Научный рецензируемый журнал

# ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

*Основан в 1997 г.*
*Периодичность 8 выпусков в год*

**Учредители**

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

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

Сибирское отделение Российской академии наук

2018 • 22 • 5

Scientific Peer Reviewed Journal

# VAVILOV JOURNAL OF GENETICS AND BREEDING

## Vavilovskii Zhurnal Genetiki i Selektsii

*Founded in 1997*
*Published 8 times annually*

**Founders**
Federal State Budget Scientific Institution "The Federal Research Center Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences"
The Vavilov Society of Geneticists and Breeders
Siberian Branch of the Russian Academy of Sciences

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

# СОДЕРЖАНИЕ • 2018 • 22 • 5

497 ОТ РЕДАКТОРА

## Молекулярная и клеточная биология

498 ОБЗОР
Современные подходы к синтезу генов: аспекты синтеза олигонуклеотидов, ферментативной сборки, проверки последовательностей и коррекции ошибок
*Г.Ю. Шевелев, Д.В. Пышный*

## Генетика растений

507 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Интрогрессии и хромосомные перестройки не влияют на активность глиадинкодирующих генов в линиях гибридов *Triticum aestivum* L. × *Aegilops columnaris* Zhuk.
*А.Ю. Новосельская-Драгович, А.А. Янковская, Е.Д. Бадаева*

515 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Изучение генетического полиморфизма диплоидной пшеницы *Triticum boeoticum* Boiss. с использованием SSR-маркеров
*М.А. Аббасов*

524 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Сравнительный анализ полных последовательностей пластидных геномов чеснока *Allium sativum* и лука репчатого *Allium cepa*
*М.А. Филюшин, А.М. Мазур, А.В. Щенникова, Е.З. Кочиева*

531 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Генетический полиморфизм тополей Московского региона на основе высокопроизводительного секвенирования ITS-последовательностей
*Е.В. Борхерт, Г.С. Краснов, Н.Л. Большева, П. Кезимана, О.Ю. Юркевич, О.В. Муравенко, А.В. Кудрявцева, Н.В. Мельникова*

## Генофонд и селекция растений

536 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Полногеномный анализ ассоциаций с устойчивостью к грибным болезням в коллекции сои в условиях Юго-Восточного и Южного Казахстана
*А. Затыбеков, С. Абугалиева, С. Дидоренко, А. Рсалиев, Е. Туруспеков*

544 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Аллоплазматические рекомбинантные линии (*H. vulgare*)-*T. aestivum* с транслокацией 1RS.1BL: исходные генотипы для создания сортов мягкой пшеницы
*Л.А. Першина, Л.И. Белова, Н.В. Трубачеева, Т.С. Осадчая, В.К. Шумный, И.А. Белан, Л.П. Россеева, В.В. Немченко, С.Н. Абакумов*

553 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Ресурсный потенциал некоторых видов рода *Miscanthus* Anderss. в условиях континентального климата лесостепи Западной Сибири
*О.В. Дорогина, О.Ю. Васильева, Н.С. Нуждина, Л.В. Буглова, Ю.А. Гисматулина, Е.В. Жмудь, Г.А. Зуева, О.В. Комина, Е.А. Цыбченко*

560 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Тенденция преодоления устойчивости к бурой ржавчине интрогрессивных линий мягкой пшеницы с генетическим материалом *Aegilops speltoides* Tausch
*Л.Я. Плотникова, Л.В. Мешкова, Е.И. Гультяева, О.П. Митрофанова, И.Ф. Лапочкина*

## Генетика животных

568 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
*Drosophila melanogaster*, обитающие на севере европейской части России, заражены *Wolbachia*, негативно влияющей на их продолжительность жизни
*Н.В. Рощина, А.В. Симоненко, А.В. Кременцова, Е.А. Цыбулько, В.Е. Алаторцев, Е.Г. Пасюкова, Д.В. Муха*

574 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Инвазивные виды *Aedes albopictus* и *Aedes aegypti* на Черноморском побережье Краснодарского края: генетика (*COI*, ITS2), зараженность *Wolbachia* и *Dirofilaria*
*Е.В. Шайкевич, И.В. Патраман, А.С. Богачева, В.М. Ракова, О.П. Зеля, Л.А. Ганушкина*

586 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Популяционно-генетическая структура красавки *Anthropoides virgo* в России
*Е.А. Мудрик, Е.И. Ильяшенко, О.А. Горошко, Т.А. Кашенцева, М.В. Корепов, И.А. Сикорский, Г.С. Джамирзоев, В.Ю. Ильяшенко, Д.В. Политов*

**Физиологическая генетика**

593 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Экспрессия пальмитилтрансфераз в структурах мозга мышей с генетически детерминированной предрасположенностью к депрессивно-подобному поведению
*Е.М. Кондаурова, Т.В. Ильчибаева, А.С. Цыбко, Е.Г. Понимаскин, В.С. Науменко*

600 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Влияние однократного введения стрептозотоцина на метаболиты гиппокампа мышей линии NODSCID
*Д.А. Тур, О.Б. Шевелев, М.Б. Шарапова, М.А. Золотых, А.Е. Акулов*

606 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Регуляция ацетилирования гистона H4 в центральной нервной системе и командных нейронах оборонительного поведения моллюска *Helix* серотонином и нейропептидом FMRFамидом
*Л.Н. Гринкевич, Т.Г. Зачепило*

