allele frequencies of these polymorphic gene variants revealed significant differences in the rare T allele frequency of *IL4* rs2243250 and *IL13* rs1800925 in children with mild and severe asthma (in the case of rs1800925) compared with healthy children ($p < 0.05$).

When comparing the frequency of *IL17A* rs2275913, *IFNG* rs2069705 and *TNFA* rs1800629 genotypes and alleles, AA homozygotes of rs2275913 (p = 0.01), TT of rs2069705 (p = 0.03) and allelic A variant of *TNFA* rs1800629 were shown to be significantly more common in the group of children with mild asthma relative to the control group.

As a result of the analysis of the *IL4* rs2243250 and *IL13* rs1800925 distribution depending on the level of asthma control, it was demonstrated that the CT of *IL4* rs2243250 and TT of *IL13* rs1800925 genotypes are more common in the uncontrolled asthma group compared to the controls ($p < 0.05$). The CT genotype of *IL4* rs2243250 is also significantly more common in controlled asthma patients than in controls. Allele frequency analysis revealed the differences in the frequency of the rare T allele of *IL13* rs1800925 between groups of children with both controlled and uncontrolled asthma, and healthy children as well ($p < 0.05$). When comparing the frequency of the *IFNG* rs2069705 genotypes, the homozygous TT was shown to be significantly higher in the controlled asthma group compared to the control group ($p < 0.05$).

Table 1. Prevalence of the genotypes and alleles of the SNPs in asthma patients and control group, % (*n*)

Genotype/ allele	Control	Asthma	OR (CI)	*p*	Genotype/ allele	Control	Asthma	OR (CI)	*p*
IL2 rs2069762					*IL13* rs1800925				
TT	38.3 (88)	45.2 (143)	1.33 (0.94–1.88)	0.103	CC	56.3 (129)	42.2 (134)	**0.57 (0.40–0.80)**	**0.002**
TG	46.9 (108)	41.8 (132)	0.81 (0.58–1.14)	0.229	CT	38.0 (87)	46.1 (146)	1.39 (0.99–1.97)	0.061
GG	14.8 (34)	13.0 (41)	0.86 (0.53–1.40)	0.545	TT	5.7 (13)	11.7 (37)	**2.20 (1.14–4.23)**	**0.017**
G	38.3 (176)	33.9 (214)	0.83 (0.64–1.06)	0.135	T	24.7 (113)	34.7 (220)	**1.62 (1.24–2.12)**	**< 0.001**
IL4 rs2243250					*IL17A* rs2275913				
CC	61.3 (141)	50.9 (161)	**0.66 (0.46–0.93)**	**0.017**	GG	41.5 (95)	38.4 (122)	0.88 (0.62–1.24)	0.462
CT	30.4 (70)	42.1 (133)	**1.66 (1.16–2.38)**	**0.006**	GA	45.4 (104)	42.4 (135)	0.89 (0.63–1.25)	0.491
TT	8.3 (19)	7.0 (22)	0.83 (0.44–1.57)	0.570	AA	13.1 (30)	19.2 (61)	1.57 (0.98–2.53)	0.060
T	23.5 (108)	28.0 (177)	**1.41 (1.00–1.98)**	**0.050**	A	35.8 (164)	40.4 (257)	1.22 (0.95–1.56)	0.123
IL5 rs2069812					*IL31* rs7977932				
CC	45.2 (104)	42.4 (134)	0.89 (0.63–1.26)	0.513	CC	63.8 (146)	64.1 (202)	1.02 (0.71–1.45)	0.929
CT	46.1 (106)	48.4 (153)	1.09 (0.78–1.54)	0.591	CG	30.1 (69)	31.8 (100)	1.08 (0.75–1.56)	0.688
TT	8.7 (20)	9.2 (29)	1.06 (0.58–1.93)	0.846	GG	6.1 (14)	4.1 (13)	0.66 (0.30–1.43)	0.293
T	31.7 (146)	33.4 (211)	1.08 (0.83–1.39)	0.567	G	21.2 (97)	20.0 (126)	0.93 (0.69–1.25)	0.635
IL6 rs1800795					*IL33* rs7044343				
GG	31.9 (73)	33.7 (106)	1.08 (0.75–1.56)	0.664	TT	32.8 (75)	29.4 (93)	0.86 (0.59–1.24)	0.408
CG	43.7 (100)	47.6 (150)	1.17 (0.83–1.65)	0.362	CT	53.7 (123)	56.0 (177)	1.10 (0.78–1.54)	0.595
CC	24.4 (56)	18.7 (59)	0.71 (0.47–1.08)	0.107	CC	13.5 (31)	14.6 (46)	1.09 (0.67–1.78)	0.736
G	53.7 (246)	57.5 (362)	1.16 (0.91–1.48)	0.219	T	59.6 (273)	57.4 (363)	0.91 (0.72–1.17)	0.474
IL10 rs1800872					*IFNG* rs2069705				
CC	55.7 (128)	59.8 (189)	1.19 (0.84–1.67)	0.331	TT	28.3 (65)	33.9 (108)	1.32 (0.91–1.88)	0.143
CA	36.9 (85)	35.1 (111)	0.92 (0.65–1.31)	0.660	TC	51.3 (118)	46.5 (147)	0.90 (0.64–1.27)	0.270
AA	7.4 (17)	5.1 (16)	0.67 (0.33–1.35)	0.260	CC	20.4 (47)	19.6 (61)	0.95 (0.62–1.45)	0.744
A	25.9 (119)	22.6 (143)	0.84 (0.63–1.11)	0.216	C	46.1 (212)	42.6 (269)	0.93 (0.61–1.42)	0.293
IL12B rs3212220					*TNFA* rs1800629				
GG	57.2 (131)	65.1 (205)	1.39 (0.98–1.98)	0.063	GG	80.9 (186)	75.2 (231)	0.72 (0.47–1.09)	0.122
GT	37.6 (86)	29.8 (94)	0.71 (0.49–1.01)	0.127	GA	16.9 (39)	21.2 (65)	1.31 (0.85–2.04)	0.222
TT	5.2 (12)	5.1 (16)	0.97 (0.45–2.09)	0.934	AA	2.2 (5)	3.6 (11)	1.67 (0.57–4.88)	0.342
T	24.0 (110)	20.0 (126)	0.79 (0.59–1.06)	0.113	A	10.7 (49)	14.2 (87)	1.38 (0.95–2.01)	0.087

Discussion

Since the inflammatory response regulation for asthma has been carried out using mediators/cytokines, the mechanisms of violation of their functionality have to be studied. The level of cytokine concentration in blood serum is affected by genetic polymorphism of the cytokine network, which turns out to have an effect on the asthma progression type. By 2022, about 1500 genes have been studied for asthma, including cytokines and their receptors (according to Phenopedia). The influence of various genes on the formation of a genetic predisposition to asthma should be noted to be significantly different in various populations, i. e. has some ethnogeographic features. Hence, there are some conflicting data in the studies on the role of genetic factors in the asthma pathogenesis. As a result of the 1000 Genomes project (http://www.1000genomes.org), data on a number of SNPs in genes, including those in promoter, exons, and intron regions have been obtained. However, there are few functional polymorphic variants (affecting the protein functions or structure) with their contribution to the pathology of asthma being ambiguous.

Table 2. Prevalence of the genotypes and alleles of the SNPs in patients with mild and severe asthma and in control group, % (*n*)

Genotype/allele	Control (1)	Mild asthma (2)	Severe asthma (3)	OR (CI)	*p*
			IL4 rs2243250		
CC	61.3 (141)	48.1 (63)	53.0 (98)	1.2 = 0.59 (0.38–0.90)	1.2 = 0.015
CT	30.4 (70)	43.5 (57)	41.1 (76)	1.2 = 1.76 (1.13–2.75) 1.3 = 1.59 (1.06–2.39)	1.2 = 0.013 1.3 = 0.024
T	23.5 (108)	30.2 (79)	26.5 (98)	1.2 = 1.41 (1.00–1.98)	1.2 = 0.050
			IL6 rs1800795		
CC	24.5 (56)	13.6 (18)	22.4 (41)	1.2 = 0.49 (0.27–0.87)	1.2 = 0.015
			IL13 rs1800925		
CC	56.3 (129)	43.6 (58)	41.3 (76)	1.2 = 0.60 (0.39–0.92) 1.3 = 0.55 (0.37–0.81)	1.2 = 0.020 1.3 = 0.003
TT	5.7 (13)	11.3 (15)	12.0 (22)	1.3 = 2.26 (1.10–4.61)	1.3 = 0.023
T	24.7 (113)	33.8 (90)	35.3 (130)	1.2 = 1.56 (1.12–2.17) 1.3 = 1.67 (1.23–2.25)	1.2 = 0.010 1.3 < 0.001
			IL17A rs2275913		
AA	13.1 (30)	23.7 (31)	16.2 (30)	1.2 = 2.06 (1.18–3.59)	1.2 = 0.011
			IFNG rs2069705		
TT	28.3 (65)	38.9 (51)	30.3 (56)	1.2 = 1.62 (1.03–2.55)	1.2 = 0.037
			TNFA rs1800629		
GG	80.9 (186)	71.4 (90)	77.9 (141)	1.2 = 0.59 (0.36–0.98)	1.2 = 0.042
A	10.7 (49)	16.7 (42)	12.4 (45)	1.2 = 1.68 (1.08–2.62)	1.2 = 0.022

Note. Genotypes and alleles are shown, with their frequency difference between the comparison groups $p \le 0.05$.

Table 3. Prevalence of the genotypes and alleles of the SNPs in patients with controlled and uncontrolled asthma and in control group, % (*n*)

Genotype/allele	Control (1)	Controlled asthma (2)	Uncontrolled asthma (3)	OR (CI)	*p*
			IL4 rs2243250		
CC	61.3 (141)	51.2 (87)	50.7 (74)	1.2 = 0.66 (0.44–0.99) 1.3 = 0.65 (0.43–0.99)	1.2 = 0.044 1.3 = 0.043
CT	30.4 (70)	43.5 (74)	40.4 (59)	1.2 = 1.76 (1.16–2.66) 1.3 = 1.55 (1.00-2.39)	1.2 = 0.007 1.3 = 0.048
			IL6 rs1800795		
CC	24.5 (56)	16.5 (28)	21.4 (31)	1.2 = 0.61 (0.37–1.01)	1.2 = 0.054
			IL12B rs3212220		
T	24.0 (110)	22.1 (75)	17.6 (51)	1.3 = 0.67 (0.47–0.98)	1.3 = 0.038
			IL13 rs1800925		
CC	56.3 (129)	43.9 (75)	40.4 (59)	1.2 = 0.61 (0.41–0.90) 1.3 = 0.53 (0.34–0.80)	1.2 = 0.014 1.3 = 0.003
TT	5.7 (13)	10.5 (18)	13.0 (19)	1.3 = 2.49 (1.19–5.20)	1.3 = 0.014
T	24.7 (113)	33.3 (114)	36.3 (106)	1.2 = 1.53 (1.12–2.08) 1.3 = 1.74 (1.26–2.39)	1.2 = 0.008 1.3 < 0.001
			IFNG rs2069705		
TT	28.3 (65)	37.7 (64)	30.1 (44)	1.2 = 1.53 (1.00–2.34)	1.2 = 0.048

Note. SNPs are given, with the frequency difference between the comparison groups $p \le 0.05$.

In our study the SNP allele and genotype frequencies of key cytokines produced by different types of the immune system cells that mediate inflammatory reactions in diseases among Russian children in Krasnoyarsk region were studied. We found significant differences in the frequency of polymorphic distribution of cytokine genes between asthma patients and the control group, allowing us to identify genetic markers that are suggestive risk factors for developing asthma, i. e. the heterozygous CT genotype and the T allele of *IL4* rs2243250, the homozygous TT variant and the T allele of *IL13* rs1800925.

As mentioned above, the disease course prediction, effectiveness of treatment, controlled course, prevention of the severe asthma development, as well as providing personalized therapy and asthma prophylaxis are of the greatest importance.

In order to find genetic markers of different types of asthma, we analyzed the SNPs distribution of cytokine genes in patients with different severity and control of the disease. The CT genotype and the T allele of *IL4* rs2243250 was found to be associated with mild asthma, the CT genotype of *IL4* rs2243250 was also associated with severe asthma; moreover, the TT genotype of *IL13* rs1800925 was associated with severe asthma, and the T allele of *IL13* rs1800925 – with both mild and severe asthma. Genetic markers predisposing to different forms of asthma depending on the control were also identified, i. e. both the CT genotype of *IL4* rs2243250 and the T allele of *IL13* rs1800925 were associated with both controlled and uncontrolled asthma, with the TT genotype of *IL13* rs1800925 being associated with with uncontrolled one.

The *IL4* and *IL13* genes are located in one cluster of chromosome 5q31.1 and encode cytokines that play a key role in the asthma pathogenesis, namely, IL-4 and IL-13 promote airway eosinophilia, mucus hyperproduction, bronchial hyperreactivity, and IgE synthesis (Zhang et al., 2015). The SNP (rs2243250) in the *IL4* promoter is associated with increased expression and production of IL-4, and SNP *IL13* rs1800925 enhances the expression of IL-13 in Th2 cells. Asthma patients with an elevated IgE level were reported to have a homozygous genotype for the rare allele T of *IL4* rs2243250. Our data obtained as a result of analysis of the frequency distribution of genotypes and alleles rs2243250 and rs1800925 in Krasnoyarsk children are consistent with the study results of other scientists.

It was previously determined that the CT genotype of *IL4* rs2243250 predominates in the group of Russian children with atopic asthma, and Arab asthma patients having this genotype were also found to have the highest incidence of eczema compared to the patients with the TT genotype (Hijazi, Haider, 2000; Smirnova et al., 2018). In the asthma children group, an increased incidence rate of the TC and TT genotypes of the *IL4* (C-590T) polymorphism (rs2243250) compared to healthy ones was shown (Prosekova et al., 2020). W. Nie et al., in the meta-analysis including 40 studies, concluded that the CT *vs*. CC was significantly associated with an increased risk of developing asthma. In addition, when analyzed by ethnicity, significant associations were found in Asians and Caucasians, but not in African Americans (Nie et al., 2013). However, some studies obtained different results, for instance, the analysis of genotypes associated with asthma for C-589T of the *IL4* gene did not reveal statistically significant differences between the control group and the group of asthma patients, which might be due to the small number of studied samples (Rudenko et al., 2021). And in a meta-analysis carried out by Chinese scientists, a rare allele was said to be a weak risk factor for asthma development in Caucasians (Liu et al., 2012).

Z. Liu et al. (2014) have shown that the CT and TT genotypes of *IL13* rs1800925 were more common in the group of asthma patients. Scientists from Malaysia have found out that the percentage of the minor T allele in asthma patients was above the frequency of the same allele in the control, being a risk factor for the development of this pathology (Radhakrishnan et al., 2013). However, the study results on a population of children in Costa Rica have demonstrated that the T allele rs1800925 led to the progression of asthma only in children taking corticosteroids and was not associated with the risk of developing the disease (Hunninghake et al., 2007).

As a result of meta-analysis, the mutation rs1800925 was associated with an increased risk of developing asthma only in the Caucasian population, and not associated with a predisposition to asthma in Asians (Omraninava et al., 2020).

There are also controversial data, which are likely to be related to the small number of studied samples, in particular, an analysis of the distribution of alleles and genotypes of *IL13* rs1800925 did not reveal statistically significant differences between the control and the group of asthma patients. However, there was a tendency to increase the proportion of allele C in the group of asthma patients (Kutlina et al., 2018). Nevertheless, polymorphic variants of the *TNFA*, *IL4*, and *IL13* cytokine genes have been shown to contribute to the formation of a genetic predisposition to asthma in the Republic of Bashkortostan (Karunas et al., 2012).

While working, we have also found that the AA genotype of *IL17A* rs2275913, the TT genotype of *IFNG* rs2069705, and the A allele of *TNFA* rs1800629 were associated with mild asthma, and the TT genotype of *IFNG* rs2069705 – with controlled asthma.

The literature available on the association of *IL17A* rs2275913, localized in the promoter region, with the expression level and cytokine activity of IL-17A are very inconsistent. Thus, an association between the SNP and susceptibility to asthma in children has been noted, i. e. the GG genotype patients have mild to moderate asthma and low levels of IL-17A (Maalmi et al., 2014). The A allele of rs2275913 has been reported to increase the activity of the *IL17A* promoter and upregulate its transcription, leading to increased airway inflammation (Espinoza et al., 2011). However, another study failed to find an association between *IL17A* rs2275913 and asthma risk (Wang et al., 2011), while J. Chen et al. have demonstrated that the level of *IL-17A* expression in peripheral blood mononuclear cells was not affected by rs2275913 (Chen et al., 2010). One of the ethnicity-specific analysis showed that the G allele of *IL17A* rs2275913 was a protective factor of the asthma in Asians, with no association being found in Africans (Zhai et al., 2018).

It is known that one of the key Th1-cytokines is IFN-γ, involved in the many features regulation of asthma pathogenesis, including suppression of the of Th2 profile cytokine release, inhibition of the recruitment of effector cells to the site of inflammation, apoptosis induction of T-cells, eosinophils, etc.

Nevertheless, there have currently been a limited number of studies investigating the role of polymorphic sites in the *IFNG* in the pathogenesis of asthma. The G-238A mutation in the *TNFA* gene has been shown to reduce the risk for developing asthma, whereas the SNP G-308A (rs1800629) was associated with the development of asthma and an increase in TNF-α production (Zedan et al., 2008). The A allele of rs1800629 has also been shown to be associated with increased *TNFA* transcription compared to the G allele, with its frequency varying significantly between ethnic groups and being rare in Japanese (less than 3 %) (Wilson et al., 1997).

Tomsk scientists, who have been studying pathogenetics of asthma for many years, found an association of the polymorphic variant of the *TNFA* gene (rs1800629) with the development of asthma, namely, the AA genotype was more often indicated in the group of patients compared to the control (Zhalsanova et al., 2020). According to a series of study results, an analysis depending on ethnodemographic data is necessary. Only in this case, the obtained markers of the diseases can be used as prognostic ones.

Some researchers have distinguished not only genetic markers of the risk of developing a disease or its forms, but also some protective markers. The CC genotypes of *IL4* rs2243250, *IL6* rs1800795, *IL13* rs1800925, as well as the GG genotype of *TNFA* rs1800629 were shown in our study to be protective against the development of mild asthma. It was also determined that the CC genotypes of *IL4* rs2243250, *IL6* rs1800795, *IL13* rs1800925 and the allelic variant T of *IL12B* rs3212220 can be considered to be potentially protective for the development of uncontrolled asthma.

Conclusion

Thus, the obtained data on the prevalence of genetic variants indicate that functional SNPs in cytokine genes are associated with asthma and various disease courses not only in adults, but also in children. However, it is evident that the results do not always agree with each other; this is due to several reasons, namely, the ethnicity of the population, the study sample size, the presence of concomitant diseases, etc. In addition, differences between children and adults can be caused by either the presence or absence, as well as the different duration of exposure of asthma patients to some environmental risk factors, including contact allergens and irritants, air pollution, smoking and occupational exposure.

An important aspect of medical practice is to achieve disease control, which is supposed to involve a personalized approach to treatment for asthma of any severity. It should be taken into account that children with severe asthma are at increased risk for adverse outcomes, including drug-related side effects, life-threatening exacerbations, and poor quality of life. As a result, the study of the distribution of allelic variants of cytokine genes in asthma among patients of different ages, representatives of different populations needs to be continued in order to find risk factors for different types of asthma. The obtained results will update with the data on the polymorphic distribution of cytokine genes in the Russian population and in asthma patients with different disease courses. This is most likely to be ultimately used for practical healthcare authorities to develop measures both in order to prevent severe uncontrolled asthma and to optimize personalized therapy.

References

Chen J., Deng Y., Zhao J., Luo Z., Peng W., Yang J., Ren L., Wang L., Fu Z., Yang X., Liu E. The polymorphism of IL-17 G-152A was associated with childhood asthma and bacterial colonization of the hypopharynx in bronchiolitis. *J. Clin. Immunol*. 2010;30(4):539-545. DOI 10.1007/S10875-010-9391-8.

Chuchalin A.G., Khaltaev N., Antonov N.S., Galkin D.V., Manakov L.G., Antonini P., Murphy M., Solodovnikov A.G., Bousquet J., Pereira M.H.S., Demko I.V. Chronic respiratory diseases and risk factors in 12 regions of the Russian Federation. *Int. J. Chron. Obstruct. Pulmon. Dis.* 2014;9:963-974. DOI 10.2147/COPD.S67283.

Espinoza J.L., Takami A., Nakata K., Onizuka M., Kawase T., Akiyama H., Miyamura K., Morishima Y., Fukuda T., Kodera Y., Nakao S. A genetic variant in the IL-17 promoter is functionally associated with acute graft-versus-host disease after unrelated bone marrow transplantation. *PloS One*. 2011;6(10):e26229. DOI 10.1371/JOURNAL.PONE.0026229.

GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir. Med.* 2017;5(9):691-706. DOI 10.1016/S2213-2600(17)30293-X.

Hancox R.J., Cowan D.C., Aldridge R.E., Cowan J.O., Palmay R., Williamson A., Town G.I., Taylor D.R. Asthma phenotypes: Consistency of classification using induced sputum. *Respirology.* 2012;17(3):461-466. DOI 10.1111/J.1440-1843.2011.02113.X.

Hijazi Z., Haider M.Z. Interleukin-4 gene promoter polymorphism [C590T] and asthma in Kuwaiti Arabs. *Int. Arch. Allergy Immunol.* 2000;122(3):190-194. DOI 10.1159/000024396. DOI 10.1165/AJRCMB.25.3.4483.

Hunninghake G.M., Soto-Quirós M.E., Avila L., Su J., Murphy A., Demeo D.L., Ly N.P., Liang C., Sylvia J.S., Klanderman B.J., Lange C., Raby B.A., Silverman E.K., Celedón J.C. Polymorphisms in IL13, total IgE, eosinophilia, and asthma exacerbations in childhood. *J. Allergy Clin. Immunol.* 2007;120(1):84-90. DOI 10.1016/J.JACI.2007.04.032.