**Генетика человека**

611 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Оценка вклада монгольской экспансии в генофонд тувинцев
*Л.Д. Дамба, Е.В. Балановская, М.К. Жабагин, Ю.М. Юсупов, Ю.В. Богунов, Ж.М. Сабитов, А.Т. Агджоян, Н.А. Короткова, М.Б. Лавряшина, Б.Б. Монгуш, У.Н. Кавай-оол, О.П. Балановский*

620 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Полноэкзомный анализ первичного иммунодефицита
*Э.Ш. Рахмани, Х. Азарпара, М. Каримипур, Х. Рахими*

VAVILOV JOURNAL OF GENETICS AND BREEDING

# CONTENTS • 2018 • 22 • 5

497 FROM THE EDITOR

## Molecular and cell biology

498 REVIEW
Modern approaches to artificial gene synthesis: aspects of oligonucleotide synthesis, enzymatic assembly, sequence verification and error correction
*G.Y. Shevelev, D.V. Pyshnyi*

## Plant genetics

507 ORIGINAL ARTICLE
Alien introgressions and chromosomal rearrangements do not affect the activity of gliadin-coding genes in hybrid lines of *Triticum aestivum* L. × *Aegilops columnaris* Zhuk.
*A.Yu. Novoselskaya-Dragovich, A.A. Yankovskaya, E.D. Badaeva*

515 ORIGINAL ARTICLE
Study of the genetic polymorphism of diploid wheat *Triticum boeoticum* Boiss. using SSR markers
*M.A. Abbasov*

524 ORIGINAL ARTICLE
Comparative analysis of the complete plastomes of garlic *Allium sativum* and bulb onion *Allium cepa*
*M.A. Filyushin, A.M. Mazur, A.V. Shchennikova, E.Z. Kochieva*

531 ORIGINAL ARTICLE
Genetics polymorphism of poplars from Moscow region based on high-throughput sequencing of ITS
*E.V. Borkhert, G.S. Krasnov, N.L. Bolsheva, P. Kezimana, O.Yu. Yurkevich, O.V. Muravenko, A.V. Kudryavtseva, N.V. Melnikova*

## Plant gene pool and breeding

536 ORIGINAL ARTICLE
GWAS of a soybean breeding collection from South East and South Kazakhstan for resistance to fungal diseases
*A. Zatybekov, S. Abugalieva, S. Didorenko, A. Rsaliyev, Y. Turuspekov*

544 ORIGINAL ARTICLE
Alloplasmic recombinant lines (*H. vulgare*)-*T. aestivum* with 1RS.1BL translocation: initial genotypes for production of common wheat varieties
*L.A. Pershina, L.I. Belova, N.V. Trubacheeva, T.S. Osadchaya, V.K. Shumny, I.A. Belan, L.P. Rosseeva, V.V. Nemchenko, S.N. Abakumov*

553 ORIGINAL ARTICLE
Resource potential of some species of the genus *Miscanthus* Anderss. under conditions of continental climate of West Siberian forest-steppe
*O.V. Dorogina, O.Yu. Vasilyeva, N.S. Nuzhdina, L.V. Buglova, Yu.A. Gismatulina, E.V. Zhmud, G.A. Zueva, O.V. Komina, E.A. Tsybchenko*

560 ORIGINAL ARTICLE
A tendency towards leaf rust resistance decrease in common wheat introgression lines with genetic material from *Aegilops speltoides* Tausch
*L.Ya. Plotnikova, L.V. Meshkova, E.I. Gultyaeva, O.P. Mitrofanova, I.F. Lapochkina*

## Animal genetics

568 ORIGINAL ARTICLE
*Drosophila melanogaster* inhabiting northern regions of European Russia are infected with *Wolbachia* which adversely affects their life span
*N.V. Roshina, A.V. Symonenko, A.V. Krementsova, E.A. Tsybul'ko, V.E. Alatortsev, E.G. Pasyukova, D.V. Mukha*

574 ORIGINAL ARTICLE
Invasive mosquito species *Aedes albopictus* and *Aedes aegypti* on the Black Sea coast of the Caucasus: genetics (*COI*, ITS2), *Wolbachia* and *Dirofilaria* infections
*E.V. Shaikevich, I.V. Patraman, A.S. Bogacheva, V.M. Rakova, O.P. Zelya, L.A. Ganushkina*

586 ORIGINAL ARTICLE
The Demoiselle crane (*Anthropoides virgo*) population genetic structure in Russia
*E.A. Mudrik, E.I. Ilyashenko, O.A. Goroshko, T.A. Kashentseva, M.V. Korepov, I.A. Sikorskiy, G.S. Dzhamirzoev, V.Yu. Ilyashenko, D.V. Politov*

## Physiological genetics

593 ORIGINAL ARTICLE
Expression of palmitoyl transferases in brain structures of mice genetically predisposed to depressive-like behavior
*E.M. Kondaurova, T.V. Ilchibaeva, A.S. Tsybko, E.G. Ponimaskin, V.S. Naumenko*

600 ORIGINAL ARTICLE
The effect of a single administration of streptozotocin on hippocampus metabolites in NODSCID mice
*D.A. Tur, O.B. Shevelev, M.B. Sharapova, M.A. Zolotykh, A.E. Akulov*

606 ORIGINAL ARTICLE
Regulation of histone H4 acetylation in the CNS and defensive behavior command neurons of the mollusk *Helix* mediated by serotonin and neuropeptide FMRFamide
*L.N. Grinkevich, T.G. Zachepilo*

## Human genetics

611 ORIGINAL ARTICLE
Estimating the impact of Mongol expansion on gene pool of Tuvans
*L.D. Damba, E.V. Balanovskaya, M.K. Zhabagin, Y.M. Yusupov, Y.V. Bogunov, Z.M. Sabitov, A.T. Agdzhoyan, N.A. Korotkova, M.B. Lavryashina, B.B. Mongush, U.N. Kavai-ool, O.P. Balanovsky*

620 ORIGINAL ARTICLE
Whole exome analysis of primary immunodeficiency
*E.S. Rahmani, H. Azarpara, M. Karimipoor, H. Rahimi*

Уважаемые коллеги, дорогие читатели! Текущий выпуск Вавиловского журнала генетики и селекции открывает раздел «Молекулярная и клеточная биология» с обзорной статьей, в которой обсуждается современное состояние молекулярных и биотехнологических подходов для синтеза и ферментативной сборки генных конструкций.

Традиционными рубриками нашего журнала являются «Генетика растений» и «Генофонд и селекция растений». Раздел «Генетика растений» состоит из четырех экспериментальных статей, в которых описаны результаты применения молекулярно-цитологических методов для изучения генетического разнообразия растений. В первой приведены данные по выявлению чужеродных замещений и транслокаций у гибридных линий пшеницы, полученных с участием дикорастущего вида *Aegilops columnaris*. Преимущество и эффективность микросателлитных маркеров для оценки внутривидового разнообразия и паспортизации *Triticum boeoticum* показано во второй публикации. Еще две статьи посвящены использованию высокопроизводительного секвенирования для сравнительного анализа пластидных геномов чеснока (*Allium sativum*) и лука репчатого (*Allium cepa*) и для характеристики генетического полиморфизма видов и межвидовых гибридов тополя (*Populus*).

Эффективность применения современных геномных технологий для повышения устойчивости сои к болезням и результаты изучения рекомбинантных аллоплазматических линий, полученных скрещиванием пшеницы *T. aestivum* и ячменя *H. vulgare*, по устойчивости к листостеблевым инфекциям и признакам продуктивности обсуждаются в рубрике «Генофонд и селекция растений». В одной из статей этой рубрики рассмотрены тенденция повышения степени восприимчивости к бурой ржавчине культивируемых сортов яровой пшеницы в Западной Сибири и необходимость создания новых форм с генетической устойчивостью. В последней публикации раздела представлены данные оценки ресурсного потенциала комплекса видов *Miscanthus*.

В разделе «Генетика животных» собраны оригинальные экспериментальные работы, выполненные группами исследователей из Института общей генетики им. Н.И. Вавилова. Представлены результаты изучения влияния эндосимбиотической бактерии *Wolbachia* на природные популяции дрозофилы *D. melanogaster* и кровососущих комаров видов *Aedes aegypti* и *Ae. albopictus*. Показана степень инфицированности популяций дрозофилы и комаров симбиотической бактерией, разнообразие штаммов бактерии и ее влияние на продолжительность жизни насекомого-хозяина. В одной из этих работ проведена молекулярно-генетическая идентификация комаров с помощью полиморфных ДНК-маркеров. В последней статье раздела описано генетическое разнообразие широко распространенного вида журавлей *Anthropoides virgo*, или красавки. Генетическая структура популяций данного вида, состоящая из особей пяти гнездовых группировок различных регионов России, была изучена с помощью молекулярных маркеров, специфичных как для ядерного, так и для митохондриального геномов.

Следующая рубрика включает оригинальные исследования в области физиологической генетики. На модели мышей с генетической предрасположенностью к депрессивно-подобному поведению авторы одной из статей попытались установить связь между заболеванием и посттрансляционными модификациями рецепторных мембранных белков, осуществляемых ферментами пальмитилтрансферазами. В другой работе с использованием химически индуцированной модели диабета у линии мышей изучены метаболические изменения, происходящие в гиппокампе под влиянием хронической гипергликемии.

В последний раздел «Генетика человека» вошли две статьи. В одной из них показано использование SNP-маркеров, специфичных для Y-хромосомы человека, для изучения генофонда тувинцев. В другой продемонстрирована эффективность технологий секвенирования экзома для изучения сложных наследственных заболеваний, таких как первичный и комбинированный иммунодефицит.

Обращаем внимание наших читателей, что оригинальные статьи по истории генетики, анонсы и итоги конференций, обзорные и дискуссионные статьи можно опубликовать в электронном издании «Письма в Вавиловский журнал» (http://pismavavilov.ru/). Рукописи принимаются по электронной почте: vavilov_journal@bionet.nsc.ru. Всем статьям присваивается индекс DOI.

*Академик В.К. Шумный*

Вавиловский журнал генетики и селекции. 2018;22(5):498-506
DOI 10.18699/VJ18.387

# Modern approaches to artificial gene synthesis: aspects of oligonucleotide synthesis, enzymatic assembly, sequence verification and error correction

G.Y. Shevelev[1, 2] ✉, D.V. Pyshnyi[1, 2]

[1] Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia

Synthetic biology is a rapidly developing field aimed at engineering of biological systems with predictable properties. Synthetic biology accumulates the achievements of modern biological sciences, programming and computational modeling as well as engineering technologies for creation of biological objects with user-defined properties. Evolution of synthetic biology has been marked by a number of technological developments in each of the mentioned fields. Thus, significant reduction in cost of DNA sequencing has provided an easy access to large amounts of data on the genetic sequences of various organisms, and decreased the price of the DNA sequence synthesis, which, analogous to Moore's law, resulted in an opportunity to create a lot of potential genes without the time – consuming and labor – intensive traditional methods of molecular biology. Development of system biology has allowed forming a deeper understanding of the functions and relationship of natural biological models, as well as of the computational models describing processes at the cell and system levels. Combination of these factors has created an opportunity for conscious changes of natural biological systems. In this review the modern approaches to oligonucleotide gene assembly synthesis are discussed, including such aspects as protocols for gene assembly, sequence verification, error correction and further applications of synthesized genes.