Karunas A.S., Fedorova Yu.Yu., Ramazanova N.N., Galimova E.S., Gimalova G.F., Guryeva L.L., Levasheva S.V., Biktasheva A.R., Mukhtarova L.A., Zagidullin Sh.Z., Etkina E.I., Khusnutdinova E.K. Evaluation of a role of cytokine gene polymorphisms in development of bronchial asthma in the Republic of Bashkortostan. *Pul'monologiya = Pulmonology.* 2012;5:37-40. (in Russian)

Kutlina T.G., Valova Ya.V., Karimov D.D., Kudoyarov E.R., Mukhammadiyeva G.F., Karimov D.O., Kabirova E.F. The role of rs1800925 gene *IL-13* in the formation of broncial asthma in residents of the Republic Bashkortostan. *Vestnik Bashkirskogo Gosudarstvennogo Meditsinskogo Universiteta = Bulletin of the Bashkir State Medical University*. 2018;6:81-84. (in Russian)

Liu S., Li T., Liu J. Interleukin-4 rs2243250 polymorphism is associated with asthma among Caucasians and related to atopic asthma. *Cytokine*. 2012;59(2):364-369. DOI 10.1016/J.CYTO.2012.05.006.

Liu Z., Li P., Wang J., Fan Q., Yan P., Zhang X., Han B. A meta-analysis of IL-13 polymorphisms and pediatric asthma risk. *Med. Sci. Monit.* 2014;20:2617. DOI 10.12659/MSM.891017.

Maalmi H., Beraies A., Charad R., Ammar J., Hamzaoui K., Hamzaoui A. IL-17A and IL-17F genes variants and susceptibility to childhood asthma in Tunisia. *J. Asthma.* 2014;51(4):348-354. DOI 10.3109/02770903.2013.876647.

Nie W., Zhu Z., Pan X., Xiu Q. The interleukin-4 − 589C/T polymorphism and the risk of asthma: A meta-analysis including 7345 cases and 7819 controls. *Gene.* 2013;520(1):22-29. DOI 10.1016/J.GENE.2013.02.027.

Omraninava M., Eslami M.M., Aslani S., Razi B., Imani D., Feyzinia S. Interleukin 13 gene polymorphism and susceptibility to asthma: a meta-regression and meta-analysis. *Eur. Ann. Allergy Clin. Immunol.* 2020. DOI 10.23822/EURANNACI.1764-1489.180.

Prosekova E.V., Dolgopolov M.S., Sabynych V.A. Gene polymorphism, spontaneous and induced production of interleukin 4 and interferon gamma by peripheral blood cells in children with asthma. *Meditsinskoye Obozreniye = Russian Medical Review.* 2020;4(1): 10-14. (in Russian)

Puzyrev V.P., Freidin M.B., Kucher A.N. Genetic Diversity of the Population and Human Diseases. Tomsk, 2007. (in Russian)

Radhakrishnan A.K., Raj V.L., Tan L.K., Liam C.K. Single nucleotide polymorphism in the promoter of the human interleukin-13 gene is associated with asthma in Malaysian adults. *BioMed Res. Int.* 2013. DOI 10.1155/2013/981012.

Rudenko K.A., Tuguz A.R., Tatarkova E.A. C-589T *IL-4* gene polymorphism in the pathogenesis of bronchial asthma. *Vyatskii Meditsinskii Vestnik = Medical Newsletter of Vyatka.* 2021;2(70):42-47. (in Russian)

Smirnova S.V., Smolnikova M.V., Konopleva O.S. Polymorphism of genes (IL4, IL5) as a genetic predisposition to asthma in children. *Eur. Respir. J.* 2018;52(Suppl. 62):PA4462. DOI 10.1183/13993003.congress-2018.PA4462.

Thomsen S.F. Exploring the origins of asthma: Lessons from twin studies. *Eur. Clin. Respir. J*. 2014;1(Suppl. 1):25535. DOI 10.3402/ECRJ.V1.25535.

Wang J., Zhou J., Lin L.H., Li J., Peng X., Li L. Association of single nucleotide polymorphism of IL-17 gene promoter with childhood asthma. *Acad. J. Second Mil. Med. Univ*. 2011;32(5):481-484. DOI 10.3724/SP.J.1008.2011.00481.

Wenzel S. Severe asthma: from characteristics to phenotypes to endotypes. *Clin. Exp. Allergy.* 2012; 42(5):650-658. DOI 10.1111/J.1365-2222.2011.03929.X.

Wilson A.G., Symons J.A., Mcdowell T.L., Mcdevitt H.O., Duff G.W. Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. *Proc. Natl. Acad. Sci. USA.* 1997;94(7):3195-3199. DOI 10.1073/PNAS.94.7.3195.

Zedan M., Settin A., Farag M.K., El-Bayoumi M., El Regal M.E., El Baz R., Osman E. Gene polymorphisms of tumor necrosis factor alpha-308 and interleukin-10-1082 among asthmatic Egyptian children. *Allergy Asthma Proc.* 2008;29(3):268-273. DOI 10.2500/AAP.2008.29.3123.

Zhai C., Li S., Feng W., Shi W., Wang J., Wang Q., Chai L., Zhang Q., Yan X., Li M. Association of interleukin-17a rs2275913 gene polymorphism and asthma risk: a meta-analysis. *Arch. Med. Sci.* 2018; 14(6):1204. DOI 10.5114/AOMS.2018.73345.

Zhalsanova I.Zh., Bragina E.Yu., Babushkina N.P., Tarasenko N.V., Nazarenko M.S., Puzyrev V.P. The role of the *TNF* (rs1800629), *TNFB* (rs2239704) and *TNFRSF1B* (rs652625) genes polymorphic variants in the allergic and infectious disease's development. *Meditsinskaya Genetika = Medical Genetics*. 2020;19(8):90-91. DOI 10.25557/2073-7998.2020.08.90-91. (in Russian)

Zhang S., Li Y., Liu Y. Interleukin-4 -589C/T polymorphism is associated with increased pediatric asthma risk: a meta-analysis. *Inflammation*. 2015;38(3):1207-1212. DOI 10.1007/S10753-014-0086-9.

ORCID ID

M.V. Smolnikova orcid.org/0000-0001-9984-2029
Ed.W. Kasparov orcid.org/0000-0002-5988-1688
M.A. Malinchik orcid.org/0000-0002-6350-8616
K.V. Kopylova orcid.org/0000-0002-5006-0429

Acknowledgements. Authors express their gratitude to the researchers of the Clinical Department of Somatic and Mental Health of Children of Scientific Research Institute of Medical Problems of the North for the help in collecting the material.

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

Received June 10, 2022. Revised August 9, 2022. Accepted October 25, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):401-410
DOI 10.18699/VJGB-23-48

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

Epigenetic regulation of bone remodeling and its role in the pathogenesis of primary osteoporosis

B.I. Yalaev[1, 2] ✉, R.I. Khusainova[1, 2, 3]

[1] Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Center of the Russian Academy of Sciences, Ufa, Russia
[2] Saint Petersburg State University, St. Petersburg, Russia
[3] Ufa University of Science and Technology, Ufa, Russia
✉ yalaev.bulat@yandex.ru

Abstract. Discovery of molecular mechanisms of primary osteoporosis development is fundamental to understand the pathogenesis of musculoskeletal diseases in general and for identifying key links in the genetic and epigenetic regulation of bone remodelling genes. The number of identified molecular genetic markers for osteoporosis is increasing but there is a need to describe their functional interactions. These interactions have been determined to be associated with the control of expression of a number of transcription factors and the differentiation of mesenchymal stem cells through the pathway of osteoblastogenesis or adipogenesis, and monocytic precursors through the pathway of osteoclastogenesis. The results of epigenetic studies have significantly increased the understanding of the role of post-translational modifications of histones, DNA methylation and RNA interference in the osteoporosis pathogenesis and in bone remodelling. However, the knowledge should be systematised and generalised according to the results of research on the role of epigenetic modifiers in the development of osteoporosis, and the influence of each epigenetic mechanism on the individual links of bone remodelling during ontogenesis of humans in general, including the elderly, should be described. Understanding which mechanisms and systems are involved in the development of this nosology is of interest for the development of targeted therapies, as the possibility of using microRNAs to regulate genes is now being considered. Systematisation of these data is important to investigate the differences in epigenetic marker arrays by race and ethnicity. The review article analyses references to relevant reviews and original articles, classifies information on current advances in the study of epigenetic mechanisms in osteoporosis and reviews the results of studies of epigenetic mechanisms on individual links of bone remodelling.
Key words: osteoporosis; methylation; microRNA; acetylation.

For citation: Yalaev B.I., Khusainova R.I. Epigenetic regulation of bone remodeling and its role in the pathogenesis of primary osteoporosis. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4): 401-410. DOI 10.18699/VJGB-23-48

Эпигенетическая регуляция ремоделирования костной ткани и ее роль в патогенезе первичного остеопороза

Б.И. Ялаев[1, 2] ✉, Р.И. Хусаинова[1, 2, 3]

[1] Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
[2] Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
[3] Уфимский университет науки и технологий, Уфа, Россия
✉ yalaev.bulat@yandex.ru

Аннотация. Раскрытие молекулярных механизмов развития первичного остеопороза имеет фундаментальное значение как с точки зрения понимания патогенеза заболеваний опорно-двигательного аппарата в целом, так и для выявления ключевых звеньев генетической и эпигенетической регуляции экспрессии генов ремоделирования костной ткани. Количество обнаруженных молекулярно-генетических маркеров остеопороза продолжает расти, однако существует очевидная необходимость описания их функциональных взаимодействий. Установлено, что такие взаимодействия сопряжены с контролем экспрессии ряда факторов транскрипции и дифференцировки мезенхимальных стволовых клеток по пути остеобластогенеза и адипогенеза, а моноцитарных предшественников – по пути остеокластогенеза. Кроме того, результаты эпигенетических исследований значительно расширили понимание роли посттрансляционных модификаций гистонов, ДНК-метилирования и РНК-интерференции как в молекулярном патогенезе первичного остеопороза, так и в регуляции развития костной ткани. Несмотря на это, знания не систематизированы и нуждаются в обобщении данных исследований роли эпигенетических модификаторов в развитии первичного остеопороза, и, что не менее важно, в описании влияния каждого известного эпигенетического механизма на отдельные молекулярные звенья про-

цесса формирования и резорбции костной ткани в течение онтогенеза человека, в том числе у лиц пожилого возраста. Понимание того, какие молекулярно-генетические механизмы и регуляторные системы вовлечены в развитие данной нозологии, представляет потенциальный интерес для создания таргетной терапии, поскольку уже сейчас рассматривается вопрос о возможности применения микроРНК для узконаправленной регуляции генов. Кроме того, систематизация этих данных важна для изучения разницы массивов эпигенетических маркеров, в зависимости от расовой и этнической принадлежности. В представленной обзорной статье проанализированы соответствующие систематические обзоры и оригинальные статьи, собрана и классифицирована информация о современных достижениях в области изучения эпигенетических механизмов и их аберраций при первичном остеопорозе, а также рассмотрены результаты исследований эпигенетических механизмов на отдельных функциональных звеньях ремоделирования костной ткани.
Ключевые слова: остеопороз; метилирование; микроРНК; ацетилирование.

Introduction

Primary osteoporosis (OP) is an age-associated disease of multifactorial aetiology, which is based on a violation of the balance of bone remodelling, leading to a decrease in the level of bone mineral density (BMD) and a violation of the structure of bone microarchitectonics, which result in the appearance of an incorrect spatial structure of the spongy and cortical bone (Yalaev et al., 2021). Bone mass is reduced to 26 % in individuals predisposed to OP. In roughly 80 % of women, the mineral content in the spine falls below the threshold value at the age of 60–70 years, and at 85 years – in more than 90 % (Sveshnikov, 2013). According to statistics, the frequency of fractures in women is 33 %, in men, 20 % (Marshall et al., 1996; Estrada et al., 2012; Lee et al., 2020; Widl et al., 2020).

Until recently, OP was diagnosed only by secondary signs, such as low height and bone pain (Lorentzon, Cummings, 2015). In 1940, the American endocrinologist F. Albright, describing postmenopausal osteoporosis, suggested that it developed due to estrogen deficiency (Albright et al., 1940). On this basis, clinicians developed the now obsolete concept of two forms of OP, one associated with estrogen deficiency during menopause and the other with calcium deficiency and skeletal ageing, both of which are characteristic of both sexes (Riggs et al., 1982).

Current evidence defines OP as a musculoskeletal disease associated with profound metabolic changes not only in bone, but also in whole-body homeostasis: micro-nutrient and endocrine dysregulation, as well as a complex interaction of genetic, endogenous and environmental factors that contribute to a complex disease phenotype (Foger-Samwald et al., 2020). Risk factors with regards to OP are assumed to be the female sex, being of the Caucasoid or Mongoloid race (Thomas, 2007), early menopause, old age, family history, insufficient insolation, comorbidities with impaired bone matrix micronutrient absorption, smoking, alcohol abuse, sedentary lifestyle, taking hormonal drugs and rheumatoid arthritis. DNA testing has not been introduced into diagnostic practice because significant population differences in the frequency distribution of risk markers prevent the development of universal test systems, despite the existence of significant and validated genetic markers for OP (de Souza, 2010; Bolland et al., 2011).

The first multicentre molecular genetic studies based on the genome-wide association search (GWAS) method confirmed that OP is associated with genes for local and systemic regulation of bone cell function. The lion's share of risk markers, specifically, single-nucleotide polymorphisms, have been identified not in exons but in introns and promoters of regulatory genes, transcription factor, receptor (*ESR2*) and growth factor (*FGF2*) genes (Rivadeneira et al., 2009; Estrada et al., 2012; Wood et al., 2015). Certain markers have been identified in non-coding RNA (ncRNA) sequences, including microRNA (Lei et al., 2011; Yalaev, Khusainova, 2020). Significant levels of associations with fractures have been identified among the genes of systemic and local regulators (e. g., *RANKL* and *OPG*), transmembrane receptors, as well as WNT signalling genes, nuclear transcription factors (*ZNF239*, etc.,) and enzymes that produce or inactivate local bone regulators (Raisz, 2005).

Bioinformatics studies have concluded that the main enrichment pathways by means of the functional affiliation of genes associated with histone modifications and microRNA patterns are significantly associated with the regulation of RUNX2, FGF2 and SOX9 transcription factors. Moreover, they are also associated with the regulation of mesenchymal stem cell (MSC) differentiation (Letarouilly et al., 2019). Thus, particular epigenetic factors in the pathogenesis of osteoporosis have been identified. However, at present, it is more interesting to consider the molecular pathogenesis of OP in terms of individual functional links in the regulation of bone remodelling, in particular how RANK/RANKL/OPG, WNT, RUNX2 transcription factor and others are epigenetically regulated.

The role of DNA methylation and microRNA in the regulation of the RANK/RANKL/OPG system in primary osteoporosis

Controlled differentiation of osteoblast and osteoclast precursors (mononuclear phagocytes) is necessary to maintain the balance of bone remodelling (Soltanoff et al., 2009). Osteoclast activity is primarily regulated by the RANK/RANKL/OPG (receptor-activator nuclear factor kβ/RANK ligand/osteoprotegerin) system. RANKL is produced by osteoblasts and its binding to RANK on the osteoclast

surface activates the expression of osteoclastogenic genes (Tobeiha et al., 2020).

The role of gene polymorphisms of this system in increased risk of fracture and the formation of low BMD has previously been shown (Yalaev, Khusainova, 2020). In the *RANKL* gene sequence, two CpG sites were detected: one in the upstream sequence with 18 CpG sites, localized at a distance of 14,415 pairs of nucleotides (bp) from the transcription start site of the TSS I major isoform and one in the downstream sequence with 59 CpG sites, which starts at 260 and ends at 615 pairs of nucleotides of the TSS I isoform. In the *OPG* gene sequence, one island of 56 sites spanning from –402 to +850 nucleotide pairs from the transcription site of the TSS isoform was observed (Delgado-Calle et al., 2012). In (Wang et al., 2018), a group of patients from Guangzhou Medical University Hospital with osteoporotic fractures had significantly higher *RANKL* gene mRNA levels from femoral bone cells compared to controls, with downregulated methylation status (Wang et al., 2018).

Dysregulation of the expression of these genes is known to be a key link in the development of steroid-induced osteonecrosis of the femoral head and was noted to be associated with increased DNA methylation levels of *OPG* and *RANK* genes and with decreased levels in the *RANKL* gene (Sun et al., 2021). Previously, J. Delgado-Calle et al. (2012) performed a comparative analysis of the expression levels and methylation profile of the *RANKL* and *OPG* genes in bone tissue samples from patients with osteoporotic femoral neck fractures and osteoarthritis (OA) of the hip joint, which obtained remarkable results. *RANKL* expression levels were significantly higher in patients with fractures ($p = 0.012$), while no significant changes in *OPG* gene expression levels were observed. The ratio of RANKL ligand to OPG was higher in bone samples with OP (7.66 ± 0.23 versus 0.92 ± 0.21, $p = 0.002$). Differential methylation analysis revealed that the upstream promoter region of the *RANKL* gene was strongly methylated in all samples, while individual CpG junctions of the gene were equally hypomethylated in the comparison groups (Delgado-Calle et al., 2012).

There is evidence of the contribution of miRNA to the regulation of the RANK/RANKL/OPG system. In 2014, C. Chen et al. published results measuring miR-503 microRNA levels in peripheral blood cells, the overexpression of which in $CD14^+$ mononuclear cells inhibited RANKL-induced osteoclastogenesis. In $CD14^+$ cells from postmenopausal women with osteoporosis, the baseline level of miR-503 was lower than in normal controls and had no change after induction by RANKL factor, suggesting the direct role of miR-503 in the regulation of *RANK* expression (Chen et al., 2014). miR-142-3p and miR-21-5p are potential biomarkers of OP as they have a high affinity with the *OPG* gene mRNA and are involved in the regulation of several signalling pathways involved in bone formation (Ge et al., 2007; Hu et al., 2020). Thus, the epigenetic regulation of the RANK/RANKL/OPG system is dynamic and dependent on the DNA methylation status and a number of microRNAs.

The role of epigenetic mechanisms in the regulation of bone cell differentiation through alteration of RUNX2 activity

Transcription factor 2 (RUNX2) plays a crucial role in osteoblast differentiation. Gene expression is predominantly high in the early stages of bone cell development when MSCs are differentiating into osteoblast precursors, but naturally decreases at the osteocyte maturation stage (Stein et al., 2004).

The gene contains several functional regions: activation domain, runt domain, PST domain, etc. (Gomathi et al., 2020). In terms of epigenetic regulators of RUNX2, microRNAs are well studied. In particular, miRNA-194 modulates MSC differentiation (Gomathi et al., 2020) and accelerates osteoblast differentiation by regulating RUNX2 nuclear translocation by way of STAT1 signal transducer (Li J. et al., 2015). miR-133a-5p inhibits *RUNX2* gene expression at the transcription and translation level by binding to the 3′-untranslated mRNA site (Zhang et al., 2018).

During the early stages of osteoblast maturation, miR-125b affects *RUNX2* expression by affinity binding to the 3′-untranslated region of the gene, indirectly participating in the formation of a complex with Cbfβ that inhibits differentiation of these cells. Using microarray technology, P. Garmilla-Ezquerra et al. (2015) discovered a significant reduction in *miR-187-3p* gene expression levels and induction of miR-518f in bone with low BMD. In their research, Y. Zhang et al. (2017) observed the involvement of miR-221 in the formation of low mineral density through regulation of RUNX2 activity based on bioinformatics analysis (Garmilla-Ezquerra et al., 2015; Zhang et al., 2017). Several microRNAs affecting the activity of this factor are presented in Table 1.

Phosphorylation of RUNX2 mobilises chromatin regulatory factors and accelerates MSC maturation. RUNX2 is phosphorylated at specific serine residues 301 and 319, inducing osteocyte maturation through MAPK-dependent signalling (Ge et al., 2009; Li Y. et al., 2017) and BMP2-sensitive transcription (Afzal et al., 2005). Phosphorylation of S104 leads to prevention of RUNX protein degradation (Huang et al., 2001; Wee et al., 2002). The ERK-MAPK signalling pathway plays a crucial role in the regulation of *RUNX2* gene expression and bone formation (Ge et al., 2007). MKK6 is a protein kinase with dual specificity that participates in the MAP kinase signal transduction pathway and promotes RUNX2 phosphorylation (Ge et al., 2012). In addition, parathormone, which is one of the main regulators of blood calcium levels, activates the phosphorylation of the RUNX2 factor by means of the PKA signalling pathway. This process is associated with the activation of the

Table 1. Factors involved in the regulation of RUNX2 expression

MicroRNA	Function	Reference
miR-590-5p	Protecting RUNX2 from degradation	Ge et al., 2012
miR-873-3p	Increases RUNX2 activity	Selvamurugan et al., 2009
miR-98	Reduces the expression of the *RUNX2* gene	Selvamurugan et al., 2000
miR-21-5p	Increases RUNX2 activity	Selvamurugan et al., 2017

promoter of the MMP13 gene, which plays an essential role in bone resorption (Selvamurugan et al., 2000, 2009).

Post-translational modification of histones, in particular histone methylation, plays an important role in bone formation. The so-called JUMONJI protein is considered to be a transcription factor and is encoded by the *JARID2* gene. The domain containing 3 JMJD3 is a histone demethylase specifically catalysing the removal of trimethylation of histone H3K27me3. JMJD3 was found to inhibit the differentiation of RUNX2 osteoblasts. Conversely, inhibition of JMJD3 activity decreases *RUNX2* promoter activity while increasing H3K27me3 activity in promoter regions (Yang et al., 2013).