Key words: oligonucleotide synthesis; artificial gene synthesis; enzymatic gene assembly; polymerase cycling assembly (PCA); ligase cycling assembly (LCA); error correction; sequence verification.

**HOW TO CITE THIS ARTICLE:**
Shevelev G.Y., Pyshnyi D.V. Modern approaches to artificial gene synthesis: aspects of oligonucleotide synthesis, enzymatic assembly, sequence verification and error correction. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):498-506. DOI 10.18699/VJ18.387

Received 02.05.2018
Accepted for publication 09.06.2018


✉ e-mail: metatezis@gmail.com

# Современные подходы к синтезу генов: аспекты синтеза олигонуклеотидов, ферментативной сборки, проверки последовательностей и коррекции ошибок

Г.Ю. Шевелев[1, 2] ✉, Д.В. Пышный[1, 2]

[1] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский государственный университет, Новосибирск, Россия

Синтетическая биология – быстро развивающаяся отрасль науки, нацеленная на создание биологических систем с предсказанными свойствами. При этом она использует достижения современной биологии, программирования и компьютерного моделирования, а также инженерной отрасли для создания биологических объектов, обладающих набором заранее заданных пользовательских свойств. Развитие синтетической биологии было обусловлено множеством технологических разработок в каждой из упомянутых отраслей. Так, значительное снижение стоимости технологии секвенирования ДНК привело к наработке больших объемов данных о генетических последовательностях различных организмов. Снижение стоимости синтеза последовательностей ДНК в соответствии с законом Мура позволило создавать библиотеки синтетических генов, представляющие потенциальный интерес в работе генных инженеров без необходимости использования традиционных и трудоемких методов молекулярной биологии. Благодаря развитию системной биологии сформировано глубокое понимание взаимосвязей и функций природных биологических моделей, а также построены прогностические модели, описывающие молекулярные процессы на клеточном и системном уровнях. Комбинация вышеперечисленных факторов создала возможность осознанного изменения природных биологических систем. В данном обзоре обсуждается современное состояние подходов к синтезу олигонуклеотидов для последующей сборки генных конструкций и к ферментативной сборке генов. Освещены аспекты использования различного программного обеспечения для подбора олигонуклеотидов для последующей сборки генов, проверки точности синтезированной последовательности генов, а также исправления ошибок.

Ключевые слова: синтез олигонуклеотидов; синтез искусственных генов; ферментативная сборка генов; полимеразная циклическая сборка; лигазная циклическая сборка; коррекция ошибок; проверка последовательности.

Since a first gene was synthesized from scratch in the 70's (Agarwal et al., 1970) and it's *in vivo* activity was tested (Ryan et al., 1979) around 50 years has passed, and the technology of artificial gene synthesis has made a significant leap beyond the approaches implemented in the classical works. Nowdays, it relies on the modern methods of oligonucleotide synthesis, the variants of Polymerase Chain Reaction (PCR) (Saiki et al., 1985; Mullis et al., 1986) with high fidelity DNA polymerases (Böhlke et al., 2000), and the sequence verification step using the Sanger sequencing (Sanger, Coulson, 1975; Smith et al., 1986) or high-throughput DNA sequencing (Church, 2006; Hall, 2007; Bosch, Grody, 2008; Schuster, 2008; Shendure, Ji, 2008; Tucker et al., 2009). During the last decade modern approaches to assemble large (up to dozens and hundreds kb) DNA molecules (Gibson et al., 2009; Gibson, 2011) and error correction within synthesized gene have been introduced (Ma et al., 2012b).

## Oligonucleotide synthesis

All the synthesis approaches rely on usage of chemically synthesized oligonucleotides as building blocks for an enzyme-mediated assembly (Ma et al., 2012b; Kosuri, Church, 2014).

### Solid-phase oligonucleotide synthesis

The most widely used approach to oligonucleotide synthesis has been the solid-phase phosphoramidite method developed in the 80's (Beaucage, Caruthers, 1981; Matteucci, Caruthers, 1981). The DNA oligonucleotides are synthesized from the 3′ to 5′ end by consecutive coupling of activated building deoxynucleoside phosphoramidites to an initial deoxynucleoside attached to a solid support (usually the support is a controlled pore glass (CPG) or highly cross-linked polystyrene beads) by its 3′-OH group (Ellington, Pollard, 2000). A solid support matrix is placed into several or several hundred parallel flow-through individual reaction columns, and while synthesis all the reagents necessary for the synthesis cycle flow through the solid support column. A full cycle to add a single nucleotide consists of several stages: (1) deprotection: acid is used to remove a Dimetoxytrityl (DMT) group from the 5′-end of a growing oligonucleotide chain; (2) coupling: the solid-phase 5′-OH group of an attached oligonucleotide reacts with tetrazole – activated nucleoside phosphoramidite producing the elongated product of reaction; (3) capping: treatment of the uncoupled 5′-OH groups by acetic anhydride to minimize the products of deletion mutations; (4) oxidation: conversion of an unstable phosphite triester into a stable phosphate prior to the next step of detritylation. The cycle repeats until a full-length oligonucleotide is obtained. The synthesized oligonucleotide is then cleaved from solid support by treatment with a strong base such as ammonia in aqueous solution and the remaining protection groups are removed during the cleavage. For gene synthesis application high quality and homogeneity of synthesized oligonucleotides is critically important (Xiong et al., 2006) since the synthesized oligonucleotides need to be purified either by the high performance liquid chromatography (HPLC) or by polyacrilamide gel electrophoresis (PAGE) methods.

Synthesis of oligonucleotides with the length up to 100 nt using solid-phase phosphoramidite method is a standard for gene synthesis industry and other applications. The limiting factor for overcoming barrier over 100 nt during oligonucleotide synthesis is stochastic depurination side reactions which accompany acid deprotection step (Hall et al., 2009). Vendors such as Integrated DNA Technologies (IDT), however, declare possibility to synthesize high quality oligonucleotides up to 200 nt using a special support and synthetic protocol for low-yield and long-oligonucleotide synthesis.

As an alternative to chemical oligonucleotide synthesis, T4 RNA ligase-mediated solid-phase enzymatic oligonucleotide synthesis approach was proposed (Schmitz, Reetz, 1999). The approach is based on coupling reaction between a solid-phase attached oligonucleotide primer and the mononucleoside 3′,5′-biphosphate mediated by T4 RNA ligase in a water solution. The terminal phosphate blocking group is removed enzymatically by alkaline phosphatase, and all excess reagents are simply washed off from the resin. The next nucleoside 3′,5′-biphosphate can then bind to the end of the chain. Following $NH_2OH$-induced cleavage from the resin gives the desired elongation of the product. The drawback of this approach is relatively low rate of elongation (48 h) due to the kinetic properties of the enzyme. However, possible automation and molecular biological optimization of enzyme using directed evolution can make the enzymatic approach a viable alternative to today's chemical synthesis approach.

### Parallel oligonucleotide synthesis using microchips

The oligonucleotide amounts necessary for gene synthesis application are relatively small (picomols) while modern DNA synthesizers can produce oligonucleotides in a higher excess (nanomolar and micromolar) making this approach redundant. Significant improvements in performance and cost-reduction may be achieved by parallelization and miniaturization of oligonucleotide synthesis platforms. Thus, microchip-based approaches previously used for DNA diagnostics and sequencing have been adopted for multi-parallel synthesis of oligonucleotide arrays. Several constructive implementations such as ink-jet DNA printing (Lausted et al., 2004), microfluidic devices (Zhou et al., 2004; Huang et al., 2009; Lee et al., 2010), light-directed (Richmond et al., 2004) and electrochemical (Egeland, Southern, 2005; Chow et al., 2009) microarray synthesis and LED-controlled capillaries (Blair et al., 2006) synthesis have been independently developed. A conditions modification during the detritylation step of chip-based oligonucleotide synthesis lead to improved quality of oligonucleotides with the possible length up to 150 nt (LeProust et al., 2010). However, the limitation of all microarray synthesis methods is the attomoles ($10^{-18}$ mol) of an individual oligonucleotide that is insufficient for a gene assembly. PCR of selected subpools allows amplifying synthesized oligos at acceptable quantities (Kosuri et al., 2010; Schmidt et al., 2015) for a subsequent multiple gene assembly using barcoded magnetic beads (Plesa et al., 2018).

## Gene assembly

Enzymatic assembly of synthetic oligonucleotides into long double-stranded DNA fragments with the length up to several hundred bp or even several kb (Ma et al., 2012b) is the next step of gene synthesis with numerous approaches developed since this technology was implemented. In this paper we will discuss several protocols for gene assembly and their parameters.

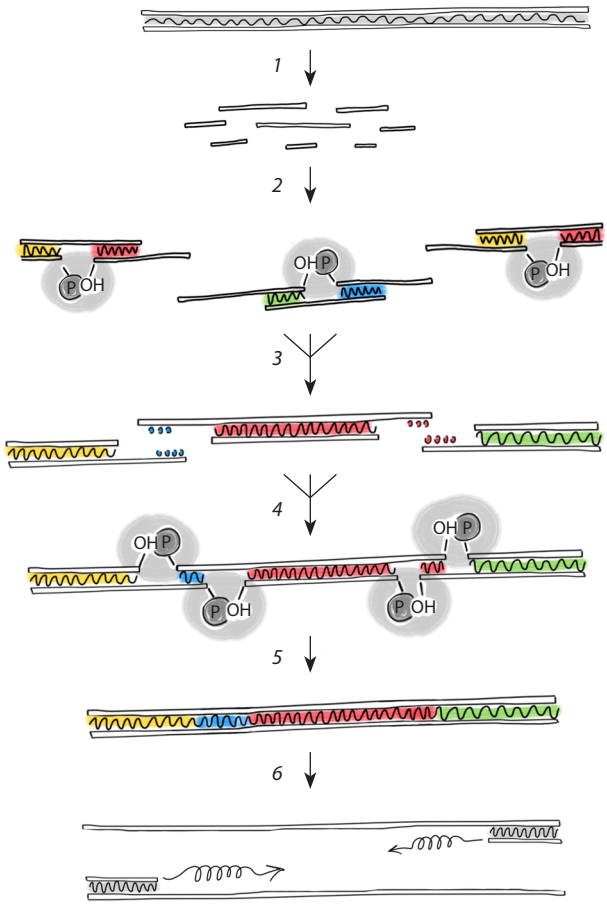



**Fig. 1.** LCA assembly of a DNA construct.

A target sequence is shown on the top of the figure. A set of oligonucleotides synthesized to build the construct (*1*). A set of 5′-phosphorylated oligonucleotides is assembled in stoichiometric ratio and annealed. Ligase seals the nicks (*2*). The double stranded DNA is melted and reannealed with extension products and any remained oligos (*3*) and ligase seals the nicks again (*4*) until full length product appears (5). Amplification of the full-length product using PCR (*6*).