Histone acetylation affects the state of chromatin compaction by neutralising the positive charge of histone tails and reducing the electrostatic interactions of histone tails with deoxyribonucleic acid. Studies have revealed that osteoporosis induced by glucocorticoid therapy leads to decreased acetylation of H3K9/K14 and H4K12 in the regulatory regions of the *RUNX2* and *OSX* genes and increases the hyperacetylation of H3K9/K14 and H4K12 in the *PPARγ2* regulatory region in bone marrow-derived MSCs from osteoporosis. The transcriptional activity of the *RUNX2* factor gene is enhanced by P300 acetyltransferase and nicotinamide phosphoribosyltransferase (NAMPT), which, in turn, promote osteogenic differentiation of MSCs by H3K14 acylation and MC3T3-E1 cells through H3K9 acylation, respectively (Xu et al., 2021).

Thus, an extremely significant number of regulators of *RUNX2* gene expression have been identified, which is a promising therapeutic gene model in culture studies, as blocking or enhancing the activity of this gene and monitoring its expression level during the induction of osteogenic lines allows the identification of key switches of mesenchymal stem cell differentiation.

The role of epigenetic regulation of the WNT-signalling pathway in the regulation of bone remodelling and the pathogenesis of osteoporosis

The WNT signalling pathway is one of the most important systems regulating embryonic development and cell differentiation. This pathway represents one of the central links in the control of bone development and remodelling. Among the various genes involved in this system, methylation of the *SOST* (encoding sclerostin) gene promoter has been comprehensively studied in osteoblast cultures. Sclerostin produced by osteocytes inhibits WNT signalling and reduces the rate of bone formation. DNA methylation of the gene is necessary for the transition of osteoblasts to osteocytes (Delgado-Calle et al., 2012). In women with primary OP, *SOST* methylation is elevated in iliac bone cells whilst sclerostin levels are decreased and the WNT pathway is enhanced (Reppe et al., 2015).

Several studies have shown that histone deacetylation under the regulation of the *WNT6*, *WNT10B*, *WNT10A* and *WNT1* genes inhibits WNT signalling and increases the risk of primary osteoporosis (Jing et al., 2018). Thus, elevated levels of HDAC5 deacetylase reduce *SOST* gene expression in osteocytes, contributing to bone mass loss. Deficiency of this deacetylase is associated with the acetylation of histone H3 lysine 27 as well as the interaction of MEF2C with the *SOST* gene enhancer, therefore suggesting the significant role of sclerostin in the regulation of osteocyte maturation (Wein et al., 2016). High levels of the zeste 2 homologue enhancer methyltransferase (EZH2) have been demonstrated to suppress osteogenic differentiation of MSCs, while low ones decrease the levels of the H3K27me3 tag near the transcription start site of osteogenesis genes, including *WNT10B* (Dudakovic et al., 2016). EZH2 increases H3K27me3 levels at the *WNT1*, *WNT6* and *WNT10A* promoters and inhibits WNT signalling (Jing et al., 2016).

Many regulatory microRNAs are associated with WNT signalling. miR-433-3p inhibits *DKK1* (Dickkopf-1) gene expression, enhancing osteoblast differentiation. Dickkopf-1 acts as an antagonist of the WNT signalling pathway and enhances bone resorption (Tang et al., 2017). miR-139-5p induces WNT signalling via the inhibition of NOTCH1 (Feng et al., 2020). R.E. Makitie et al. (2018) screened a specially designed panel of 192 microRNAs in patients with a genetically determined WNT signalling disorder with a heterozygous missense mutation c. 652 T>G (p. C218G) in exon 4 of the *WNT1* gene. It was determined that miR-22-3p, miR-34a-5p and miR-31-5p levels were lower in mutation carriers compared to controls (Makitie et al., 2018).

Another common inhibitor of osteogenic differentiation is miR-31, the level of which drops in MSCs differentiating into osteoblasts. This was confirmed by S. Weilner et al. (2016), who observed increased levels of this microRNA

in plasma in elderly patients with OP (Weilner et al., 2016; Amjadi-Moheb, Akhavan-Niaki, 2019). miR-31 is released from extracellular vesicles of endothelial cells and inhibits osteogenesis in stromal stem cells by binding to the Freisled 3 protein. Furthermore, decreased levels of miR-199a-5p result in glucocorticoid-mediated inhibition of osteogenesis (Shi et al., 2015). More recently, L. Duan et al. (2018) identified that high levels of miR-16-2* may contribute to the development of primary OP: knockdown of this microRNA may promote RUNX2 activation. This microRNA has an affinity for the *WNT5A* gene mRNA (Duan et al., 2018). miR-148a-3p has been revealed to enhance both osteoclastogenesis (Cheng et al., 2013) and adipogenesis in osteogenic progenitor cells (Gao et al., 2011). Plasma levels of this microRNA are significantly higher in patients with OP compared to controls without OP (Bedene et al., 2016). miR-30e is another important microRNA in the pathogenesis of OP, playing a role in regulating adipocyte and osteoblast differentiation through the inhibition of LRP6 (Wang et al., 2013). Thus, the WNT signalling pathway is regulated by a complex epigenetic system, notably microRNAs.

The role of non-coding RNAs in bone remodelling

MicroRNAs are considered to be the most studied epigenetic factors in osteoporosis (Yalaev, Khusainova, 2020). They are conventionally divided into two classes: those that promote bone formation or bone resorption. In particular, several microRNAs that slow down the progression of OP have been identified. For example, miR-33-5p is a mechanosensitive microRNA that positively regulates osteoblastogenesis by way of inhibition of HMGA2 high mobility group proteins (Wang et al., 2016). miR-96 enhances osteogenic differentiation by inhibiting phosphorylation of epidermal growth factor receptor EGFR and the expression of major osteoblast factors RUNX2 and OSTERIX (Yang et al., 2014). miR-216a enhances bone formation by regulating the c-Cbl-mediated PI3K/AKT pathway (Li H. et al., 2015). Table 2 shows the microRNAs separated by direction of action in bone remodelling (Yalaev, Khusainova, 2018).

miR-124 is a positive regulator of adipogenic and neurogenic differentiation, while being a negative regulator of myogenic and osteogenic differentiation. It directly targets the *DLX3*, *DLX5* and *DLX2* homeobox genes (Qadir et al., 2015). In one pharmacogenetic study, patients with OP, after three months of treatment with the parathormone analogue Teriparatide, had reduced expression levels of miR-33 and one year later, miR-133a. Simultaneously, there was a general tendency: the increase in the level of BMD increased with a decrease in the level of expression of these microRNAs. Post-transcriptional regulation of *DKK-1* changes due to a decrease in miR-33 microRNA levels and the action of parathyroid hormone, which leads to a decrease in the negative impact of DKK-1 on an alternative regulatory mechanism that improves optimal control of WNT signaling.(Anastasilakis et al., 2018).

Of interest are the results of studies on the effects of long non-coding RNAs on the regulation of Sirtuins, nicotinamide adenine dinucleotide-dependent deacetylases. These molecules have a broad spectrum of action and are associated with longevity and resistance to age-related diseases. It was established that *SIRT1* gene expression is inversely related to HIF1A-AS1 ncRNA expression, while HOXA-AS3 ncRNA interacts with EZH2 and is required for RUNX2 trimethylation of lysine-27 H3 (H3K27me3) factor. Hence, HOXA-AS3 is important for bone formation in general (Yang et al., 2020).

The long ncRNA HOTAIR reduces protein expression and inhibits WNT signalling. DKK1 reduces the protein levels of C-myc, β-catenin, HOTAIR and RUNX2, which theoretically counteracts the regulatory effect of HOTAIR (Zhang et al., 2019). If the expression level of p21 ncRNA is low, WNT signalling becomes more active due to increased E2 secretion, which ultimately increases the rate of bone formation (Yang et al., 2019). Also, reduced levels of H19 ncRNA reduce the level of *DKK4* gene expression (Li B. et al., 2017). AK045490 ncRNA levels are significantly elevated and inhibit bone formation by inhibiting nuclear translocation of β-catenin and suppressing *TCF1*, *LEF1* and *RUNX2* expression (Li et al., 2019). Similarly, AK016739 ncRNA inhibits osteogenic differentiation as it can reduce the expression and activity of osteoblastogenesis transcription factors (Yin et al., 2019).

Inhibition of UCA1 ncRNA promotes bone formation via activation of the BMP-2/(Smad1/5/8) pathway in osteoblasts (Zhang et al., 2019). As a result, microRNAs and long ncRNAs remain among the most studied epigenetic factors involved in the pathogenesis of primary OP but require further systematisation.

Table 2. MicroRNAs stimulating bone formation or resorption

Osteoblast effectors	Stimulants of osteoclastogenesis	Osteoclast inhibitors	WNT antagonist inhibitors
miR-21	miR-214	miR-126-p	miR-29a
miR-216a	miR-183	miR-34a	miR-218
miR-96	miR-9718	miR-7b	miR-355-5p

Epigenetic regulation of adipogenesis and osteoblastogenesis

The mechanisms of the relationship between bone and adipose cell formation are complex and remain an area of active research. Works performed on the cell cultures of osteoblasts and MSCs convincingly show an inverse relationship between differentiation of bone marrow MSCs into adipocytes or osteoblasts. An imbalance between adipogenesis and osteogenesis has been proposed as a mechanism for the development of OP, but obesity itself is not always a predictor of an increased risk of osteoporosis.

Post-translational modifications of histones play a key role in this system. Among them, histone methylation is crucial, in particular in chromatin reorganisation. In particular, lysine methylation in H4K20, H3K27 and H3K9 is associated with decreased levels of transcription, whereas methylation of H3K79, H3K36 and H3K4, with active gene transcription (Huang et al., 2015).

However, concerning osteogenic inducers, the homeobox gene *HOXA10*, which contributes to osteogenic clone determination by enriching the trimethylation of the 4th lysine residue in histone, activating RUNX2, alkaline phosphatase and osteocalcin, thus stimulating bone cell maturation, is important (Hassan et al., 2007). It is known that the combination of methylation and demethylation can function as an epigenetic switch of osteogenesis to adipogenesis based on EZH2 activity, catalysing the trimethylation of histone H3 on lysine 27 key regulatory genes (such as *RUNX2*). Simultaneously, removal of this tag by lysindemethylase 6A inhibits adipogenesis and induces osteoblastogenesis. Proteins that can be targeted by EZH2 and are involved in MSC switching are HDAC9c and HDAC (Chen et al., 2011). A direct correlation was realised between increased levels of EZH2 and decreased levels of *HDAC9c* gene expression (Chen et al., 2016). The methyltransferase activity of EZH2 is reduced by phosphorylation and is associated with osteogenic induction (Wei et al., 2011).

It is acknowledged that during osteocyte aging there is an accumulation of adipose tissue in the bone marrow and, at the same time, the number of mesenchymal stem cells increases in the intercellular phase. From this perspective, it is interesting that overexpression of miR-1292 accelerates the senescence of human adipose-derived stem cells and inhibits bone formation via the Wnt/β-catenin signalling pathway, while miR-10b inhibits adipose stem cell differentiation via the TGF-β pathway (Xu et al., 2020).

Several bone-associated cells, including multipotent bone mesenchymal stem cells, osteoblasts that form bone tissue and osteoclasts that break it down, are in a symbiotic relationship throughout life. A growing body of evidence suggests that epigenetic cell modifications induced by aging contribute to impaired bone remodelling and lead to osteoporosis. A variety of epigenetic mechanisms are involved, including DNA/RNA modifications, histone modifications, microRNAs (miRNAs) and long non-coding RNAs (dnRNAs), and chromatin remodelling (Yu et al., 2022). Thus, epigenetic mechanisms can switch the direction of mesenchymal stem cell differentiation between osteoblastogenesis and adipogenesis.

Results of full-genome studies of methylome

Molecular genetic determinants of endophenotypes of osteoporosis, such as fracture risk and BMD levels, can be converged through the whole epigenome. In the research they conducted, J.A. Morris et al. (2017), leveraging the technological capabilities of Infinium HumanMethylation450, performed a whole-genome methylome analysis, measuring site-specific DNA methylation in 5515 individuals of European descent. Following a meta-analysis of their results, they were able to identify the CpG site cg23196985, which was significantly associated with low MPCT adjusted for multiple comparisons without regard to gender ($p_{BH} = 1.30 \times 10^{-2}$) and in females ($p_{BH} = 3.41 \times 10^{-5}$).

The CpG cg23196985 site is localized to the 5′-translated region of the hepatic carboxylase 1 gene *CES1*, which is expressed in the liver and peripheral blood (Morris et al., 2017). J.G. Zhang et al. (2015), performing transcriptome analysis based on Affymetrix GeneChip Human Exon 1.0 ST Array and microRNA analysis on Capitalbio Cor. microarrays, as well as methylome sequencing in patients with low MPCT hip and controls, identified the most enriched functional molecular pathways associated with OP or MPCT variability: a network of 12 interacting genes and 11 microRNAs. Among the genes are *AKT1*, *STAT5A*, *PIK3R5*, etc., while among the microRNAs, miR-141 and miR-675 – their levels correlate with the expression of these genes and global DNA methylation status (Zhang et al., 2015).

D. Cheishvili et al. (2018) performed a full-epigenomic analysis in women without OP and in women with early postmenopausal OP. Genes in which CpG sites with significant levels of differential methylation were identified were *ZNF267*, *ABLIM2*, *RHOJ*, *CDKL5*, *PDCD1*, *ABRA* and *HOJ* (Cheishvili et al., 2018). At the human femur level, DNA methylation profile studies using pyrosequencing and qRT-PCR-based gene expression studies proved that DNA methylation status was inversely correlated with the expression of the *iNOS* and *COL9A1* genes, but not catabolic genes including *MMP13* and *IL1B*. Significant demethylation of the osteocalcin gene promoter was also shown between the embryonic and adult stages of development, demonstrating the importance of DNA methylation at the tissue level (Curtis et al., 2022). It is anticipated that the results of these studies will confirm the whole-epigenome approach as being sufficiently robust to allow large-scale studies of women at risk of developing OP.

Conclusion

Despite the great advances and the wide range of work performed in the study of the epigenetics of primary osteo-

porosis, understanding remains fragmentary. The research describes key epigenetic regulators of bone remodelling, but it is difficult to build a coherent picture of the pathogenesis of osteoporosis from these data, common to all key pathogenetic processes in bone tissue.

It is recognised that the key epigenetic changes in osteoporosis are converging on the regulation of MSCs differentiation and that the multi-step system of activity regulation of the transcription factor RUNX2, sclerostin, DKK1 protein, RANKL-RANK-OPG system and factors involved in the regulation of WNT signalling is of great importance. A separate issue is the systematisation and creation of a unified genetic network of a vast number of regulatory microRNAs that affect key signalling pathways and the transcription factors associated with osteoporosis. However, these microRNAs influence the ultimate risk of osteoporosis indirectly and it remains to be understood how they can be systematised for relevance and functional involvement in pathogenesis due to their dynamism and different levels depending on tissue specificity.

Approaches necessary to implement early diagnosis or targeted therapies for osteoporosis still need to be developed. The task is complicated by experimental data from methylome studies, in which genes other than critical regulatory factors, such as RUNX2, sclerostin or DKK1, were observed to be key and most important. The data presented in this review show that epigenetic modifications can have a strong influence on the determination, differentiation and activity of mesenchymal stem cells and, therefore, may contribute to the pathophysiology of age-related bone mass loss. The results update further basic research on osteoporosis and, with the available data, broaden the horizons for a more insightful and detailed approach to new epigenetic studies of this disease and the prospects for new and effective personalised therapies.

References

Afzal F., Pratap J., Ito K., Ito Y., Stein J.L., Wijnen A.J., Stein G.S., Lian J.B., Javed A. Smad function and intranuclear targeting share a Runx2 motif required for osteogenic lineage induction and BMP2 responsive transcription. *J. Cell. Physiol*. 2005;204(1):63-72. DOI 10.1002/jcp.20258.

Albright F., Bloomberg E., Smith P.H. Postmenopausal osteoporosis. *Trans. Assoc. Am. Phys*. 1940;55:298-305.

Amjadi-Moheb F., Akhavan-Niaki H. Wnt signaling pathway in osteoporosis: Epigenetic regulation, interaction with other signaling pathways, and therapeutic promises. *J. Cell. Physiol*. 2019;234(9): 14641-14650. DOI 10.1002/jcp.28207.

Anastasilakis A.D., Makras P., Pikilidou M., Tournis S., Makris K., Bisbinas I., Tsave O., Yovos J.G., Yavropoulou M.P. Changes of circulating MicroRNAs in response to treatment with teriparatide or denosumab in postmenopausal osteoporosis. *J. Clin. Endocrinol. Metab*. 2018;103(3):1206-1213. DOI 10.1210/jc.2017-02406.

Bedene A., Bedrac S.M., Jese L., Marc J., Vrtacnik P., Prezelj J., Kocjan T., Kranjc T., Ostanek B. MiR-148a the epigenetic regulator of bone homeostasis is increased in plasma of osteoporotic postmenopausal women. *Wien. Klin. Wochenschr*. 2016;128(7):519-526. DOI 10.1007/s00508-016-1141-3.

Bolland M.J., Siu A.T., Mason B.H., Horne A.M., Ames R.W., Grey A.B., Gamble G.D., Reid I.R. Evaluation of the FRAX and Garvan fracture risk calculators in older women. *J. Bone Miner. Res*. 2011;26(2):420-427. DOI 10.1002/JBMR.215.

Cheishvili D., Parashar S., Mahmood N., Arakelian A., Kremer R., Goltzman D., Szyf M., Rabbani S.A. Identification of an epigenetic signature of osteoporosis in blood DNA of postmenopausal women. *J. Bone Miner. Res*. 2018;33(11):1980-1989. DOI 10.1002/jbmr.3527.

Chen C., Cheng P., Xie H., Zhou H.D., Wu X.P., Liao E.Y., Luo X.H. MiR-503 regulates osteoclastogenesis via targeting RANK. *J. Bone Miner. Res*. 2014;29(2):338-347. DOI 10.1002/JBMR.2032.

Chen Y.H., Chung C.C., Liu Y.C., Yeh S.P., Hsu J.L., Hung M.Ch., Su H.L., Li L.Y. Enhancer of zeste homolog 2 and histone deacetylase 9c regulate age-dependent mesenchymal stem cell differentiation into osteoblasts and adipocytes. *Stem Cells*. 2016;34(8):2183-2193. DOI 10.1002/stem.2400.

Chen Y.H., Yeh F.L., Yeh S.P., Ma H.T., Hung S.C., Hung M.C., Li L.Y. Myocyte enhancer factor-2 interacting transcriptional repressor (MITR) is a switch that promotes osteogenesis and inhibits adipogenesis of mesenchymal stem cells by inactivating peroxisome proliferator-activated receptor γ-2. *J. Biol. Chem*. 2011;286(12):10671-10680. DOI 10.1074/jbc.M110.199612.

Cheng P., Chen C., He H.B., Hu R., Zhou H.D., Xie H., Zhu W., Dai R.C., Wu X.P., Liao E.Y., Luo X.H. miR-148a regulates osteoclastogenesis by targeting V-maf musculoaponeurotic fibrosarcoma oncogene homolog B. *J. Bone Miner. Res*. 2013;28(5):1180-1190. DOI 10.1002/jbmr.1845.

Curtis E.M., Fuggle N.R., Cooper P., Harvey N.C. Epigenetic regulation of bone mass. *Best Pract. Res. Clin. Endocrinol. Metab*. 2022; 36(2):101612. DOI 10.1016/j.beem.2021.101612.

Delgado-Calle J., Sanudo C., Fernandez A.F., Garcia-Renedo R., Fraga M.F., Riancho J.A. Role of DNA methylation in the regulation of the RANKL-OPG system in human bone. *Epigenetics*. 2012; 7(1):83-91. DOI 10.4161/epi.7.1.18753.

Duan L., Zhao H., Xiong Y., Tang X., Yang Y., Hu Z., Li C., Chen S., Yu X. miR-16-2* interferes with WNT5A to regulate osteogenesis of mesenchymal stem cells. *Cell. Physiol. Biochem*. 2018;51(3):1087-1102. DOI 10.1159/000495489.

Dudakovic A., Camilleri E.T., Riester S.M., Paradise C.R., Gluscevic M., O'Toole T.M., Thaler R., Evans J.M., Yan H., Subramaniam M., Hawse J.R., Stein G.S., Montecino M., McGee-Lawrence M.E., Westendorf J.J., Wijnen A.J. Enhancer of zeste homolog 2 inhibition stimulates bone formation and mitigates bone loss caused by ovariectomy in skeletally mature mice. *J. Biol. Chem*. 2016; 291(47):24594-24606. DOI 10.1074/jbc.M116.740571.

Dudakovic A., Camilleri E.T., Xu F., Riester S.M., McGee-Lawrence M.E., Bradley E.W., Paradise C.R., Lewallen E.A., Thaler R., Deyle D.R., Larson A.N., Lewallen D.G., Dietz A.B., Stein G.S., Montecino M.A., Westendorf J.J., Wijnen A.J. Epigenetic control of skeletal development by the histone methyltransferase Ezh2. *J. Biol. Chem*. 2015;290(46):27604-27617. DOI 10.1074/jbc.M115.672345.

Estrada K., Styrkarsdottir U., Evangelou E., … Ioannidis J.P.A., Kiel D.P., Rivadeneira F. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet*. 2012;44(5):491-501. DOI 10.1038/ng.2249.

Feng Y., Wan P., Yin L., Lou X. The inhibition of microRNA-139-5p promoted osteoporosis of bone marrow-derived mesenchymal stem cells by targeting Wnt/beta-catenin signaling pathway by Notch1. *J. Microbiol. Biotechnol*. 2020;30(3):448-458. DOI 10.4014/jmb.1908.08036.

Foger-Samwald U., Dovjak P., Azizi-Semrad U., Kerschan-Schindl K., Pietschmann P. Osteoporosis: pathophysiology and therapeutic options. *EXCLI J*. 2020;19:1017-1037. DOI 10.17179/excli2020-2591.

Gao J., Yang T., Han J., Yan K., Qiu X., Zhou Y., Fan Q., Ma B. MicroRNA expression during osteogenic differentiation of human multipotent mesenchymal stromal cells from bone marrow. *J. Cell. Biochem*. 2011;112(7):1844-1856. DOI 10.1002/jcb.23106.