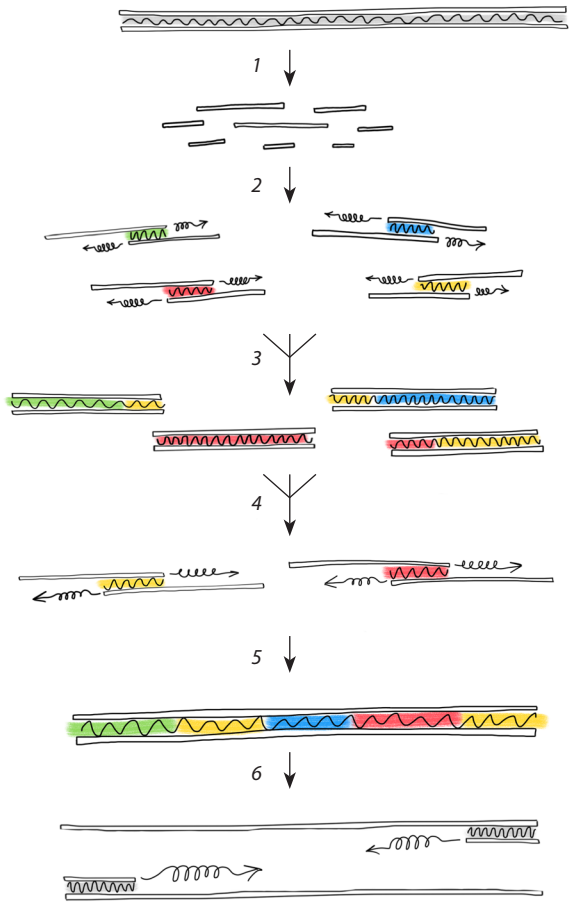



**Fig. 2.** PCA assembly of a DNA construct.

A target sequence is shown on the top of the figure. A set of oligonucleotides is synthesized to build the construct (*1*). A set of oligonucleotides is assembled in stoichiometric ratio and annealed (*2*). Polymerase extends the chain in 3′ direction until the end of the template oligo is reached (*3*). The double stranded DNA is melted and reannealed with extension products and any remained oligos (*4*). Each extension cycle results in a longer products until the full-length product is synthesized (*5*). The full-length product is amplified with PCR using terminal primers (*6*).

### Ligase cycling assembly (LCA)

LCA utilizes the properties of *Taq* DNA ligase to repair single stranded 5′-P, 3′-OH breaks (nicks) within a double stranded DNA at increased temperature (50–60 °C). Conjunction of two oligonucleotides (one 5′-phosphorilated and another with 3′-OH) takes place when they are annealed to a third template oligonucleotide. After joining of oligonucleotides the reaction mixture contains multiple oligonucleotides heated to 95 °C until all the complexes are denaturated and further ligation of longer fragments takes place. These steps (annealing, ligation and denaturation) repeat a number of cycles and after that the reaction mixture is amplified with terminal primers and DNA polymerase until the full-length product is obtained (Fig. 1).

One of the alternative implementation of LCA is blunt-end ligation of double stranded DNA blocks using streptavidin-coated magnetic beads and biotin bonded oligonucleotides joined with subsequently added 5′-phosphorylated DNA duplexes (Dietrich et al., 1998). The latest modifications of this approach utilize 5′-phosphorilated single-stranded oligonucleotides (Pengpumkiat et al., 2016).

### Polymerase cycling assembly (PCA)

PCA is the most widely used method of gene synthesis based on the polymerase chain reaction (PCR) simultaneously carried out in multiple places within the set of short complementary overlapping oligonucleotides forming paired complexes with sticky ends (Fig. 2).

Each step of filling the gaps by polymerase is followed by denaturation and annealing steps and leads to formation of another set of elongated complexes with sticky ends. The process repeats until a full-size product appears (with some side products formed) in the mixture and the final product is amplified with conventional PCR using terminal primers. In the earlier experiments using PCA, in the process of gene amplification with terminal primers gene assembly was performed in a separate tube. Later on this approach was adopted for one-tube assembly (Wu et al., 2006; TerMaat et al., 2009). The desired fragment assembled in a “single-pot” cycled enzymatic reaction or in a several step assembly of separately preassembled structural blocks of the final genetic circuit. Similarity with recursion process gave it its name, so

it is known as recursive PCR. The approach was proposed in 1992 and used for the synthesis of human lysozyme gene (Prodromou, Pearl, 1992). The structure of oligonucleotides overlapping ends was specially designed to obtain the melting temperature within the range 52–56 °C to minimize heterodimers and the oligonucleotides were purified with PAGE. Gapless implementation of recursive PCR was used for 2703-bp plasmid assembly using 134 unpurified 40nt oligonucleotides with 20nt sticky ends (Stemmer et al., 1995).

Since the first usage, PCA approach has been thoroughly investigated and several improvements of the approach have been developed, e. g. primeSTAR HS DNA polymerase was proposed as the best-choice polymerase for efficient high-fidelity gene synthesis (Cherry et al., 2008). Utilization of real-time PCR with intercalating dyes has given the insight information concerning the optimal amount of PCR cycles during gene assembly: the authors declared the amount of PCR cycles of about two times more than the theoretically calculated minimal amount of cycles necessary for assembling a full-length product (Ye et al., 2009). Also, the optimal oligonucleotide concentrations while gene assembly was found to be in a range 10–60 nM, external primer concentration – 0.2–1 μM (Wu et al., 2006) and dNTP increased concentration – up to 1 mM each (Ye et al., 2009).