Garmilla-Ezquerra P., Sanudo C., Delgado-Calle J., Perez-Nunez M.I., Sumillera M., Riancho J.A. Analysis of the bone MicroRNome in osteoporotic fractures. *Calcif. Tissue Int*. 2015;96(1):30-37. DOI 10.1007/s00223-014-9935-7.

Ge C., Xiao G., Jiang D., Franceschi R.T. Critical role of the extracellular signal-regulated kinase-MAPK pathway in osteoblast differentiation and skeletal development. *J. Cell Biol*. 2007;176(5):709-718. DOI 10.1083/JCB.200610046.

Ge C., Xiao G., Jiang D., Yang Q., Hatch N.E., Roca H., Franceschi R.T. Identification and functional characterization of ERK/MAPK phosphorylation sites in the Runx2 transcription factor. *J. Biol. Chem*. 2009;284(47):32533-32543. DOI 10.1074/jbc.M109.040980.

Ge C., Yang Q., Zhao G., Yu H., Kirkwood K.L., Franceschi R.T. Interactions between extracellular signal-regulated kinase 1/2 and P38 Map kinase pathways in the control of RUNX2 phosphorylation and transcriptional activity. *J. Bone Miner. Res*. 2012;27(3):538-551. DOI 10.1002/JBMR.561.

Gomathi K., Akshaya N., Srinaath N., Moorthi A., Selvamurugan N. Regulation of Runx2 by post-translational modifications in osteoblast differentiation. *Life Sci*. 2020;245:117389. DOI 10.1016/J.LFS.2020.117389.

Hassan M.Q., Tare R., Lee S.H., Mandeville M., Weiner B., Montecino M., Wijnen A.J., Stein J.L., Stein G.S., Lian J.B. HOXA10 controls osteoblastogenesis by directly activating bone regulatory and phenotypic genes. *Mol. Cell. Biol*. 2007;27(9):3337-3352. DOI 10.1128/mcb.01544-06.

Hu H., He X., Zhang Y., Wu R., Chen J., Lin Y., Shen B. MicroRNA alterations for diagnosis, prognosis, and treatment of osteoporosis: a comprehensive review and computational functional survey. *Front. Genet*. 2020;11:181. DOI 10.3389/FGENE.2020.00181/BIBTEX.

Huang B., Li G., Jiang X.H. Fate determination in mesenchymal stem cells: A perspective from histone-modifying enzymes. *Stem Cell Res. Ther*. 2015;6(1):35. DOI 10.1186/s13287-015-0018-0.

Huang G., Shigesada K., Ito K., Wee H.J., Yokomizo T., Ito Y. Dimerization with PEBP2β protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation. *EMBO J*. 2001;20(4):723-733. DOI 10.1093/emboj/20.4.723.

Jing H., Liao L., An Y., Su X., Liu S., Shuai Y., Zhang X., Jin Y. Suppression of EZH2 prevents the shift of osteoporotic MSC fate to adipocyte and enhances bone formation during osteoporosis. *Mol. Ther*. 2016;24(2):217-229. DOI 10.1038/mt.2015.152.

Jing H., Su X., Gao B., Shuai Y., Chen J., Deng Z., Liao L., Jin Y. Epigenetic inhibition of Wnt pathway suppresses osteogenic differentiation of BMSCs during osteoporosis. *Cell Death Dis*. 2018;9(2):176. DOI 10.1038/s41419-017-0231-0.

Lei S.F., Papasian C.J., Deng H.W. Polymorphisms in predicted miRNA binding sites and osteoporosis. *J. Bone Miner. Res*. 2011;26(1):72-78. DOI 10.1002/jbmr.186.

Letarouilly J.G., Broux O., Clabaut A. New insights into the epigenetics of osteoporosis. *Genomics*. 2019;111(4):793-798. DOI 10.1016/J.YGENO.2018.05.001.

Li B., Liu J., Zhao J., Ma J.X., Jia H.B., Zhang Y., Xing G.S., Ma X.L. LncRNA-H19 modulates Wnt/β-catenin signaling by targeting Dkk4 in hindlimb unloaded rat. *Orthop. Surg*. 2017;9(3):319-327. DOI 10.1111/os.12321.

Li D., Tian Y., Yin C., Huai Y., Zhao Y., Su P., Wang X., Pei J., Zhang K., Yang C., Dang K., Jiang S., Zhiping M., Li M., Hao Q., Zhang G., Qian A. Silencing of lncRNA AK045490 promotes osteoblast differentiation and bone formation via β-catenin/TCF1/Runx2 signaling axis. *Int. J. Mol. Sci*. 2019;20(24):6229. DOI 10.3390/ijms20246229.

Li H., Li T., Fan J., Li T., Fan L., Wang S., Weng X., Han Q., Zhao R.C. mIR-216a rescues dexamethasone suppression of osteogenesis, promotes osteoblast differentiation and enhances bone formation, by regulating c-Cbl-mediated PI3K/AKT pathway. *Cell Death Differ*. 2015;22(12):1935-1945. DOI 10.1038/cdd.2015.99.

Li H., Xiao Z., Quarles L.D., Li W. Osteoporosis: mechanism, molecular target and current status on drug development. *Curr. Med. Chem*. 2020;28(8):1489-1507. DOI 10.2174/0929867327666200330142432.

Li J., He X., Wei W., Zhou X. MicroRNA-194 promotes osteoblast differentiation via downregulating STAT1. *Biochem. Biophys. Res. Commun*. 2015;460(2):482-488. DOI 10.1016/J.BBRC.2015.03.059.

Li Y., Ge C., Franceschi R.T. MAP kinase-dependent RUNX2 phosphorylation is necessary for epigenetic modification of chromatin during osteoblast differentiation. *J. Cell. Physiol*. 2017;232(9):2427-2435. DOI 10.1002/jcp.25517.

Lorentzon M., Cummings S.R. Osteoporosis: the evolution of a diagnosis. *J. Intern. Med*. 2015;277(6):650-661. DOI 10.1111/joim.12369.

Makitie R.E., Hackl M., Niinimaki R., Kakko S., Grillari J., Makitie O. Altered microRNA profile in osteoporosis caused by impaired WNT signaling. *J. Clin. Endocrinol. Metab*. 2018;103(5):1985-1996. DOI 10.1210/jc.2017-02585.

Marshall D., Johnell O., Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *Br. Med. J*. 1996;312(7041):1254-1259. DOI 10.1136/bmj.312.7041.1254.

Morris J.A., Tsai P.C., Joehanes R., Zheng J., Trajanoska K., Soerensen M., Forgetta V., Castillo-Fernandez J.E., Frost M., Spector T.D., Christensen K., Christiansen L., Rivadeneira F., Tobias J.H., Evans D.M., Kiel D.P., Hsu Y.H., Richards J.B., Bell J.T. Epigenome-wide association of DNA methylation in whole blood with bone mineral density. *J. Bone Miner. Res*. 2017;32(8):1644. DOI 10.1002/JBMR.3148.

Qadir A.S., Um S., Lee H., Baek K., Seo B.M., Lee G., Kim G.S., Woo K.M., Ryoo H.M., Baek J.H. miR-124 negatively regulates osteogenic differentiation and *in vivo* bone formation of mesenchymal stem cells. *J. Cell. Biochem*. 2015;116(5):730-742. DOI 10.1002/JCB.25026.

Raisz L.G. Pathogenesis of osteoporosis: concepts, conflicts, and prospects. *J. Clin. Invest*. 2005;115(12):3318-3325. DOI 10.1172/JCI27071.

Reppe S., Noer A., Grimholt R.M., Halldorsson B.V., Carolina M.G., Gautvik V.T., Olstad O.K., Berg J.P., Datta H., Estrada K., Hofman A., Uitterlinden A.G., Rivadeneira F., Lyle R., Collas P., Gautvik K.M. Methylation of bone SOST, its mRNA, and serum sclerostin levels correlate strongly with fracture risk in postmenopausal women. *J. Bone Miner. Res*. 2015;30(2):249-256. DOI 10.1002/JBMR.2342.

Riggs B.L., Wahner H.W., Seeman E., Offord K.P., Dunn W.L., Mazess R.B., Johnson K.A., Melton L.J. Changes in bone mineral density of the proximal femur and spine with aging. Differences between the postmenopausal and senile osteoporosis syndromes. *J. Clin. Invest.* 1982;704(4):716-723. DOI 10.1172/JCI110667.

Rivadeneira F., Styrkarsdottir U., Estrada K., … Stefansson K., Ioannidis J.P.A., Uitterlinden A.G. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat. Genet*. 2009;41(11):1199-1206. DOI 10.1038/ng.446.

Selvamurugan N., He Z., Rifkin D., Dabovic B., Partridge N. Pulsed electromagnetic field regulates microRNA 21 expression to activate TGF-β signaling in human bone marrow stromal cells to enhance osteoblast differentiation. *Stem Cells Int*. 2017;2017:2450327. DOI 10.1155/2017/2450327.

Selvamurugan N., Pulumati M.R., Tyson D.R., Partridge N.C. Parathyroid hormone regulation of the rat collagenase-3 promoter by protein kinase A-dependent transactivation of core binding factor α1. *J. Biol. Chem*. 2000;275(7):5037-5042. DOI 10.1074/JBC.275.7.5037.

Selvamurugan N., Shimizu E., Lee M., Liu T., Li H., Partridge N.C. Identification and characterization of Runx2 phosphorylation sites involved in matrix metalloproteinase-13 promoter activation. *FEBS Lett*. 2009;583(7):1141-1146. DOI 10.1016/J.FEBSLET.2009.02.040.

Shi C., Huang P., Kang H., Hu B., Qi J., Jiang M., Zhou H., Guo L., Deng L. Glucocorticoid inhibits cell proliferation in differentiating osteoblasts by microRNA-199a targeting of WNT signaling. *J. Mol. Endocrinol*. 2015;54(3):325-337. DOI 10.1530/JME-14-0314.

Soltanoff C.S., Yang S., Chen W., Li Y.P. Signaling networks that control the lineage commitment and differentiation of bone cells. *Crit. Rev. Eukaryot*. *Gene Expr.* 2009;19(1):1-46. DOI 10.1615/CritRevEukarGeneExpr.v19.i1.10.

Souza M.P.G. Osteoporosis diagnosis and treatment. *Rev. Bras. Ortop.* 2010;45(3):220-229. DOI 10.1016/S2255-4971(15)30361-X.

Stein G.S., Lian J.B., Wijnen A.J., Stein J.L., Montecino M., Javed A., Zaidi S.K., Young D.W., Choi J.Y., Pockwinse S.M. Runx2 control of organization, assembly and activity of the regulatory machinery for skeletal gene expression. *Oncogene*. 2004;23(24):4315-4329. DOI 10.1038/sj.onc.1207676.

Sun M., Cao Y., Yang X., An F., Wu H., Wang J. DNA methylation in the *OPG/RANK/RANKL* pathway is associated with steroid-induced osteonecrosis of the femoral head. *BMC Musculoskelet. Disord*. 2021;22(1):599. DOI 10.1186/s12891-021-04472-6.

Sveshnikov A.A. Mineral Density of Skeletal Bones, Muscle Mass, and Issues of Fracture Prevention. Moscow: Akademiya Yestestvoznaniya Publ., 2013. (in Russian)

Tang X., Lin J., Wang G., Lu J. MicroRNA-433-3p promotes osteoblast differentiation through targeting DKK1 expression. *PLoS One*. 2017;12(6):0179860. DOI 10.1371/JOURNAL.PONE.0179860.

Thomas P.A. Racial and ethnic differences in osteoporosis. *J. Am. Acad. Orthop. Surg*. 2007;15(1):26-30. DOI 10.5435/00124635-200700001-00008.

Tobeiha M., Moghadasian M.H., Amin N., Jafarnejad S. RANKL/RANK/OPG pathway: a mechanism involved in exercise-induced bone remodeling. *Biomed. Res. Int*. 2020;2020:6910312. DOI 10.1155/2020/6910312.

Wang H., Sun Z., Wang Y., Hu Z., Zhou H., Zhang L., Hong B., Zhang S., Cao X. miR-33-5p, a novel mechano-sensitive microRNA promotes osteoblast differentiation by targeting Hmga2. *Sci. Rep*. 2016;6:23170. DOI 10.1038/SREP23170.

Wang J., Guan X., Guo F., Zhou J., Chang A., Sun B., Cai Y., Ma Z., Dai C., Li X., Wang B. miR-30e reciprocally regulates the differentiation of adipocytes and osteoblasts by directly targeting low-density lipoprotein receptor-related protein 6. *Cell Death Dis*. 2013;4(10):845. DOI 10.1038/cddis.2013.356.

Wang P., Cao Y., Zhan D., Wang D., Wang B., Liu Y., Li G., He W., Wang H., Xu L. Influence of DNA methylation on the expression of OPG/RANKL in primary osteoporosis. *Int. J. Med. Sci*. 2018;15(13):1480-1485. DOI 10.7150/ijms.27333.

Wee H.J., Huang G., Shigesada K., Ito Y. Serine phosphorylation of RUNX2 with novel potential functions as negative regulatory mechanisms. *EMBO Rep*. 2002;3(10):967-974. DOI 10.1093/embo-reports/kvf193.

Wei Y., Chen Y.H., Li L.Y., Lang J., Yeh S.P., Shi B., Yang C.C., Yang J.Y., Lin C.Y., Lai C.C., Hung M.C. CDK1-dependent phosphorylation of EZH2 suppresses methylation of H3K27 and promotes osteogenic differentiation of human mesenchymal stem cells. *Nat. Cell Biol*. 2011;13(1):87-94. DOI 10.1038/ncb2139.

Weilner S., Schraml E., Wieser M., Messner P., Schneider K., Wassermann K., Micutkova L., Fortschegger K., Maier A., Westendorp R., Resch H., Wolbank S., Redl H., Jansen-Durr P., Pietschmann P., Grillari-Voglauer R., Grillari J. Secreted microvesicular miR-31 inhibits osteogenic differentiation of mesenchymal stem cells. *Aging Cell*. 2016;15(4):744-754. DOI 10.1111/acel.12484.

Wein M.N., Liang Y., Goransson O., Sundberg T.B., Wang J., Williams E.A., O'Meara M.J., Govea N., Beqo B., Nishimori S., Nagano K., Brooks D.J., Martins J.S., Corbin B., Anselmo A., Sadreyev R., Wu J.Y., Sakamoto K., Foretz M., Xavier R.J., Baron R., Bouxsein M., Gardella T.J., Divieti-Pajevic P., Gray N.S., Kronenberg H.M. SIKs control osteocyte responses to parathyroid hormone. *Nat. Commun*. 2016;7:13176. DOI 10.1038/ncomms13176.

Wein M., Spatz J., Nishimori S., Doench J., Root D., Babij P., Nagano K., Baron R., Brooks D., Bouxsein M., Pajevic P.D., Kronenberg H.M. HDAC5 controls MEF2C-driven sclerostin expression in osteocytes. *J. Bone Miner. Res*. 2015;30(3):400-411. DOI 10.1002/jbmr.2381.

Wiedl A., Forch S., Fenwick A., Mayr E. Fractures' associated mortality risk in orthogeriatric inpatients: a prospective 2-year survey. *Eur. Geriatr. Med*. 2020;12(1):61-68. DOI 10.1007/s41999-020-00392-1.

Wood C.L., Stenson C., Embleton N. The developmental origins of osteoporosis. *Curr. Genomics*. 2015;16(6):411. DOI 10.2174/1389202916666150817202217.

Xu F., Li W., Yang X., Na L., Chen L., Liu G. The roles of epigenetics regulation in bone metabolism and osteoporosis. *Front Cell Dev. Biol*. 2020;8:619301. DOI 10.3389/fcell.2020.619301.

Xu Y., Ma J., Xu G., Ma D. Recent advances in the epigenetics of bone metabolism. *J. Bone Miner. Metab*. 2021;39(6):914-924. DOI 10.1007/s00774-021-01249-8.

Yalaev B.I., Khusainova R.I. Epigenetics of osteoporosis. *Meditsinskaya Genetika = Medical Genetics*. 2018;17(6):3-10. DOI 10.25557/2073-7998.2018.06.3-10. (in Russian)

Yalaev B.I., Khusainova R.I. Study of the rs2910164 polymorphic variant of *miR-146a* microRNA gene in patients with primary osteoporosis. *Geny i Kletki = Genes and Cells*. 2020;15(4):40-45. DOI 10.23868/202012007. (in Russian)

Yalaev B.I., Tyurin A.V., Mirgalieva R.I., Khusnutdinova E.K., Khusainova R.I. Investigating the role of osteoprotegerin gene polymorphic variants in osteoporosis. *Russ. Open Med. J*. 2021;10(1):e0101. DOI 10.15275/RUSOMJ.2021.0101.

Yang D., Okamura H., Nakashima Y., Haneji T. Histone demethylase Jmjd3 regulates osteoblast differentiation via transcription factors *Runx2* and osterix. *J. Biol. Chem*. 2013;288(47):33530-33541. DOI 10.1074/jbc.M113.497040.

Yang K., Tian N., Liu H., Tao X.Z., Wang M.X., Huang W. LncRNAp21 promotes osteogenic differentiation of mesenchymal stem cells in the rat model of osteoporosis by the Wnt/β-catenin signaling pathway. *Eur. Rev. Med. Pharmacol. Sci*. 2019;23(10):4303-4309. DOI 10.26355/EURREV_201905_17935.

Yang M., Pan Y., Zhou Y. miR-96 promotes osteogenic differentiation by suppressing HBEGF-EGFR signaling in osteoblastic cells. *FEBS Lett*. 2014;588(24):4761-4768. DOI 10.1016/J.FEBSLET.2014.11.008.

Yang Y., Yujiao W., Fang W., Linhui Y., Ziqi G., Zhichen W., Zirui W., Shengwang W. The roles of miRNA, lncRNA and circRNA in the development of osteoporosis. *Biol. Res*. 2020;53(1):40. DOI 10.1186/s40659-020-00309-z.

Yin C., Tian Y., Yu Y., Wang H., Wu Z., Huang Z., Zhang Y., Li D., Yang C., Wang X., Li Y., Qian A. A novel long noncoding RNA AK016739 inhibits osteoblast differentiation and bone formation. *J. Cell. Physiol*. 2019;2347(7):11524-11536. DOI 10.1002/jcp.27815.

Yu W., Wang H.L., Zhang J., Yin C. The effects of epigenetic modifications on bone remodeling in age-related osteoporosis. *Connect. Tissue Res.* 2022;1-12. DOI 10.1080/03008207.2022.2120392.

Zhang J.G., Tan L.J., Xu C., He H., Tian Q., Zhou Y., Qiu C., Chen X.D., Deng H.W. Integrative analysis of transcriptomic and epigenomic data to reveal regulation patterns for BMD variation. *PLoS One*. 2015;10(9):0138524. DOI 10.1371/journal.pone.0138524.

Zhang R.F., Liu J.W., Yu S.P., Sun D., Wang X.H., Fu J.S., Xie Z. LncRNA UCA1 affects osteoblast proliferation and differentiation by regulating BMP-2 expression. *Eur. Rev. Med. Pharmacol. Sci*. 2019;23(16):6774-6782. DOI 10.26355/eurrev_201908_18715.

Zhang W., Wu Y., Shiozaki Y., Sugimoto Y., Takigawa T., Tanaka M., Matsukawa A., Ozaki T. miRNA-133a-5p inhibits the expression of osteoblast differentiation-associated markers by targeting the 3′ UTR of *RUNX2*. *DNA Cell Biol*. 2018;37(3):199-209. DOI 10.1089/dna.2017.3936.

Zhang Y., Gao Y., Cai L., Li F., Lou Y., Xu N., Kang Y., Yang H. MicroRNA-221 is involved in the regulation of osteoporosis through regulates RUNX2 protein expression and osteoblast differentiation. *Am. J. Transl. Res*. 2017;9(1):126-135.

ORCID ID

B.I. Yalaev orcid.org/0000-0003-4337-1736
R.I. Khusainova orcid.org/0000-0002-8643-850X

Acknowledgements. This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation (No. 122041400169-2), with partial support from the megagrant of the Government of the Russian Federation (agreement No. 075-15-2021-595), and the project of St. Petersburg University (No. 94034528).

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

Received October 21, 2022. Revised December 2, 2022. Accepted December 2, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):411-420
DOI 10.18699/VJGB-23-49

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

Taxonomic and ecophysiological characteristics of actinobacteria in soils of the dry steppe zone of the Selenga Highlands (Western Transbaikalia)

E.P. Nikitina[1, 2] ✉, L.B. Buyantueva[2], E.Yu. Abidueva[3], C.H. Sun[4]

[1] Baikal Institute of Nature Management of the Siberian Branch of the Russian Academy of Sciences, Ulan-Ude, Russia
[2] Banzarov Buryat State University, Ulan-Ude, Russia
[3] Institute of General and Experimental Biology of the Siberian Branch of the Russian Academy of Sciences, Ulan-Ude, Russia
[4] Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences, Beijing, China
✉ lenauude@mail.ru

Abstract. Arid habitats have recently attracted increasing attention in terms of biodiversity research and the discovery of new bacterial species. These habitats are among the target ecosystems suitable for isolating new strains of actinobacteria that are likely to produce new metabolites. This paper presents the results on the isolation of actinobacteria from soils of the dry steppe zone of the Selenga Highlands, the characterization of their taxonomic diversity, as well as ecological and trophic properties. The bacterial counts on ISP 4 medium ranged from 6.6×10^5 to 7.1×10^6 CFU/g. The highest bacterial counts were observed in the subsurface and middle horizons of the studied soils. 28 strains of Gram-positive bacteria represented by thin-branched mycelium, coccoid and bacilliform forms were isolated. According to the results of 16S rRNA gene analysis, the isolated strains were representatives of *Streptomyces*, *Arthrobacter*, *Glycomyces*, *Kocuria*, *Microbacterium*, *Micromonospora*, *Nocardioides*, *Pseudarthrobacter*, and *Rhodococcus* (Actinomycetota). One isolate that showed low 16S rRNA gene sequence similarity with previously isolated and validly described species was a new species of the genus *Glycomyces*. It was shown that all tested strains are mesophilic, prefer neutral or slightly alkaline conditions, have growth limits in the temperature range of 5–45 °C and pH 6–9. The optimal NaCl concentration for growth of most strains was 0–1 %. The strains under study were capable of utilizing a wide range of mono- and disaccharides and polyatomic alcohols as a carbon source. The isolated strains were capable of using both organic (proteins and amino acids) and inorganic (ammonium salts and nitrates) compounds as nitrogen sources. The examinations of extracellular enzymes showed that all isolates were capable of producing catalase and amylase; 78.6 % of the total number of isolates produced protease and lipase; 53.6 %, cellulase; and 28.6 %, urease. The data obtained expand current knowledge about the diversity of microbial communities in soils of the Selenga Highlands and also confirm the potential of searching for new actinobacteria species in these soils.
Key words: chestnut soils; the Selenga Highlands; Actinomycetota; 16S rRNA gene; ecological and trophic properties of bacteria.

For citation: Nikitina E.P., Buyantueva L.B., Abidueva E.Yu., Sun C.H. Taxonomic and ecophysiological characteristics of actinobacteria in soils of the dry steppe zone of the Selenga Highlands (Western Transbaikalia). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):411-420. DOI 10.18699/VJGB-23-49

Таксономическая и экофизиологическая характеристика актинобактерий почв сухостепной зоны Селенгинского среднегорья (Западное Забайкалье)

Е.П. Никитина[1, 2] ✉, Л.Б. Буянтуева[2], Е.Ю. Абидуева[3], Ч.-Х. Сун[4]

[1] Байкальский институт природопользования Сибирского отделения Российской академии наук, Улан-Удэ, Россия
[2] Бурятский государственный университет им. Доржи Банзарова, Улан-Удэ, Россия
[3] Институт общей и экспериментальной биологии Сибирского отделения Российской академии наук, Улан-Удэ, Россия
[4] Институт медицинской биотехнологии Китайской академии медицинских наук, Пекин, Китай
✉ lenauude@mail.ru

Аннотация. Засушливые местообитания привлекают все больше внимания с точки зрения исследования биоразнообразия и обнаружения новых видов бактерий. Они являются одними из целевых экосистем для выделения новых штаммов актинобактерий, которые с большой вероятностью могут продуцировать новые метаболиты. В настоящей работе представлены результаты по выделению актинобактерий из почв сухостепной зоны Селенгинского среднегорья, их таксономическому разнообразию и эколого-трофическим свойствам. Численность бактерий на крахмало-аммиачной среде колебалась от 6.6×10^5 до 7.1×10^6 КОЕ/г. Максимальные

значения численности были отмечены в подповерхностных и срединных горизонтах исследуемых почв. Получено 28 штаммов грамположительных бактерий, представленных тонким разветвленным мицелием, кокковидными и палочковидными формами. По результатам анализа последовательностей гена 16S рРНК выделенные культуры были отнесены к родам *Streptomyces*, *Arthrobacter*, *Glycomyces*, *Kocuria*, *Microbacterium*, *Micromonospora*, *Nocardioides*, *Pseudarthrobacter* и *Rhodococcus* филума Actinomycetota. Один изолят, показавший низкое сходство последовательности гена 16S рРНК с ранее выделенными и достоверно описанными видами, представлял собой новый вид рода *Glycomyces*. Все исследуемые штаммы мезофильны, предпочитают нейтральные или слабощелочные условия, имеют границы роста в диапазоне температур от 5 до 45 °С и значений pH от 6 до 9. Оптимальная концентрация NaCl для роста культур составляла от 0 до 1 %. Исследуемые штаммы были способны утилизировать в качестве источника углерода достаточно широкий спектр моно- и дисахаридов, многоатомных спиртов. В качестве источника азота выделенные культуры использовали как органические (белки и аминокислоты), так и неорганические (соли аммония и нитраты) соединения. Исследование наличия внеклеточных ферментов показало, что все культуры могли продуцировать каталазу и амилазу, 78.6 % от общего количества изолятов продуцировали протеазу и липазу, 53.6 % – целлюлазу, 28.6 % – уреазу. Полученные данные расширяют знания о разнообразии микробных сообществ почв Селенгинского среднегорья и подтверждают, что данные почвы представляют интерес с точки зрения поиска новых видов актинобактерий.
Ключевые слова: каштановые почвы; Селенгинское среднегорье; актинобактерии; 16S рРНК; эколого-трофические свойства бактерий.

Introduction

Actinobacteria (Actinomycetota) are a morphologically diverse group of predominantly gram-positive bacteria widely distributed in various terrestrial and aquatic ecosystems (Ventura et al., 2007; Hazarika, Thakur, 2020). They play an important role in the organic matter cycle, especially in soils, contributing to the decomposition of natural polymers such as starch, chitin, pectin, cellulose, hemicellulose and lignocellulose (McCarthy, Williams, 1992; Manucharova et al., 2004; Wang et al., 2016; Leo et al., 2018; Bao et al., 2021). Some actinobacteria species participate in the synthesis and mineralization of humus substances (Tepper, 1981; Wu et al., 2011).

In early studies, actinobacteria (in particular, actinomycetes) were considered to be unstable to the influence of extreme environmental factors, and therefore unable to occupy certain ecological niches (Kalakuckiy, Agre, 1977; Lechevalier, 1981). Later, as bacterial cultivation methods were improved, and modern molecular research methods became available, actinobacteria resistant to one or another environmental factor became known (Zenova et al., 2009, 2016; Yaradoddi et al., 2021). A large diversity of Actinomycetota is now reported in arid habitats (Kurapova et al., 2012; Zenova et al., 2014; Mohammadipanah, Wink, 2016; Xie, Pathom-aree, 2021). Certain actinobacteria can grow in soils in dry climates due to such properties as xerophilicity, resistance to ultraviolet light, mycelial structure, and spore-forming ability (Zenova, Zvyagintsev, 2002; Zenova et al., 2014; Yaradoddi et al., 2021).

In this context, the soils of the dry steppe zone of the Selenga Highlands are interesting to study because they are formed in pronounced continental and arid climates (Nogina, 1964; Batuev et al., 2000). This area is characterized by high levels of solar radiation, low and irregular precipitation, and sharp average daily and monthly fluctuations in air temperature (Chimitdorzhieva G.D., Chimitdorzhieva E.O., 2021). These conditions may have contributed to a great taxonomic diversity of actinobacteria, which may include new species and possess unique physiological mechanisms of adaptation. However, culturable soil actinobacteria of the dry steppe soils of Transbaikalia remain relatively unstudied: there are only few publications devoted mainly to the study of actinomycetes abundance (Nimaeva, 1992; Zvyagintsev et al., 1999; Buyantueva et al., 2014).

Given the above, this work aimed to isolate culturable actinobacteria from the soils of the dry steppe zone of the Selenga Highlands, to determine their taxonomic diversity, as well as ecological and trophic characteristics.

Materials and methods

Subjects of the study. Actinobacteria strains were isolated from soil samples from the dry steppe zone of the Selenga Highlands. Chestnut soils are typical of these areas, which are characterized by a sharply continental climate, long seasonal permafrost, limited rainfall (180–250 mm/year), and dry steppe vegetation. These areas are characterized by a significant accumulated temperature during the growth period (1700–1800 °C) and the length of the frost-free period (106–116 days). Winter precipitation is not more than 10 % of the annual amount, resulting in poor snow cover and prolonged spring droughts. In July and August, up to 60–70 % of the total annual precipitation falls (Nogina, 1964; Ecological Atlas…, 2015).

Four soil profiles were established. Soil profiles 1T (coordinates 51°08′58.62″ N, 107°24′25.38″ E; 613 m a.s.l.) and 3T (51°11′15.24″ N, 107°34′46.08″ E; 698 m a.s.l.) were established in the western part of the Tugnui Basin at the base of the southern slope of the Tsagan-Daban Range; soil profiles 4I (51°34′50.94″ N, 107°03′56.34″ E; 637 m a.s.l.) and 5I (51°37′1.98″ N, 107°07′42.06″ E; 686 m a.s.l.) – at the foot of the southwestern slope of the Khamar-Daban Range near the Ivolginskaya Depression.

Sampling. Soil samples were collected in the summer of 2017 according to genetic horizons. For physicochemical analyses, the soil samples were dried to an air-dry state. For microbiological studies, samples were collected in sterile containers, from three walls of each soil section in three replicates. After transportation in a cooling box, the samples were delivered to the laboratory within 12 hours. Soil samples were stored at 4 °C for no more than a week before the study. Immediately before inoculation, soil samples were dried to air-dry in a sterile laminar flow cabinet.

Physicochemical properties of soil. Soil pH was measured in water according to GOST 26423-85 (Soils. Methods for Determination of Specific Electric Conductivity, pH and Solid Residue of Water Extract); total organic carbon (TOC) content was measured according to Tyurin (Manual on Agrochemistry, 2001); total nitrogen (TN) – according to GOST 26107-84 (Soils. Methods for Determination of Total Nitrogen). Soil particle size distribution was determined using a laser diffraction particle size analyzer Analysette 22 MicroTec Plus (FRITSCH, Germany).

Pure strains isolation. Actinobacteria were isolated using a dilution plate technique. Samples were inoculated on inorganic salts-starch agar ISP 4 (Shirling, Gottlieb, 1966). The media was supplemented with nystatin (50 μg/mL) to limit the growth of fungi. The plates were incubated at 30 °C for 2–3 weeks. The actinobacteria isolates were preliminarily characterized by their morphological characteristics using a Zeiss AxioStar Plus light microscope (Carl Zeiss, Germany) with a magnification of 1000×. Further routine isolation and culturing of the dominant morphotypes were performed on yeast extract-malt extract agar ISP 2 (Shirling, Gottlieb, 1966).

DNA extraction, amplification, and sequencing of the 16S rRNA gene. DNA was isolated according to the method described by Zhou et al. (2010). Two universal primers 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′) were used for the amplification of 16S rRNA gene fragments (DeLong, 1992). Amplification was performed in a reaction mixture of 50 μL containing 25 μL of 2× EasyTaq PCR SuperMix (TransGen Biotech, China), 1.5 μL of each primer (10 mM, Sangon Biotech, PRC), 2 μL DNA, and 20 μL deionized water in a Veriti™ 96-Well Thermal Cycler DNA Amplifier (Applied Biosystems, USA). The temperature-time profile of PCR was as follows: the first cycle was 95 °C × 5 min; the subsequent 35 cycles were 94 °C × 1 min, 55 °C × 1 min, and 72 °C × 2 min; the final cycle was 72 °C × 10 min. PCR products were purified and sequenced at Sangon Biotech Company (Beijing, China) using an ABI PRISM 3730xl Genetic Analyzer (Thermo Fisher Scientific).

Taxonomic and phylogenetic analysis. The 16S rRNA gene sequence similarity was analyzed using EzTaxon-e (Yoon et al., 2017) and BLAST (Camacho et al., 2009) services. Then, the sequences of closely related species were retrieved from the GenBank database using the EzBioCloud server. Multiple sequence alignment was performed using ClustalW software. The phylogenetic trees were constructed using the neighbor-joining method using MEGA 7.0 (Kumar et al., 2016), and the branching relationships were confirmed by maximum likelihood and maximum parsimony methods. The statistical reliability of the phylogenetic reconstructions was assessed using bootstrap analysis by constructing 1000 alternative trees. The obtained nucleotide sequences were deposited in GenBank with accession numbers assigned to the strains (MN314472–MN314496, MW410748, MW410749).

Ecophysiological characteristics of the isolated bacteria. The isolates were cultured at different temperatures (5 to 55 °C, at 5 °C intervals) and NaCl concentrations (0, 1, 3, 5, 6, 7, 8, 9, and 10 %) to identify optimal parameters and growth limits. The pH range (5.0–10.0, at 0.5 intervals) was set at 30 °C by adding a buffer solution system (Xu et al., 2005). The ability to consume various carbon sources was tested according to Shirling and Gottlieb (1966). The ability to grow on a medium with organic acids was tested according to Gordon et al. (1974). The presence of extracellular enzymes (amylase, catalase, lipase, protease, cellulase, and urease), as well as the ability to release hydrogen sulfide and ammonia, were tested according to Williams et al. (1983). The tests were performed in three replicates; the corresponding sterile nutrient media were used as control samples.

Results

Physicochemical properties of soils and the total bacterial count

The contents of total organic carbon and total nitrogen were maximum in the upper horizons of the studied soils (Table 1). Down the profile, a rather sharp decrease in the content of both indicators was observed. The pH in the upper horizons was almost neutral (6.85–7.54), while a gradual alkalization was observed down the profile. The granulometric analysis of the soils showed a predominance of light loam, except for light-humic soil, which had a sandy loam composition.

The bacterial counts on ISP 4 medium reached several million colony-forming units per 1 g of soil (CFU/g soil) and ranged from 6.6×10^5 to 7.1×10^6 CFU/g. The highest bacterial counts were observed in the subsurface and middle horizons of the studied soils. Mycelial actinobacteria (actinomycetes) colonies were noticeably predominant on nutrient media, accounting for 40–80 % of the total number of bacterial colonies on the plates.

Pure strains and cell morphology of actinobacteria

34 strains of aerobic bacteria were isolated from the soil samples examined. Based on colony morphology and cell microscopy, 28 isolates were selected for further studies. The strains grown on ISP 2 medium formed rounded (0.3–1.0 cm) colonies of predominantly white, beige, yellowish, orange, brown, and maroon colors.

Several different bacterial morphotypes were observed in microscopy: cocci (4c-3-1, 5c-3-3, 13p-4-1), bacilli (3c-1-1, 6c-4-2), and branched mycelium (all other strains). Most mycelial strains were characterized by the release of water-soluble light yellow and light brown pigments into the medium.

Taxonomy and phylogeny of the isolates

Nine genera of Actinomycetota were identified as a result of 16S rRNA gene sequence analysis. Most isolates belonged to *Streptomyces*, a genus widely distributed in soil. Representatives of the genera *Arthrobacter*, *Glycomyces*, *Kocuria*, *Microbacterium*, *Micromonospora*, *Nocardioides*, *Pseudarthrobacter*, and *Rhodococcus* were also isolated along with them (Table 2). The isolated strains showed 98.10–100 % similarity with the previously described type strains. One isolate that showed low 16S rRNA gene sequence similarity (<98.65 %) with previously isolated and validly described species was a new species of the genus *Glycomyces* (Nikitina et al., 2020).

To analyze the phylogenetic relatedness of the isolates and their closest validly described species, three phylogenetic trees were constructed using neighbor-joining, maximum

Table 1. Physical and chemical properties of soils and the number of bacteria on ISP 4

Soil horizon	Depth, cm	pH_{H_2O}	TOC, %	TN, %	∑ soil particle size, %		Bacterial counts, million CFU/g
					<0.01 mm	<0.001 mm	
Soil profile 1T. Chestnut typical soil (Haplic Kastanozems)							
AJ	0–7(9)	7.54	2.63	0.28	37.17	4.38	4.90
BMK	7(9)–21	7.66	1.30	0.12	22.72	4.16	7.10
CAT	21–39	7.92	0.42	0.04	28.59	5.07	4.52
BCca	39–72	7.68	0.19	–	24.15	4.12	3.09
Soil profile 3T. Light-humic soil (Eutric Leptosols)							
AJ1	0–10(16)	7.02	1.39	0.18	17.40	2.43	2.45
AJ2	10(16)–31(45)	7.17	1.33	0.13	17.30	2.53	2.90
Cca,m	31(45)–44(58)	7.20	0.57	–	17.39	2.54	4.50
Cca	44(58)–79	7.25	0.35	–	21.49	2.94	1.91
Soil profile 4I. Chestnut soil with a buried hydrometamorphosed chernozem profile (Haplic Kastanozems, Calcic Chernozems)							
AJ	0–7	6.85	2.45	0.30	23.43	2.97	3.34
BMK	7–15	7.77	1.21	0.12	21.78	2.97	3.10
[AU]	15–39	7.63	0.70	0.08	24.95	3.10	3.26
[AU/BCA]	39–55	7.69	0.48	0.05	24.07	2.84	4.60
BCAq	55–74	7.86	0.55	0.05	29.14	3.83	1.61
BCq	74–95	7.73	0.12	–	29.67	3.77	0.66
Soil profile 5I. Chestnut quasi-gley soil (Gleyic Kastanozems)							
AJ1	0–7	7.18	3.31	0.42	22.03	2.62	2.87
AJ2	7–18	7.36	1.86	0.19	23.18	3.05	1.96
BMK	18–42	7.57	1.08	0.10	34.31	4.12	2.98
CATq	42–60	7.30	0.78	0.06	28.68	3.45	2.50
BCq	60–75(80)	7.58	0.45	–	30.66	3.89	2.45
Cca,q	75(80)–125	8.02	0.33	–	23.56	2.73	0.78

likelihood, and maximum parsimony methods (Fig. 1). All three phylograms had similar basic topologies. According to the phylogenetic analysis, strains belonging to the genera *Arthrobacter* (5c-3-3, 13p-4-1), *Kocuria* (4c-3-1), *Microbacterium* (8c-1-4), *Micromonospora* (2pp-5-2) and *Rhodococcus* (3c-1-1) apparently belonged to known species: it was confirmed by the high similarity (99.36–100 %) and clustering reliability (86–100 %). Strains 3a-1-3 and 15a-4-5 of the genus *Nocardioides* were combined with *Nocardioides luteus* DSM 43366^T and *Nocardioides albus* ATCC 27980^T with high confidence, respectively. Strain 14p-5-5 belonged to the subcluster uniting strains of the genus *Nocardioides* and demonstrated a low level of similarity with the closest homolog (98.90 %). The reliability of combining this nucleotide sequence with *Nocardioides jensenii* JCM 1364^T into one cluster was 96 %, which implies their phylogenetic proximity. Isolates 16am-5-2 and 6c-4-2 were characterized by a high similarity with already known species. However, the low reliability of nucleotide sequence association between 6c-4-2 and *Pseudarthrobacter phenanthrenivorans* DSM 18606^T, as well as the difference in evolutionary distance between 16am-5-2 and *Pseudarthrobacter scleromae* DSM 17756^T do not allow a clear conclusion on the species identity of these strains.

The remaining strains, according to 16S rRNA gene sequence analysis, belonged to the genus *Streptomyces* and were united in one cluster with all streptomycetes collection strains on the phylogenetic tree. The 16S rRNA gene sequences of strains 6a-3-2, 7a-3-2, and 11a-4-1 were identical to each other (100 %), showing high similarity to *Streptomyces brevispora* KACC 21093^T (99.79 %). The reliability of combining these nucleotide sequences into one cluster was 100 %, indicating that the strains could belong to this species. The same assumption may be true for strains 1a-1-2, 4a-1-4, 8a-3-3, 10a-3-3, 13a-4-3, 13c-5-2, and 4k-1-2, which had a high level

Table 2. Taxonomic position of strains based on 16S rRNA gene sequences analysis

Soil sample	Strain	Fragment length, bp	Nearest homologue	Similarity, %
			Soil profile 1T	
AJ	3c-1-1	785	*Rhodococcus fascians* ATCC 12974^T	99.36
BMK	1a-1-2	880	*Streptomyces pseudogriseolus* ATCC 12770^T	99.89
BMK	4k-1-2	815	*Streptomyces qinglanensis* DSM 42035^T	99.51
CAT	3a-1-3	950	*Nocardioides albus* ATCC 27980^T	99.79
BCca	4a-1-4	940	*Streptomyces violaceochromogenes* DSM 40181^T	99.57
BCca	8c-1-4	1404	*Microbacterium saccharophilum* NBRC 108778^T	100
			Soil profile 3T	
AJ1	4c-3-1	830	*Kocuria rosea* DSM 20447^T	99.64
AJ2	6a-3-2	950	*Streptomyces brevispora* KACC 21093^T	99.79
AJ2	7a-3-2	935	*Streptomyces brevispora* KACC 21093^T	99.79
Cca,m	5c-3-3	932	*Arthrobacter ruber* CGMCC 1.9772^T	99.56
Cca,m	8a-3-3	940	*Streptomyces seymenliensis* DSM 42117^T	99.57
Cca,m	10a-3-3	920	*Streptomyces candidus* DSM 40141^T	99.46
Cca,m	20a-3-3	940	*Streptomyces glomeroaurantiacus* DSM 41782^T	99.26
Cca	9a-3-4	950	*Streptomyces monticola* NEAU-GS4^T	98.63
			Soil profile 4I	
AJ	11a-4-1	950	*Streptomyces brevispora* KACC 21093^T	99.79
AJ	13p-4-1	601	*Arthrobacter humicola* JCM 15921^T	100
BMK	6c-4-2	920	*Pseudarthrobacter phenanthrenivorans* DSM 18606^T	99.02
[AU]	13a-4-3	940	*Streptomyces aureocirculatus* JCM 4454^T	98.83
BCAq	15a-4-5	1394	*Nocardioides luteus* DSM 43366^T	99.57
BCq	18	1455	*Glycomyces paridis* CPCC 204357^T	97.18
			Soil profile 5I	
AJ2	27a-5-2	930	*Streptomyces sioyaensis* ATCC 13989^T	99.68
AJ2	16am-5-2	624	*Pseudarthrobacter scleromae* DSM 17756^T	99.19
AJ2	13c-5-2	758	*Streptomyces gobitricini* NBRC 15419^T	99.08
AJ2	2pp-5-2	970	*Micromonospora luteifusca* DSM 100204^T	99.59
BMK	21a-5-3	1399	*Streptomyces phaeochromogenes* ATCC 23945^T	98.34
BMK	22a-5-3	950	*Streptomyces cremeus* DSM 40147^T	99.05
BMK	28a-5-3	940	*Streptomyces galbus* JCM 4222^T	99.26
BCq	14p-5-5	816	*Nocardioides jensenii* JCM 1364^T	98.90

of similarity with their closest homologs and clustered with them with high reliability.