Modification of PCA approach known as TopDown one-step gene synthesis was proposed as advanced technique (Ye et al., 2009). The essential idea of the approach is in a difference between the $T_m$ values of complexes corresponding to the overlapping sites of oligonucleotides and terminal primers. The outer primers ($T_m$ ~ 50 °C) and inner oligonucleotides ($T_m$ ~ 65 °C) were designed with a melting temperature difference of ~ 15 °C. A higher annealing temperature (65 °C) was used for the first 20 cycles of product assembly and after that the annealing temperature was lowered to 50 °C for the next cycles of full-length product amplification using outer primers. Further improvement of the TopDown synthesis led to the automatic kinetics switch approach (Cheong et al., 2010). The additional feature of this kinetic approach is flanking tail at the outer primer sequence. Primers has two regions with melting temperatures $T_{p1}$ and $T_{p2}$, where $T_{p1}$ is the temperature of a gene-specific region without a flanking tail, and $T_{p2}$ is the melting temperature of a primer including a flanking tail. $T_{p1}$ is lower than the melting temperature of the oligonucleotides assembly and $T_{p2}$ is ~ 72 °C. When the full-length template appears, the outer primers first created full-length DNA with flanked tails, causing the shift in melting temperatures of outer flanked template to $T_{p2}$. The approach provided kinetic switch from inner oligonucleotides assembly to full-length template amplification during one-step gene assembly.

### Thermodynamically balanced inside-out (TBIO) approach

Another PCR-based approach is the TBIO gene synthesis (Gao et al., 2003). The approach utilizes the process of primer dimers extension under PCR conditions in the absence of template. The process of gene synthesis starts in the center of gene sequence between two overlapping oligonucleotides (Fig. 3).

After extension by PCR, the annealing and extension of the next pair of primers between previously synthesized DNA duplex hybridized through specially designed overlapped sequences with adjusted melting temperatures (i. e. thermodynamically balanced) begins. The next iteration with the third pair of primers leads to further elongation of the fragment. The primers in a TBIO assembly are added with gradient concentration: the most internal primers are added in the lowest concentration and the most external primers – in the highest concentration facilitating reaction in the direction inside out. Several gene fragments up to 400–500 bp are independently assembled using a TBIO approach and gel-purified to remove PCR byproducts. Then, these fragments are assembled in one larger gene circuit.

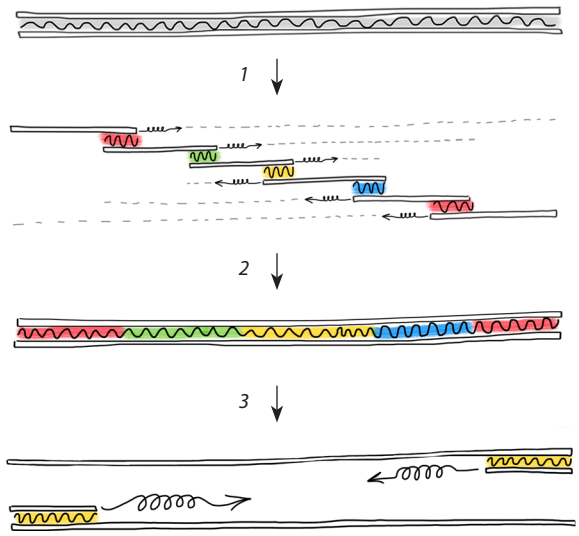



**Fig. 3.** TBIO assembly of a DNA construct.

A target sequence is shown on the top of the figure. A set of synthesized oligonucleotides is mixed in concentration increasing from inside to the outside primers (*1*). Polymerase extends the chain in 3′ direction until the end of the oligo template is reached. The double stranded DNA is melted and reannealed with external primers for further elongation by polymerase. Each extension cycle results in a longer products until the full-length product is synthesized (*2*). The full-length product is amplified with PCR using terminal primers (*3*).

Another strategy for gene synthesis utilizes a combination of Dual Asymmetric (DA) and Overlap Extension PCR (OE-PCR) (Young, Dong, 2004). The oligonucleotides are designed in the same manner as it is done in the PCA approach to contain overlapping regions with or without a gap. Two pairs of complementary oligonucleotides are mixed in one tube and extended by *Pfu* polymerase to double-stranded blunt-ended blocks (Fig. 4).

Several blocks are assembled independently and then mixed with each other to make a whole gene assembly. Reduction of the gap size or its total elimination leads to higher fidelity of the gene assembly during PCR process. To further decrease the mutation product ratio, the full length PCR products are denatured, reannealed and cleaved by T7 endonuclease I. In this manner several genes from 470 bp to 1.2 kb are synthesized. The approach is relatively low-cost because it does not require gel purification of enzymatic phosphorylation.

## Error correction

The source of the errors presented in a synthetic gene can be either the product of oligonucleotide synthesis or of enzymatic gene assembly. Various strategies can be utilized to reduce the error occurrence during different stages of gene synthesis.