Strains 27a-5-2 and 28a-5-3 showed high 16S rRNA gene sequence similarity with closely related species (99.68 and 99.26 %, respectively), but the clustering reliability was low. Strain 9a-3-4 showed a relatively low level of similarity (98.63 %) and formed a cluster with the unvalidated species *Streptomyces monticola* NEAU-GS4. Strain 21a-5-3 had a level of similarity with the closest described species below the threshold (98.34 %), and strains 20a-3-3 and 22a-5-3 did not

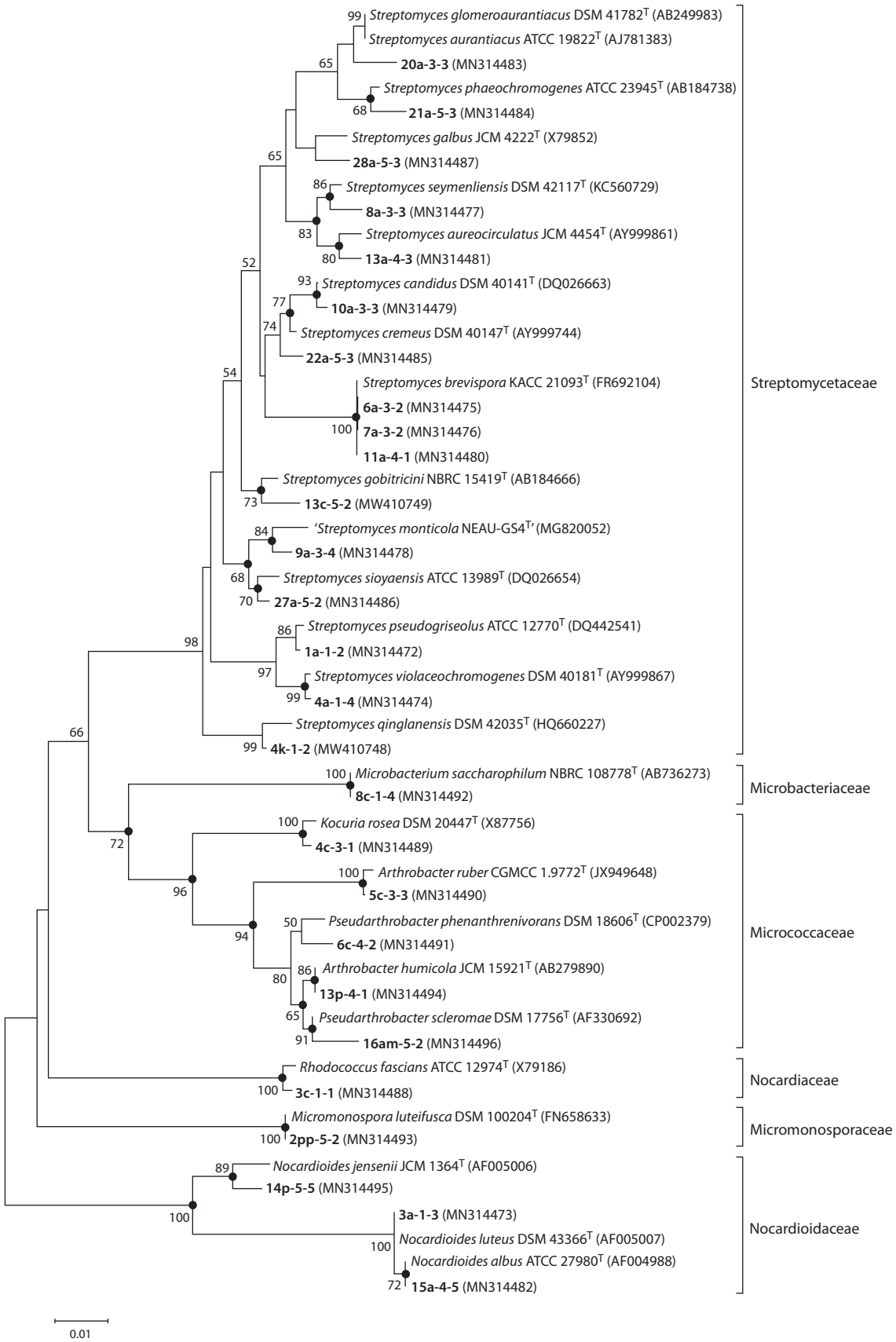


Fig. 1. Neighbour-joining phylogenetic tree showing the phylogenetic position of actinobacteria strains.

Bar, 0.01 substitutions per nucleotide position. Numbers at branch nodes refer to bootstrap values based on 1000 replicates (only values >50 % are shown). Filled circles at nodes indicate corresponding branches that were recovered by using the maximum likelihood and maximum parsimony algorithms.

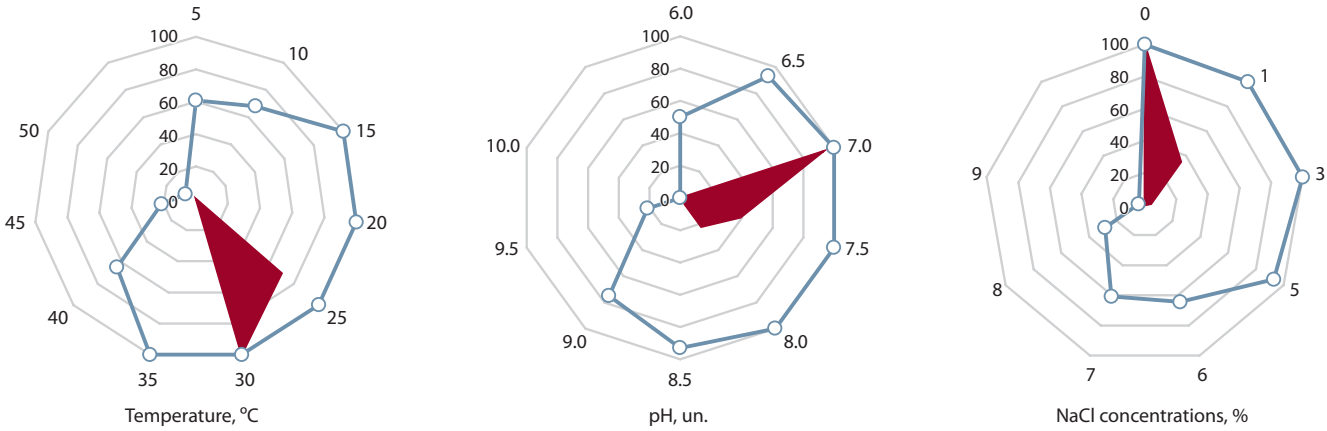


Fig. 2. Basic growth conditions for isolated strains.

The values are given in % of the total number of strains. The selected sector corresponds to the optimal values of the growing conditions.

form a cluster with any collection strains, despite a relatively high level of similarity with the closest homologs. These isolates could probably represent new species of the genus *Streptomyces*. However, the identification of streptomycetes at the species level based solely on the analysis of the 16S rRNA gene is rather complicated: an earlier study by Labeda et al. (2012) showed that the nucleotide sequences of this gene have high similarity for representatives of all taxa within the family Streptomycetaceae. Because of the complex systematics of the genus *Streptomyces*, which currently includes more than seven hundred validly described species, additional tests are needed to accurately determine the species identity of the isolated strains.

Ecophysiological characteristics of the actinobacteria strains

The isolated strains were characterized by different sensitivity to temperature, pH, and NaCl concentration (Fig. 2). The optimal temperature values for growth range from 25 to 30 °C, which allows us to assign the isolated strains to the group of mesophiles. In general, growth was observed in the range from 5 to 45 °C. Regarding pH tolerance, the isolates behaved predominantly as neutrophils, having growth limits from 6 to 9 with an optimum pH of 7–8. The optimal NaCl concentration for growth of most strains was 0 to 1 %. Strains 1a-1-2, 15a-4-5, 20a-3-3, 3c-1-1, 4c-3-1, 13c-5-2, 4k-1-2 and 18 were halotolerant, being able to grow at salt concentration ranging from 0 to 8 %. However, strong growth retardation of the isolated strains was observed at the NaCl concentration of 5 % or more.

Almost all strains were capable of using monosaccharides: glucose, fructose, galactose, D-xylose, and α-rhamnose. Most were able to grow on media with disaccharides (sucrose, D-maltose, lactose) and alcohols (glycerol, mannitol, sorbitol, dulcitol). Less than half of the strains showed the ability to use acetate and succinate. Only a few isolates were able to grow on oxalate (4a-1-4 and 6a-3-2), and citrate (6c-4-2 and 13p-4-1).

The isolated strains were capable of using both organic and inorganic nitrogen. The growth of most strains on meat-peptone broth was accompanied by the release of ammonia and hydrogen sulfide, which indicated their ability to use proteins and amino acids as nitrogen sources. The ability to assimilate ammonium salts and nitrates was detected in almost all strains, except for 4c-3-1 and 5c-3-3, which did not use ammonium salts, and 5c-3-3 with 8c-1-4, which did not use nitrates.

All isolated strains were capable of producing catalase and amylase. The presence of protease, lipase, and cellulase was noted in most isolates. Only a few strains were able to produce urease (Fig. 3).

Discussion

The studied soils are formed in a sharply continental climate with low precipitation and a short period of biological activity. The water regime of chestnut soils depends mainly on atmospheric precipitation and is usually unfavorable due to the light granulometric composition and gravel content in the soil. The studied soils are characterized by low stocks of total organic carbon and total nitrogen, concentrated mainly in the upper humus horizons. All this probably causes a wide distribution of oligotrophic bacteria in the microbial community, particularly mycelial prokaryotes – actinomycetes. The results obtained are consistent with earlier studies in Transbaikalia, which noted that the average actinomycete content in soils of the steppe and dry steppe zones exceeds more than half of the total abundance of cultivated prokaryotes (Nimaeva, 1992; Buyantueva et al., 2014).

As a result of this work, pure strains of actinobacteria were obtained, the closest homologs of which were isolated from soil and plant rhizosphere. Most strains were capable of forming branching mycelium. These are representatives of the genera *Streptomyces*, *Nocardioides*, *Micromonospora*, and *Glycomyces*. According to Zenova et al. (2009), such forms of actinobacteria form the basis of the hydrolytic block of prokaryotic microorganisms in soils with intermittent moisture and nutrient supply regimes. They have advantages over other bacteria, as they are capable of cell differentiation and formation of mycelium able to penetrate through phase boundaries in the soil medium.

More than half of the isolated strains belonged to the genus *Streptomyces*, which is quite natural: this genus is commonly associated with the soil microbiota and is most easily isolated on synthetic nutrient media. Streptomycetes strains were iso-

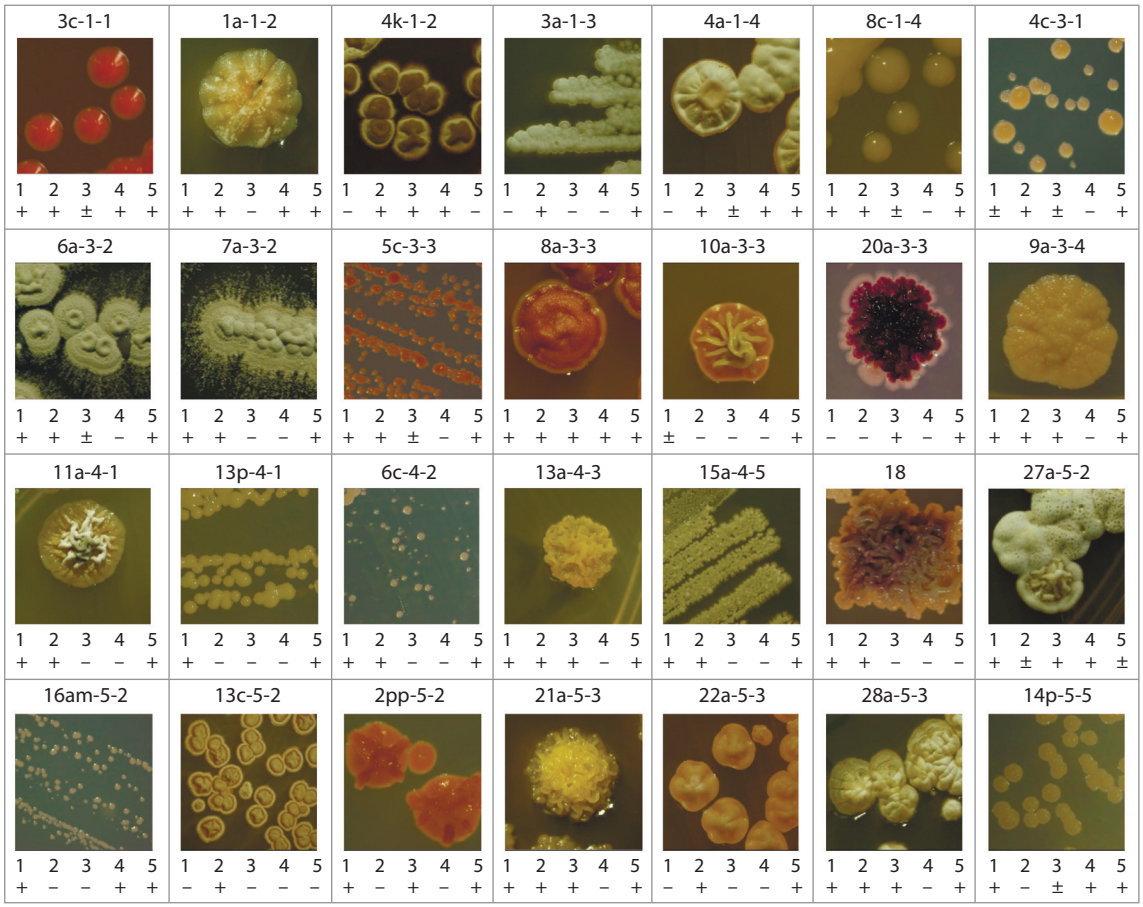


Fig. 3. Morphology of colonies and enzymatic activity of isolated strains.

1 – protease; 2 – lipase; 3 – cellulase; 4 – urease; 5 – H_2S production. (+) – positive reaction, (–) – negative, (±) – weakly positive.

lated from the upper, middle, and lower horizons of all soil sections. The widespread distribution of streptomycetes in soils is due to their mycelial structure, oligotrophy, and ability to produce arthrospores that promote dispersal and help them tolerate stress conditions (Zvyagincev et al., 2005; Cockell et al., 2013). Strains belonging to the genus *Arthrobacter* and the recently separated genus *Pseudarthrobacter* were isolated from the surface and middle horizons of the studied soils. Although their representatives do not form specific dormant forms like streptomycetes, they are capable of surviving under low-nutrient and soil desiccation conditions due to their special strategy and metabolism (Dobrovol'skaya, 2002; Wink et al., 2017). They are capable of forming cyst-like resting cells with extremely reduced metabolism under unfavorable conditions (Wink et al., 2017). One strain each of *Rhodococcus*, *Kocuria*, and *Microbacterium* was also isolated. Actinobacteria belonging to these genera are unable to form spores, but *Rhodococcus*, for example, can form mycelium capable of disintegrating into coccoid or bacilliform elements, which increase species survival (Wink et al., 2017). UV-resistant *Rhodococcus* (Urbano et al., 2013) and radio-resistant, psychrotrophic members of the genus *Kocuria* have also been reported (Asgarani et al., 2012).

All strains under study were isolated at 30 °C, with neutral pH and negligible concentration of sodium chloride in the medium; nevertheless, they demonstrate wide limits of tolerance to these factors. This indicates their high adaptation potential to abiotic factors.

At present, the ability of actinobacteria to assimilate certain sources of carbon and nitrogen is not a significant taxonomic feature, but it can provide a basis for studying the functional role of prokaryotes in the community. The isolates exhibited broad metabolic activity to the substrates, indicating their active participation in the degradation of organic matter. Almost all isolated strains were able to consume mono- and disaccharides, and less frequently, polyatomic alcohols. The isolated strains were capable of using both organic (proteins, amino acids) and inorganic (ammonium salts, nitrates) compounds as sources of nitrogen. Amylolytic and catalytic activity was observed for all strains examined. Most isolates were characterized by proteolytic and lipolytic activity. More than half of the strains produced cellulase, and one-third produced urease.

Conclusion

Certain characteristics of actinobacteria indicate that isolated bacteria play an important role in the degradation of organic matter and also have adaptive capabilities to environmental changes. These characteristics include features of morphology and life cycle (formation of aerial mycelium, spores, and dormant forms), the ability to use various substrates, the presence

of extracellular enzymes, and a wide range of strain growth. For decades, actinobacteria have attracted the attention of many researchers due to their biotechnological potential. Further research could include additional examinations of isolated actinobacteria and evaluate the prospects of their use in biotechnology (in particular, as producers of antimicrobial components). These data not only expand knowledge about the diversity of microbial communities in soils of the Selenga Highlands but also confirm the potential of searching for new actinobacteria species in these soils.

References

Asgarani E., Soudi M.R., Borzooee F., Dabbagh R. Radio-resistance in psychrotrophic *Kocuria* sp. ASB 107 isolated from Ab-e-Siah radioactive spring. *J. Environ. Radioact*. 2012;113:171-176. DOI 10.1016/j.jenvrad.2012.04.009.

Bao Y., Dolfing J., Guo Z., Chen R., Wu W., Li Z., Lin X., Feng Y. Important ecophysiological roles of non-dominant *Actinobacteria* in plant residue decomposition, especially in less fertile soils. *Microbiome*. 2021;9:84. DOI 10.1186/s40168-021-01032-x.

Batuev A.R., Buyantuev A.B., Snytko V.A. Geosystems and Mapping of Ecogeographical Situations in Selenga Basins of the Baikal Region. Novosibirsk, 2000. (in Russian)

Buyantueva L.B., Nikitina E.P., Namsaraev B.B. Actinomycete communities of chestnut soils of steppe pastures in Buryatia. *Vestnik Buryatskogo Gosuniversiteta = Herald of the Buryat State University*. 2014;4(2):55-58. (in Russian)

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T.L. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421. DOI 10.1186/1471-2105-10-421.

Chimitdorzhieva G.D., Chimitdorzhieva E.O. Structure of humic acids in dry-steppe soils of Transbaikalia. *Uspekhi Sovremennogo Yestestvoznaniya = Advances in Modern Natural Science.* 2021;12:89-94. (in Russian)

Cockell C.S., Kelly L.C., Marteinsson V. *Actinobacteria* – an ancient phylum active in volcanic rock weathering. *Geomicrobiol. J*. 2013; 30:706-720. DOI 10.1080/01490451.2012.758196.

DeLong E.F. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA*. 1992;89:5685-5689. DOI 10.1073/pnas.89.12.5685.

Dobrovol'skaya T.G. Structure of Soil Bacterial Communities. Moscow, 2002. (in Russian)

Ecological Atlas of Lake Baikal Basin. Irkutsk, 2015.

Gordon R.E., Bennett D.A., Handerhan J.E., Pang C.H. *Nocardia coeliaca*, *Nocardia autotrophica*, and the nocardin strain. *Int. J. Syst. Bacteriol*. 1974;24:54-63.

Hazarika S.N., Thakur D. Actinobacteria. In: Amaresan N., Senthil Kumar M., Annapurna K., Kumar K., Sankaranarayanan A. (Eds.) Beneficial Microbes in Agro-Ecology. Cambridge, 2020;443-476. DOI 10.1016/B978-0-12-823414-3.00021-6.

Kalakuckiy L.V., Agre N.S. Development of Actinomycetes. Moscow, 1977. (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.

Kurapova A.I., Zenova G.M., Sudnitsyn I.I., Kizilova A.K., Manucharova N.A., Norovsuren Zh., Zviagintsev D.G. Thermotolerant and thermophilic Actinomycetes from soils of Mongolia desert steppe zone. *Microbiology*. 2012;81:98-108. DOI 10.1134/S00262617120 10092.

Labeda D.P., Goodfellow M., Brown R., Ward A.C., Lanoot B., Vanncanneyt M., Swings J., Kim S.-B., Liu Z., Chun J., Tamura T., Oguchi A., Kikuchi T., Kikuchi H., Nishii T., Tsuji K., Yamaguchi Y., Tase A., Takahashi M., Sakane T., Suzuki K.I., Hatano K. Phylogenetic study of the species within the family *Streptomycetaceae*. *Anton. Leeuw. Int. J.* 2012;1(101):73-104.

Lechevalier M.P. Ecological associations involving actinomycetes. In: Schaal K., Pulverer G. (Eds.) Actinomycetes. Stuttgart, New York, 1981;159-166.

Leo V.V., Asem D., Zothanpuia, Singh B.P. Actinobacteria: a highly potent source for holocellulose degrading enzymes in Actinobacteria. In: New and Future Developments in Microbial Biotechnology and Bioengineering. 2018;191-205. DOI 10.1016/b978-0-444-63994-3.00013-8.

Manucharova N.A., Belova E.V., Polyanskaya L.M., Zenova G.M. A chitinolytic actinomycete complex in chernozem soil. *Microbiology*. 2004;73(1):56-59.

McCarthy A.J., Williams S.T. Actinomycetes as agents of biodegradation in the environment – a review. *Gene*. 1992;115(1-2):189-192.

Mineev V.G. (Ed.) Manual on Agrochemistry. Moscow, 2001. (in Russian)

Mohammadipanah F., Wink J. *Actinobacteria* from arid and desert habitats: diversity and biological activity. *Front. Microbiol.* 2016;6: 1-10.

Nikitina E., Liu S.W., Li F.N., Buyantueva L., Abidueva E., Sun C.H. *Glycomyces buryatensis* sp. nov., an actinobacterium isolated from steppe soil. *Int. J. Syst. Evol. Microbiol*. 2020;70:1356-1363. DOI 10.1099/ijsem.0.003923.

Nimaeva S.Sh. Microbiology of Cryoarid Soils. Novosibirsk, 1992. (in Russian)

Nogina N.A. Soils of Transbaikalia. Moscow, 1964. (in Russian)

Shirling E.B., Gottlieb D. Methods for characterization of *Streptomyces* species. *Int. J. Syst. Bacteriol.* 1966;16(3):313-340.

Tepper E.Z. Microorganisms of the genus *Nocardia* and humus decomposition. *Agrokhimiya = Agricultural Chemistry*. 1981;5:156-157. (in Russian)

Urbano S.B., Albarracin V.H., Urbano O.F., Farias M.E., Alvarez H.M. Lipid storage in high-altitude Andean Lakes extremophiles and its mobilization under stress conditions in *Rhodococcus* sp. A5, a UV-resistant actinobacterium. *Extremophiles.* 2013;17(2):217-227. DOI 10.1007/s00792-012-0508-2.

Ventura M., Canchaya C., Tauch A., Chandra G., Fitzgerald G.F., Chater K.F., van Sinderen D. Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. *Microbiol. Mol. Biol. Rev*. 2007;71(3):495-548.

Wang C., Dong D., Wang H., Muller K., Qin Y., Wang H., Wu W. Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. *Biotechnol. Biofuels*. 2016;9:22. DOI 10.1186/s13068-016-0440-2.

Williams S.T., Goodfellow M., Alderson G., Wellington E.M.H., Sneath P.H.A., Sackin M.J. Numerical classification of *Streptomyces* and related genera. *J. Gen. Microbiol*. 1983;129:1743-1819.

Wink J., Mohammadipanah F., Hamedi J. (Eds.) Biology and Biotechnology of Actinobacteria. Berlin: Springer, 2017. DOI 10.1007/978-3-319-60339-1.

Wu C.Y., Zhuang L., Zhou S.G., Li F.B., He J. *Corynebacterium humireducens* sp. nov., an alkaliphilic, humic acid-reducing bacterium isolated from a microbial fuel cell. *Int. J. Syst. Evol. Microbiol*. 2011;61:882-887. DOI 10.1099/ijs.0.020909-0.

Xie F., Pathom-aree W. Actinobacteria from desert: diversity and biotechnological applications. *Front. Microbiol.* 2021;12:765531. DOI 10.3389/fmicb.2021.765531.

Xu P., Li W.J., Tang S.K., Zhang Y.Q., Chen G.Z., Chen H.H., Xu L.H., Jiang C.L. *Naxibacter alkalitolerans* gen. nov., sp. nov., a novel member of the family *Oxalobacteraceae* isolated from China. *Int J. Syst. Evol. Microbiol*. 2005;55:1149-1153. DOI 10.1099/ijs.0.63407-0.

Yaradoddi J.S., Kontro M.H., Banapurmath N.R., Ganachari S.V., Sulochana M.B., Hungund B.S., Kazi Z.K. Anilkumar S.K., Oli A. Extremophilic Actinobacteria. In: Yaradoddi J.S., Kontro M.H., Ganachari S.V. (Eds.) Actinobacteria. Rhizosphere Biology. Singapore: Springer, 2021;55-68. DOI 10.1007/978-981-16-3353-9_4(2021).

Yoon S.H., Ha S.M., Kwon S., Lim J., Kim Y., Seo H., Chun J. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* 2017;67(5):1613-1617. DOI 10.1099/ijsem.0.001755.

Zenova G.M., Dubrova M.S., Gracheva T.A., Kuznetsova A.I., Stepanova O.A., Chernov I.Yu., Manucharov A.S. Ecological and taxonomic features of soil actinomycete complexes in the Near-Elton region. *Vestnik Moskovskogo Universiteta. Ser. 17. Pochvovedenie = Herald of the Moscow University. Series 17: Soil Science*. 2016;4: 43-46. (in Russian)

Zenova G.M., Kozhevin P.A., Manucharova N.A., Lubsanova D.A., Dubrova M.S. Ecophysiological features of actinomycetes in desert soils in Mongolia. *Izvestiya RAN. Ser. Biol. = Proceedings of the Russian Academy of Sciences. Biological Series*. 2014;3:246-253. (in Russian)

Zenova G.M., Kurapova A.I., Zvyagintsev D.G., Lysenko A.M. The structural-functional organization of thermotolerant complexes of actinomycetes in desert and volcanic soils. *Eurasian Soil Sci.* 2009; 42(5):531-535.

Zenova G.M., Zvyagintsev D.G. The Diversity of Actinomycetes in Land Ecosystems. Moscow, 2002. (in Russian)

Zhou S.Q., Huang X.L., Huang D.Y., Hu X. W., Chen J.L. A rapid method for extracting DNA from actinomycetes by Chelex-100. *Biotechnol. Bull*. 2010;2:123-125.

Zvyagincev D.G., Bab'eva I.P., Zenova G.M. Soil Biology. Moscow, 2005. (in Russian)

Zvyagintsev D.G., Dobrovolskaya T.G., Chernov I.Yu., Sardanashvili E.S., Gonchikov G.G., Korsunov V.M. The peculiarities of taxonomic composition of microbial complexes in soils of Baikal Region. *Eurasian Soil Sci.* 1999;6:732-737.

ORCID ID

E.P. Nikitina orcid.org/0000-0003-2431-8999
L.B. Buyantueva orcid.org/0000-0003-2942-4037
E.Yu. Abidueva orcid.org/0000-0001-6312-4076
C.H. Sun orcid.org/0000-0001-6813-8274

Acknowledgements. The work was partially carried out with the support of the state assignment of IGEB SB RAS No. 121030100229-1.

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

Received May 8, 2022. Revised August 1, 2022. Accepted August 3, 2022.

Вавиловский журнал генетики и селекции. 2023;27(4):421-427
DOI 10.18699/VJGB-23-50

The effect of thioredoxin and prochymosin coexpression on the refolding of recombinant alpaca chymosin

S.V. Belenkaya[1, 2] ✉, D.N. Shcherbakov[1, 2], A.I. Chapoval[2], T.I. Esina[1], V.V. Elchaninov[3]

[1] State Research Center of Virology and Biotechnology "Vector", Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
[2] Altai State University, Barnaul, Russia
[3] Federal Altai Scientific Center for Agrobiotechnology, Siberian Research Institute of Cheesemaking, Barnaul, Russia
✉ belenkaya.sveta@gmail.com

Abstract. The milk-clotting enzyme chymosin is a member of the group of aspartate proteinases. Chymosin is the main component of rennet traditionally obtained from the stomachs of dairy calves and widely used to coagulate milk in the production of various types of cheese. Another source of chymosin, which does not require the killing of animals, is based on recombinant DNA technology. Recombinant alpaca chymosin has a number of valuable technological properties that make it attractive for use in cheese-making as an alternative to recombinant bovine chymosin. The purpose of this work is to study the effect of coexpression of thioredoxin and prochymosin on the refolding of the recombinant zymogen and the activity of alpaca chymosin. To achieve this goal, on the basis of the pET32a plasmid, an expression vector was constructed containing the thioredoxin A gene fused to the N-terminal sequence of the marker enzyme zymogen, alpaca prochymosin. Using the constructed vector, pET-TrxProChn, a strain-producer of the recombinant chimeric protein thioredoxin-prochymosin was obtained. The choice of prochymosin as a model protein is due to the ability of autocatalytic activation of this zymogen, in which the pro-fragment is removed, together with the thioredoxin sequence attached to it, with the formation of active chymosin. It is shown that *Escherichia coli* strain BL21 transformed with the pET-TrxProChn plasmid provides an efficient synthesis of the thioredoxin-prochymosin chimeric molecule. However, the chimeric protein accumulates in inclusion bodies in an insoluble form. Therefore, a renaturation procedure was used to obtain the active target enzyme. Fusion of thioredoxin capable of disulfide-reductase activity to the N-terminal sequence of prochymosin provides optimal conditions for zymogen refolding and increases the yield of recombinant alpaca chymosin immediately after activation and during long-term storage by 13 and 15 %, respectively. The inclusion of thioredoxin in the composition of the chimeric protein, apparently, contributes to the process of correct reduction of disulfide bonds in the prochymosin molecule, which is reflected in the dynamics of the increase in the milk-clotting activity of alpaca chymosin during long-term storage.
Key words: thioredoxin (Trx); recombinant chymosin (rChn); inclusion bodies; milk-clotting activity; renaturation.

For citation: Belenkaya S.V., Shcherbakov D.N., Chapoval A.I., Esina T.I., Elchaninov V.V. The effect of thioredoxin and prochymosin coexpression on the refolding of recombinant alpaca chymosin. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(4):421-427. DOI 10.18699/VJGB-23-50

Влияние коэкспрессии тиоредоксина и прохимозина на рефолдинг рекомбинантного химозина альпака

С.В. Беленькая[1, 2] ✉, Д.Н. Щербаков[1, 2], А.И. Шаповал[2], Т.И. Есина[1], В.В. Ельчанинов[3]

[1] Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия
[2] Алтайский государственный университет, Барнаул, Россия
[3] Федеральный Алтайский научный центр агробиотехнологий, Сибирский научно-исследовательский институт сыроделия, Барнаул, Россия
✉ belenkaya.sveta@gmail.com

Аннотация. Молокосвертывающий фермент химозин является представителем группы аспартатных протеиназ. Химозин – основной компонент сычужного фермента, традиционно получаемого из желудков телят-молокопоек и широко используемого для свертывания молока при производстве разнообразных видов сыра. Другой источник химозина, не требующий умерщвления животных, базируется на технологии рекомбинантных ДНК. Рекомбинантный химозин альпака обладает рядом ценных технологических свойств, делающих его привлекательным для использования в сыроделии в качестве альтернативы рекомбинантному химозину коровы. Цель настоящей работы – исследование влияния коэкспрессии тиоредоксина и прохимозина на рефолдинг рекомбинантного зимогена и активность химозина альпака. Для достижения поставленной цели на основе плазмиды pET32a был сконструирован экспрессионный вектор, содержащий ген тиоредоксина А, слитый с N-концевой последовательностью зимогена маркерного фермента – прохимозина альпака. С помощью сконструированного вектора pET-TrxProChn создан штамм-продуцент рекомбинантного химерного белка тиоредоксин-прохимозин. Выбор прохимозина в качестве модельного белка обусловлен способностью

этого зимогена к автокаталитической активации, при которой происходит удаление профрагмента вместе с присоединенной к нему последовательностью тиоредоксина с образованием активного химозина. Показано, что штамм *Escherichia coli* BL21, трансформированный плазмидой pET-TrxProChn, обеспечивает эффективный синтез химерной молекулы тиоредоксин-прохимозин. Однако химерный белок накапливается в тельцах включения в нерастворимой форме. Поэтому для создания активного целевого фермента была использована процедура ренатурации. Присоединение тиоредоксина, обладающего способностью восстанавливать дисульфидные связи, к N-концевой последовательности прохимозина обеспечивает оптимальные условия для рефолдинга зимогена и увеличивает выход рекомбинантного химозина альпака сразу после активации и при длительном хранении на 13 и 15 % соответственно. Включение тиоредоксина в состав химерного белка, по всей видимости, способствует процессу корректного восстановления дисульфидных связей в молекуле прохимозина, что отражается на динамике роста молокосвертывающей активности химозина альпака при длительном хранении.
Ключевые слова: тиоредоксин (Trx); рекомбинантный химозин (rChn); тельца включения; молокосвертывающая активность; ренатурация.

Introduction

The development of genetic engineering and biotechnology methods makes it possible to obtain a plethora of technological and therapeutic proteins not from raw natural materials, but by synthesizing their recombinant analogs in various expression systems. Prokaryotes use for producing recombinant proteins is well studied and has important features such as rapid growth, a high level of protein synthesis in addition to a simple culturing protocol. Even though prokaryotes do not perform some post-translational modifications (glycosylation, phosphorylation, proteolytic processing) of synthesized protein as eukaryotes, these expression systems are used for industrial-scale production of various recombinant proteins (Huang et al., 2012; Rosano, Ceccarelli, 2014; Baeshen et al., 2015). This method for recombinant protein production provides a more simple and cost-effective approach compared to obtaining and maintaining stable eukaryotic producers (Rosano et al., 2019). The *Escherichia coli* expression system was used in our laboratory to produce and characterize new milk-clotting enzymes encoded by alpaca and maral chymosins genes (Belenkaya et al., 2018, 2020a, b).

When a heterologous protein is expressed in the *E. coli* system, there are several scenarios for its final localization: the synthesis and transport of the native product into periplasmic space; the accumulation in an inactive (aggregated) state in inclusion bodies; accumulation of the target recombinant protein in the cytoplasm and sometimes secretion into the culture medium (Baneyx, 1999). In the systems using strong promoters (such as the T7 promoter) and the corresponding *E. coli* strains, proteins usually accumulate in inclusion bodies and can be extracted with denaturing and chaotropic agents. To restore the native conformation of the denatured protein obtained from inclusion bodies, the renaturation (refolding) procedure is required. The key point of refolding is the gradual removal of denaturing and chaotropic components from the solution.

In systems with reversed micelles described earlier (Sakono et al., 2004), refolding can be carried out using dialysis, gel filtration (Li et al., 2004) or adsorption chromatography (Nara et al., 2009). Removal of the denaturating or chaotrope agents does not always lead to the restoration of the target product's native structure. To increase the efficiency of renaturation, attempts were made to use proteins with chaperone activity (Wei et al., 2000; Rosano, Ceccarelli, 2014). Construction of a recombinant protein containing a whole or a part of chaperon protein followed by the chaperone separation from the target product can be an effective strategy (Li, Sousa, 2012; Emamipour et al., 2019; Rosano et al., 2019).

One of the widely used proteins with chaperone activity is thioredoxin (Trx). Thioredoxin is a product of trxA gene, a small enzyme that exhibits chaperone properties. It is important to note that this protein also has disulfide-reductase activity that participates in reversible oxidation of cysteine SH-groups to disulfide. In the case of Chn – chymosin (s), this is of particular importance because it contains three disulfide bridges and their correct closure is important for its enzymatic activity (Chen et al., 2000). Coexpression of Trx and the target protein enhances production, accelerates correct folding, increases solubility, and improves the functional properties of some cysteine-containing proteins and polypeptides such as: scFv antibodies (Jurado et al., 2006), *Balanus albicostatus* adhesive protein (Balcp19k) (Liang et al., 2015), full-length and fragmented tissue plasminogen activator (Bessette et al., 1999). We hypothesized that the expression of a chimeric molecule containing the thioredoxin and prochymosin sequence (Trx-ProChn) may affect its localization and refolding.

Chymosin (EC 3.4.23.4), an aspartic endopeptidase, is widely used in the cheese-making industry for curdling milk and producing a milk curd. The native structure of chymosin (Chn) is stabilized by three -S-S- bridges, which is an important process during renaturation of recombinant protein produced in the *E. coli* expression system (Tang et al., 1994; Chen et al., 2000). We choose Chn as a model protein for verification effects of Trx on enzymes production because of the presence of disulfide bridges in the enzyme structure and possibility of easy Trx removing from recombinant fusion protein at pH 2.5–3.0 after autocatalytic breaking of the peptide bond in the position F58-G59. That makes this model simple, economical and conventional for checking the enzymatic activity of Chn after refolding.

The aim of this work was to study dynamics of changes in the milk-clotting activity (MA) of rChn – recombinant chymosin (s) synthesized as part of the fusion sequence of recombinant thioredoxin-prochymosin (rTrx-ProChn) or in the form of recombinant prochymosin (rProChn).

Materials and methods

Construction of plasmids, production, and cultivation of producer strains. Earlier, we constructed a plasmid vector containing the sequence of the alpaca prochymosin gene

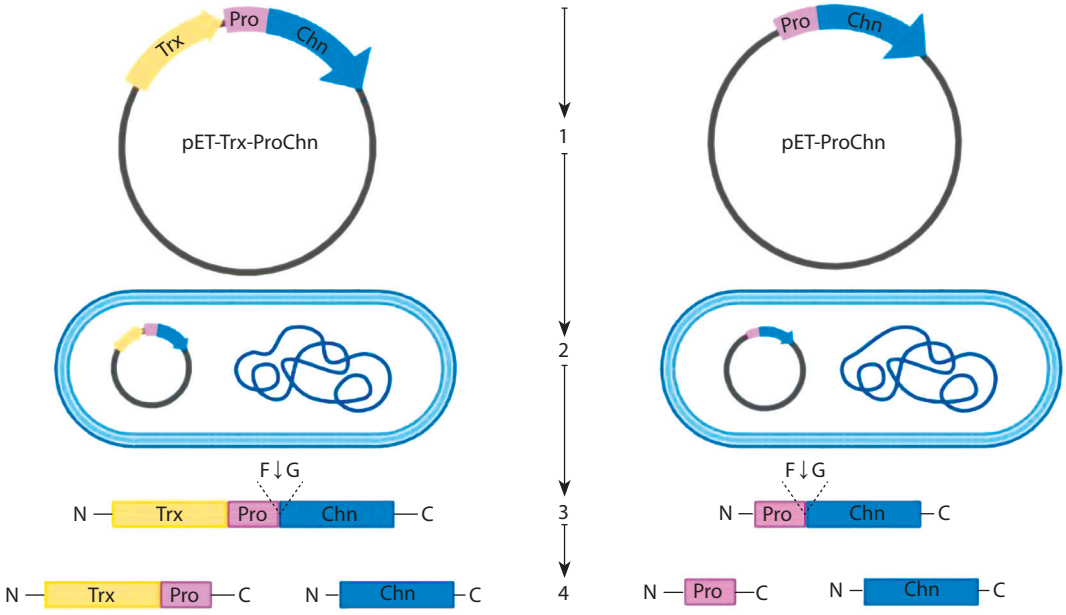


Fig. 1. Production and cleavage of rTrx-ProChn and rProChn proteins.

1 – pET-Trx-ProChn and pET-ProChn recombinant plasmid construction; 2 – *E. coli* transformation; 3 – renaturation and activation of rTrx-ProChn and rProChn; 4 – rTrx-ProChn and rProChn cleavage at F58-G59 site.

with optimized codon composition for the *E. coli* expression system (Belenkaya et al., 2018). To construct a vector containing the thioredoxin sequence, the alpaca prochymosin gene was amplified using a pair of primers Trx-Vic-F 5′-CAGC GGTATTACCAGAATCCCAC-3′ and Trx-Vic-R 5′-AAA AAAAAGCTTCTAAATGGCTTTGGCCAG-3′. Amplification was carried out according to the following program: 95 °C – 5 min (one cycle), 95 °C – 30 s, 59 °C – 30 s, 72 °C – 60 s (30 cycles). The resulting PCR product (1095 bp long) was cloned into the pET32a vector at the unique Msp20I and HindIII restriction sites so that the alpaca prochymosin gene was located in the same reading frame as the thioredoxin gene.

The *E. coli* BL21 (DE3) strain (Invitrogen Corp., USA) was chemically transformed with pET-TrxProChn and pET-ProChn plasmids (Fig. 1).

The culture conditions for both producers and the procedures used for the renaturation of both recombinant proteins were identical. Individual colonies containing pET-TrxProChn and pET-ProChn plasmids were cultured in LB medium containing 100 μg/ml ampicillin overnight on an orbital shaker at 180 rpm and 37 °C. The inoculum (1.0 % v/v) was transferred to an Erlenmeyer flask containing LB medium and grown to an optical density of 0.8 at 600 nm. Next, the inductor, isopropyl-β-D-1-thiogalactopyranoside, was added to the inoculum, to a final concentration of 0.1 mM, and additionally cultured on a shaker (180 rpm) for 6 h at 37 °C.

Isolation and solubilization of inclusion bodies. The cell biomass was separated from the culture medium by centrifugation at 5000 g and 4 °C for 20 min. The bacterial pellet was resuspended in STET buffer (8 % sucrose; 50 mM Tris; 20 mM EDTA; 1 % Triton X-100 pH 8.0) at the rate of 20 ml of buffer per gram of biomass and incubated overnight at 4 °C; at the end of the incubation, the cells were disrupted using an ultrasonic homogenizer Soniprep 150 Plus, and inclusion bodies were pelleted by centrifugation at 20,000 g and 4 °C for 20 min.

The pelleted inclusion bodies were dissolved in Buffer A (50 mM KH_2PO_4, 150 mM NaCl, pH 10.7) containing 8 M urea, which was added in a ratio of 15 ml of buffer per gram of pellet and incubated for 2 h at 30 °C. The resulting solution containing chymosin was centrifuged at 20,000 g and 4 °C for 15 min. The supernatant was separated and protein concentration was determined by the Bradford method (Bradford, 1976).

Renaturation of the target proteins was performed according to three protocols (Table 1). The supernatant (target protein solution) was diluted with buffer A to a final protein concentration of 180 μg/ml. In protocol No. 1, the resulting solution was incubated for 24 h at 4 °C, in protocols No. 2 and 3, incubation was carried out for 24 h at 15 °C. At the end of incubation, the pH of the solution was adjusted to 8.0 using 1M HCl solution, it was kept for 1 h at 15 °C and dialyzed against buffer B (50 mM Tris HCl, 150 mM NaCl, pH 8.0).

The ratio of target protein solution volumes and buffer, the duration of dialysis, the frequency of changing the buffer and some other parameters varied depending on the renaturation protocols (see Table 1). As a result, the obtained solutions contained rProChn and rTrx-ProChn variants of the target protein.

Determining the localization of recombinant proteins. The localization of rProChn and rTrx-ProChn proteins was determined by electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE) according to the Laemmli method. Mixes of standard proteins from the PageRuler Unstained Protein Ladder kit (Thermo Scientific, USA) were used as molecular weight markers (MW). To determine relative content of the target protein on electrophoregrams, the Gel-Pro Analyzer 3.1 software was used.

Table 1. Chymosin proteins production and protocols for renaturation

Proteins production steps	Protocol No. 1	Protocol No. 2	Protocol No. 3
Inclusion bodies isolation	1. Sedimentation from biomass and dilution in STET buffer 2. Disintegration by ultrasound and centrifugation 3. Dissolving in Buffer A with 8 M urea and pH 10.7		
Dilution of target protein solution	Buffer A to a final protein concentration of 150 μg/ml		
Incubation	24 h at 4 °C	24 h at 15 °C	
pH change	Bringing the pH of the solution to 8.0		
Target protein and buffer ratio (v/v)	1/20	1/10	1/5
Dialysis stirring	No		Yes
Buffer change frequency	Every 12 h		
Duration of dialysis	36 h	72 h	

Activation of rProChn and rTrx-ProChn. The activation of rProChn and the chimeric protein rTrx-ProChn was carried out in parallel, by the method of stepwise change in pH (Marciniszyn et al., 1976). As a result, samples of rChn alpaca, activated from rProChn (without Trx) – designated as rChn (Trx–), and samples of rChn alpaca, activated from rTrx-ProChn – designated as rChn (Trx+), were obtained.

Evaluation of total and relative milk-clotting activity. A 10.0 % solution of standardized skimmed milk powder (MZSF OJSC, Russia) in 5 mM $CaCl_2$, pH 6.5, was used as a substrate. A 0.5 % aqueous solution of a dry bovine rChn with a certified MA value was used as a control. Prior to determining the MA, the control sample and the liquid preparation of rChn were kept in a water bath at 35 °C for 15 min and cooled to room temperature. The procedure for determining MA was carried out in a water bath at 35 °C. Substrate solution (2.5 ml) was placed into a glass tube and heated at 35 °C for 5 min. An aliquot (0.25 ml) of an enzyme was added to the substrate, a stopwatch was turned on, and the resulting reaction mixture was immediately thoroughly mixed. The time when the first flakes of the coagulated substrate were observed in the drops of the reaction mixture applied onto the tube wall was considered to be the clotting time. The milk-clotting activity was expressed in arbitrary units (AU) per 1 ml (AU/ml) and calculated using the equation:

$$MA\ (AU/ml) = 0.005*C*T1/T2,$$

C = certified MA value of the control rChn sample in AU per 1 gram, 0.005 = the dilution factor, T1 = coagulation time for the control rChn sample of chymosin, T2 = coagulation time for the test rChn sample. Determination of total MA in each sample was performed in triplicate ($n = 3$).

Dynamics of relative milk-clotting activity. After activation and determination of the starting MA, samples were stored in plastic tubes at a temperature of 8 ± 1 °C. After 17, 25, 43, 60 days, the relative MA of the samples was determined. To visualize the obtained data, graphs of the dependence of the relative MA on the duration of storage were plotted. The starting MA values were taken as 100 %.

Statistical processing of the obtained data was carried out in the computing environment of an Excel 2007 spreadsheet processor (Microsoft Corporation, USA). For quantitative variables, the results are presented as the arithmetic mean (M) with an indication of standard deviation (± SD). M was used for plotting, values ± SD are given in the tables.

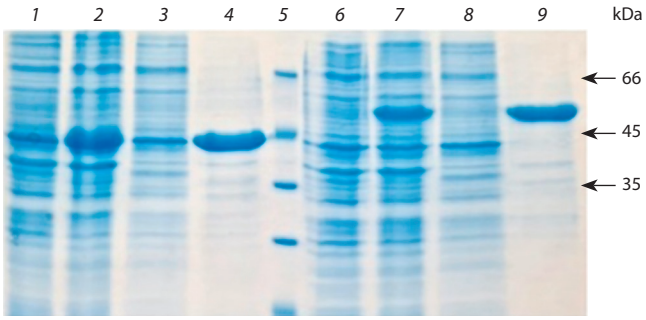


Fig. 2. SDS-PAGE analysis of protein preparations from *E. coli* transfected with pET-ProChn (lanes *1–4*) or pET-Trx-ProChn (lanes *6–9*).

Lanes *1* and *6* – cell biomass before adding the inducer; lanes *2* and *7* – cell biomass 6 h after adding the inducer; lanes *3* and *8* – soluble fraction of cell biomass after treatment with STET buffer and centrifugation; lanes *4* and *9* – insoluble fraction (inclusion bodies) of cell biomass after treatment with buffer A and centrifugation; lane *5* – molecular weight markers.

Results

Construction of plasmids and producer strains

The pET production system was chosen as one of the most convenient systems for expression of recombinant proteins in *E. coli* (Hayat et al., 2018). A specific feature of this expression system is that the target genes are cloned into specialized pET plasmids under a strong promoter control of bacteriophage T7. The promoter is specifically recognized by T7 RNA polymerase and is not recognized by *E. coli* RNA polymerases. Thus, expression of the target gene is induced in the presence of a source of T7 RNA polymerase in the host cell. RNA synthesis by T7 RNA polymerase occurs so selectively and efficiently that almost all of the cell's resources are switched to this process. The content of the target product can reach 50 % of total cellular protein (Chen, 2012).

In our work, the relative content of target proteins (Chn) in case of rProChn producer and rTrx-ProChn was respectively 36 and 19 % of the total amount of protein (Fig. 2, lanes *2* and *7*).

Recombinant proteins production

The bacterium *E. coli* strain BL21 was chosen as the producer of the chimeric protein. The characteristic features of this expression system are: 1) strict control of protein synthesis;

2) high growth rate of culture and high yield of recombinant proteins; 3) high density of viable bacteria in culture; 4) easy of exogenous DNA transformation (Chen, 2012).

To obtain the target protein, *E. coli* BL21 strain was transformed with the recombinant plasmid pET-TrxProChn. Then induction was performed by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) (Studier et al., 1990). During cultivation, three different temperatures were used, 18, 25 and 37 °C. However, since no differences in the level and localization of the protein between them were observed, all the main experiments on protein production were carried out at 37 °C.

Localization of recombinant proteins

SDS-PAGE electrophoresis of proteins indicates that the target protein is localized in the insoluble fraction of cell biomass in the form of inclusion bodies (see Fig. 2, lanes *4* and *9*). Thus, the introduction of the Trx sequence into the structure of the alpaca Chn gene does not lead to the accumulation of the target protein in the periplasmic space.

We have previously shown that when receiving pChn of alpaca, cow and Altai maral in the expression system of *E. coli* (strain BL21) using the pET21 vector, the target proteins also accumulate in the producer cells in the form of inclusion bodies (Belenkaya et al., 2018, 2020a, b).

Chymosin concentration

The relative content of target proteins in the inclusion body fractions was 82 % for rProChn and 84 % for rTrx-ProChn (see Fig. 2, lanes *4* and *9*). After dilution at the refolding stage, the concentration of rProChn and rTrx-ProChn in the solution was ≈150 μg/ml. The actual MW of rProChn and rTrx-ProChn differed by 1.3 times and amounted to 40.5 kDa and 52.5 kDa, respectively, which corresponded to predicted values. Therefore, to calculate the relative MA of Chn derived from rTrx-ProChn rChn (Trx+), we used a lowering factor of 1.3 to reflect the actual enzyme concentration. Consequently, for the calculation of the relative MA of rChn (Trx–) we used concentration of 150 μg/ml while for rChn (Trx+) we considered actual enzyme concentration at 114 μg/ml.

Activation of recombinant proteins

To prevent the autocatalytic activation of the target proteins, after the completion of refolding, the rProChn and rTrx-ProChn samples were stored at pH 8.0 (buffer B). To activate zymogens, the titration procedure to pH = 3.0 was chosen, since in preliminary experiments maximum MA values were observed at this pH. The final pH value of 5.8 was selected as chymosin is stable at moderately acidic pH values = 5.3–6.3 and loses activity at pH > 6.5 (http://www.brenda-enzymes.org/enzyme.php?ecno=3.4.23.4). Before the start of activation, MA of the samples was < 1.0 CU/ml.

Data presented in Table 2 indicate that immediately after activation, the relative MA of the rChn (Trx–) and rChn (Trx+) samples increases by more than 700 times, which indicates the efficiency of the procedure of zymogen activation.

Influence of refolding parameters on the activity of Chns, derived from rProChn and rTrx-ProChn

The MA changes in the rChn (Trx–) and rChn (Trx+) samples was monitored in order to track the possible prolonged effect of Trx on the activity of marker enzymes after refolding and activation of zymogens.

Using refolding protocol No. 1 and 2, starting (immediately after activation) and finishing (after long-term storage) values of the MA in rChn (Trx–) were higher than in rChn (Trx+). At the same time, final coagulation abilities of rChn (Trx–) and rChn (Trx+) were different (see Table 2). After 60 days of incubation, the relative MA of rChn (Trx–) increased by 1.9–2.1 times, while the activity of the rChn (Trx+) samples, during the same time, increased by more than 2.5 times.

The main difference of protocol No. 3 from protocols No. 1 and No. 2 was the reduced ratio of the target protein and dialysis buffer volumes (see Table 1). It is curious that under the conditions of protocol No. 3, the efficiency of refolding of the rProChn sample decreases and turns out to be lower than when using refolding protocols No. 1 and 2 (see Table 2). Immediately after activation, the relative MA of the rChn (Trx+) sample, renatured according to protocol No. 3, was significantly ($p > 0.05$) higher than activity of rChn (Trx–) obtained using protocols No. 2 and 1.

Table 2. Changes in the relative MA of rChn (Trx–) and rChn (Trx+) obtained using the renaturation protocols No. 1–3

Samples	Relative MA in CU/μg (the multiplicity of the increase in activity, compared with the starting value)				
	Start (0 day)	17 day	25 day	43 day	60 day
Protocol No. 1					
rChn (Trx–)	809 ± 2 (1.00)	873 ± 4 (1.08)	888 ± 5 (1.10)	1075 ± 10 (1.33)	1554 ± 10 (1.92)
rChn (Trx+)	579 ± 7 (1.00)	792 ± 7 (1.37)	1056 ± 6 (1.82)	1223 ± 7 (2.11)	1490 ± 15 (2.57)
Protocol No. 2					
rChn (Trx–)	818 ± 3 (1.00)	824 ± 3 (1.01)	917 ± 4 (1.12)	1110 ± 7 (1.36)	1701 ± 9 (2.08)
rChn (Trx+)	543 ± 4 (1.00)	723 ± 4 (1.33)	895 ± 7 (1.65)	1127 ± 8 (2.08)	1396 ± 11 (2.57)
Protocol No. 3					
rChn (Trx–)	767 ± 4 (1.00)	772 ± 6 (1.05)	810 ± 4 (1.06)	1100 ± 12 (1.43)	1490 ± 12 (1.94)
rChn (Trx+)	869 ± 4 (1.00)	1190 ± 7 (1.37)	1240 ± 9 (1.43)	1420 ± 10 (1.63)	1710 ± 12 (1.97)

Note. MA – milk-clotting activity; CU – conventional units.

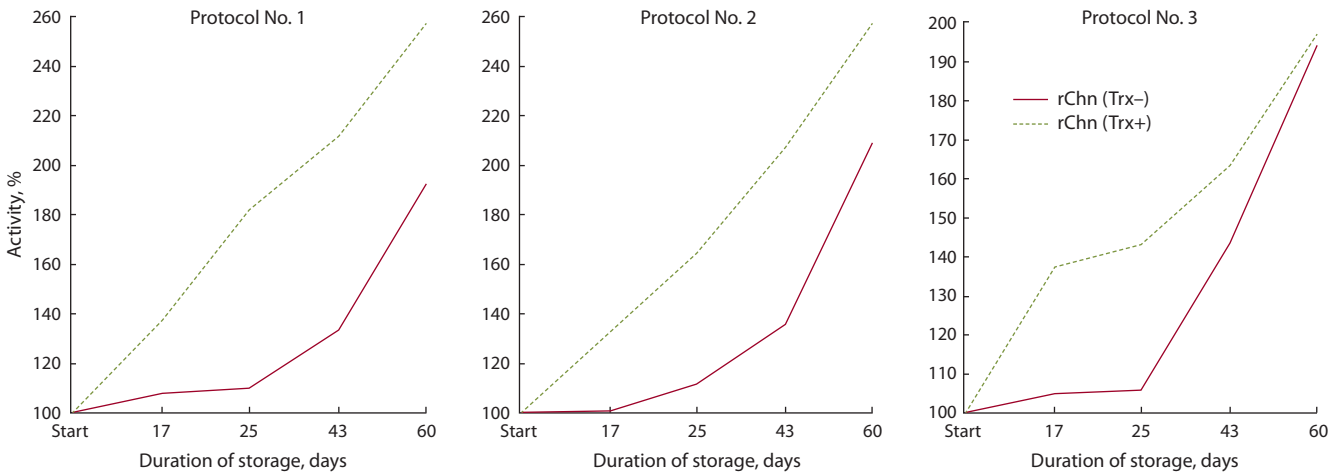


Fig. 3. Dependency of the relative MA on the duration of storage of rChn (Trx–) and rChn (Trx+), obtained using the renaturation protocols No. 1–3.

Utilization of the renaturation protocol No. 3 makes it possible to increase efficiency of the rTrx-ProChn refolding stage (see Table 2). As a consequence, immediately after activation of zymogens and after 60 days of incubation, the values of the relative MA in case of the rChn (Trx+) were 13 and 15 % higher than in the rChn (Trx–). This allows us to conclude that in the expression system used by us, application of Trx under certain refolding conditions (protocol No. 3) makes it possible to elevate the MA of rChn (Trx+) both immediately after the activation of the zymogen and in the process of long-term storage of the enzyme.

During long-term storage, there was a constant increase in the MA of rChn (Trx–) and rChn (Trx+) samples, probably due to refolding of a part of the enzyme molecules that did not keep the correct tertiary structure after the renaturation procedure. To analyze the dynamics of MA in the rChn (Trx–) and rChn (Trx+) obtained using different renaturation protocols, a graph of MA (%) versus storage duration was plotted. The data presented in Fig. 3 show that the dynamics of MA changes are different, especially in the first 25–43 days of storage. Since MA of rChn (Trx+) is ahead of rChn (Trx+) in terms of growth dynamics, it can be assumed that the introduction of Trx into the structure of zymogen additionally stimulates the refolding of the enzyme during long-term storage.

Conclusion

Thus, it was found that in the *E. coli* expression system (strain BL21), attachment of Trx to the N-terminal region of ProChn in alpaca provides accumulation of the target protein exclusively in the form of inclusion bodies.

Coexpression of thioredoxin and prochymosin in the composition of the chimeric molecule rTrx-ProChn and optimization of the conditions of zymogen refolding make it possible to increase the yield of alpaca rChn immediately after activation and during long-term storage by ~13 and ~15 %, respectively.

References

Baeshen M.N., Al-Hejin A.M., Bora R.S., Ahmed M.M.M., Ramadan H.A.I., Saini K.S., Baeshen N.A., Redwan E.M. Production of biopharmaceuticals in *E. coli*: Current scenario and future perspectives. *J. Microbiol. Biotechnol*. 2015;25(7):953-962. DOI 10.4014/jmb.1412.12079.

Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* 1999;10(5):411-421. DOI 10.1016/S0958-1669(99)00003-8.

Belenkaya S.V., Bondar A.A., Kurgina T.A., Elchaninov V.V., Bakulina A.Yu., Rukhlova E.A., Lavrik O.I., Ilyichev A.A., Shcherbakov D.N. Characterization of the Altai maral chymosin gene, production of a chymosin recombinant analog in the prokaryotic expression system, and analysis of its several biochemical properties. *Biochemistry (Moscow).* 2020a;85(7):781-791. DOI 10.1134/S0006297920070068.

Belenkaya S.V., Rudometov A.P., Shcherbakov D.N., Balabova D.V., Kriger A.V., Belov A.N., Koval A.D., Elchaninov V.V. Biochemical properties of recombinant chymosin in alpaca (*Vicugna pacos* L.). *Appl. Biochem. Microbiol.* 2018;54(6):569-576. DOI 10.1134/S0003683818060054.

Belenkaya S.V., Shcherbakov D.N., Balabova D.V., Belov A.N., Koval A.D., Elchaninov V.V. Production of maral (*Cervus elaphus sibiricus* Severtzov) recombinant chymosin in the prokaryotic expression system and the study of the aggregate of its biochemical properties relevant for the cheese-making industry. *Appl. Biochem. Microbiol.* 2020b;56(6):647-656. DOI 10.1134/S0003683820060034.

Bessette P.H., Åslund F., Beckwith J., Georgiou G. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl. Acad. Sci. USA.* 1999;96(24):13703-13708. DOI 10.1073/pnas.96.24.13703.

Bradford M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 1976;72(1-2):248-254. DOI 10.1016/0003-2697(76)90527-3.

Chen H., Zhang G., Zhang Y., Dong Y., Yang K. Functional implications of disulfide bond, Cys206−Cys210, in recombinant prochymosin (chymosin). *Biochemistry.* 2000;39(40):12140-12148. DOI 10.1021/bi000976o.

Chen R. Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol. Adv.* 2012;30(5):1102-1107. DOI 10.1016/j.biotechadv.2011.09.013.

Emamipour N., Vossoughi M., Mahboudi F., Golkar M., Fard-Esfahani P. Soluble expression of IGF1 fused to DsbA in *SHuffle™ T7* strain: optimization of expression and purification by Box-Behnken design. *Appl. Microbiol. Biotechnol*. 2019;103(8):3393-3406. DOI 10.1007/s00253-019-09719-w.

Hayat S.M., Farahani N., Golichenari B., Sahebkar A. Recombinant protein expression in *Escherichia coli* (*E. coli*): what we need to know. *Curr. Pharm. Des*. 2018;24(6):718-725. DOI 10.2174/1381612824666180131121940.

Huang C.J., Lin H., Yang X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J. Ind.*

Microbiol. Biotechnol. 2012;39(3):383-399. DOI 10.1007/s10295-011-1082-9.

Jurado P., de Lorenzo V., Fernández L.A. Thioredoxin fusions increase folding of single chain Fv antibodies in the cytoplasm of *Escherichia coli*: Evidence that chaperone activity is the prime effect of thioredoxin. *J. Mol. Biol*. 2006;357(1):49-61. DOI 10.1016/j.jmb.2005.12.058.

Li M., Su Z.G., Janson J.C. *In vitro* protein refolding by chromatographic procedures. *Protein Expr. Purif.* 2004;33(1):1-10. DOI 10.1016/j.pep.2003.08.023.

Li Y., Sousa R. Expression and purification of *E. coli* BirA biotin ligase for *in vitro* biotinylation. *Protein Expr. Purif*. 2012;82(1):162-167. DOI 10.1016/j.pep.2011.12.008.

Liang C., Li Y., Liu Z., Wu W., Hu B. Protein aggregation formed by recombinant cp19k homologue of *Balanus albicostatus* combined with an 18 kDa N-terminus encoded by pET-32a(+) plasmid having adhesion strength comparable to several commercial glues. *PLoS One*. 2015;10(8):e0136493. DOI 10.1371/journal.pone.0136493.

Marciniszyn J., Huang J.S., Hartsuck J.A., Tang J. Mechanism of intramolecular activation of pepsinogen. Evidence for an intermediate δ and the involvement of the active site of pepsin in the intramolecular activation of pepsinogen. *J. Biol. Chem*. 1976;251(22):7095-7102. DOI 10.1016/s0021-9258(17)32946-0.

Nara T.Y., Togashi H., Sekikawa C., Kawakami M., Yaginuma N., Sakaguchi K., Mizukami F., Tsunoda T. Use of zeolite to refold a disulfide-bonded protein. *Colloids Surf. B Biointerfaces*. 2009;68(1):68-73. DOI 10.1016/j.colsurfb.2008.09.012.

Rosano G.L., Ceccarelli E.A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol*. 2014;5:172. DOI 10.3389/fmicb.2014.00172.

Rosano G.L., Morales E.S., Ceccarelli E.A. New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Sci.* 2019;28(8):1412-1422. DOI 10.1002/pro.3668.

Sakono M., Kawashima Y.M., Ichinose H., Maruyama T., Kamiya N., Goto M. Efficient refolding of inclusion bodies by reversed micelles. *Biotechnol. Prog.* 2004;20(6):1783-1787. DOI 10.1252/kakoronbunshu.30.468.

Studier F.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 1990;185:60-89. DOI 10.1016/0076-6879(90)85008-C.

Tang B., Zhang S., Yang K. Assisted refolding of recombinant prochymosin with the aid of protein disulphide isomerase. *Biochem. J.* 1994;301(1):17-20. DOI 10.1042/bj3010017.

Wei C., Zhang Y., Yang K. Chaperone-mediated refolding of recombinant prochymosin. *J. Protein Chem.* 2000;19(6):449-456. DOI 10.1023/a:1026593113633.

ORCID ID

S.V. Belenkaya orcid.org/0000-0002-3883-6335
D.N. Shcherbakov orcid.org/0000-0001-8023-4453
A.I. Chapoval orcid.org/0000-0002-3426-1036
T.I. Esina orcid.org/0000-0001-9006-8313

Acknowledgements. Sanger DNA sequencing was performed in Genomics Core Facility (Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk).
The study was supported by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-15-2021-1355 dated October 12, 2021) as part of the implementation of certain activities of the Federal Scientific and Technical Program for the Development of Synchrotron and Neutron Research and Research Infrastructure for 2019–2027.

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

Received August 23, 2022. Revised December 2, 2022. Accepted December 7, 2022.

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, Russian Science Citation Index на платформе Web of Science, Российский индекс научного цитирования, ВИНИТИ, Web of Science CC, Scopus, PubMed Central, DOAJ, ROAD, Ulrich's Periodicals Directory, Google Scholar.

Открытый доступ к полным текстам:
русскоязычная версия – на сайте ИЦиГ СО РАН, https://vavilovj-icg.ru/
и платформе Научной электронной библиотеки, elibrary.ru/title_about.asp?id=32440

англоязычная версия – на сайте vavilov.elpub.ru/index.php/jour
и платформе PubMed Central, https://www.ncbi.nlm.nih.gov/pmc/journals/3805/

✉ 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 22.06.2023. Выход в свет 30.06.2023. Формат 60 × 84 $^1/_8$. Усл. печ. л. 15.81.
Уч.-изд. л. 17.9. Тираж 150 экз. (1-й завод 1–45 экз.) Заказ № 140. Цена свободная.

Отпечатано в Сибирском отделении РАН, Морской проспект, 2, Новосибирск, 630090.