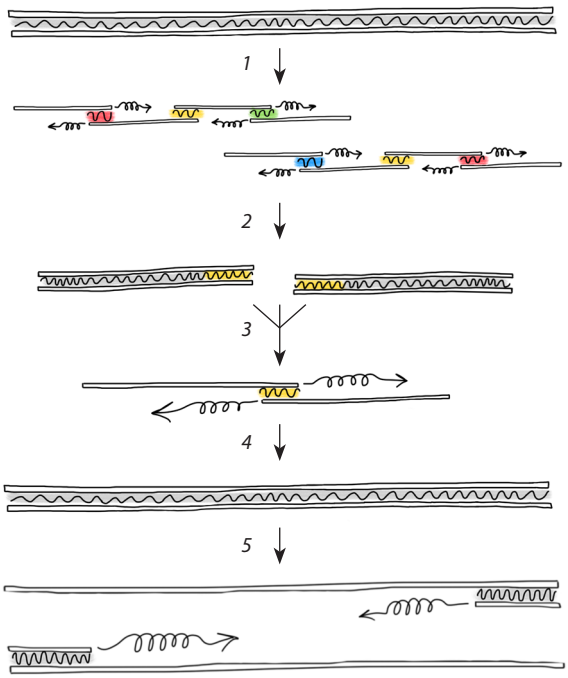



**Fig. 4.** DA OE-PCR assembly of a DNA construct.

A target sequence is shown on the top of the figure. Two pairs from the set of synthesized oligonucleotides are mixed in one tube and extended by polymerase to double-stranded blunt-ended blocks (*1*). Several blocks are assembled independently then mixed with each other to make a whole gene assembly (*2–4*). The full-length product is amplified with PCR using terminal primers (*5*).

### Correction of errors in synthetic oligonucleotides

Since the efficacy of solid-phase oligonucleotide synthesis is less than 100 %, the most common types of errors are deletions and insertions. The deletions occur due to an incomplete capping step of solid-phase oligonucleotide synthesis with the rate up to 0.5 % per nucleotide, while the insertions happen due to tetrazole cleavage of DMT group with the rate below 0.4 % per nucleotide. Another error, depurination (cleavage of the N-glycosidic bound), occurs under acidic conditions during detritylation and causes apurinic sites within an oligonucleotide chain (Ellington, Pollard, 2000). Such modifications of synthesis protocol as increased coupling time of the phosphoramidite, additional methylene chloride wash steps prior and subsequent to deblocking, using of dichloroacetic acid (DCA) instead of trichloroacetic acid (TCA) for deblocking, help to increase the product yield.

Purification of oligonucleotides with HPLC or PAGE helps to minimize the amount of side products containing deletions and insertions (Andrus, Kuimelis, 2001). However, depurination-type errors does not alter the mobility of long oligonucleotide at PAGE and retention time during HPLC purification.

### Correction of errors from synthetic genes

Synthetic oligonucleotides accumulate the errors remained after solid-phase synthesis. Additionally, enzymatic gene assembly may introduce more errors. One of the possible solutions is cloning of a synthesized gene and identifying a correct sequence. If a correct clone cannot be isolated, site-directed mutagenesis needs to be applied to correct all the remaining errors. This step is costly and time-consuming. Another approach is to use DNA mismatch recognition proteins to remove incorrect products from synthetic genes. There are two types of such proteins differ in action: (1) mismatch binding proteins, such as *Taq* MutS and (2) mismatch cleaving proteins, for instance, T7 endonuclease or MutHLS complex (Fig. 5).

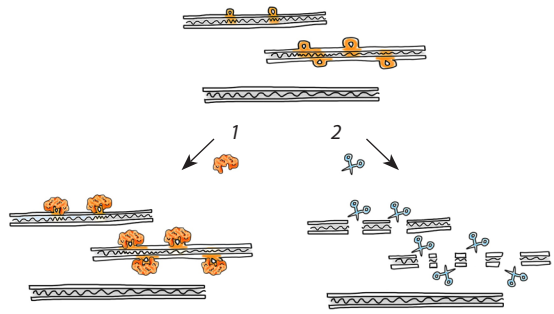



**Fig. 5.** Error correction in synthesized DNA constructs using mismatch recognition proteins.

MutS binds to gene sequences contained mismatches (*1*). The T7 endonuclease cleaves mismatched sequences (*2*).

The MutS protein is a part of bacterial MutHLS DNA repair system (Smith, Modrich, 1997). It detects and binds to a DNA molecule containing single-strand loops and mismatches *in vivo*. In one approach MutS protein isolated from *Thermus aquaticus* was utilized for selection of error-free gene sequences using gel-shift assay (Carr et al., 2004). This method has reduced errors by >15-fold relative to conventional gene synthesis techniques, yielding DNAs with one error per 10 kb. The mismatch cleaving proteins (2) cleave mismatch regions within incorrect DNA complexes. The cleaved complexes can be built into an error-free complex by a polymerase chain assembly (Binkowski et al., 2005) of removed by size selection or exonuclease degradation (Bang, Church, 2008). A combination of mismatch-specific endonucleases such as single-strand-specific nucleases S1 and P1, mung bean nuclease and CEL I nuclease (Desai, Vepatu, 2003), mismatch repair endonuclease MutH (Smith, Modrich, 1997), and resolvases, such as T7 endonuclease I, *E. coli* endonuclease V and T4 endonuclease VII (Ma et al., 2012a) can be used for error removal from synthetic genes. This approach was also applied for error correction in both column and microarray synthesized oligonucleotides using Surveyor nuclease (Saaem et al., 2012; Currin et al., 2014). However, it is unclear which type of mismatch-specific endonucleases is the most effective for error removal (Sequeira et al., 2016).

Another approach for correct gene harvesting is functional selection. The gene of purpose could be fused to a selection marker, such as luciferase gene (Yarimizu et al., 2015). In the case when a marker gene is expressed after cloning, the luciferin-producing clones could be sequenced further. This approach, however, works only for protein-coding genes and is not effective for non-functional mutations.

## Blocks assembly

The typical length of a gene fragment that can be synthesized using the above-mentioned methods is about 1 kb. The

synthesized fragments need to be cloned into a plasmid for sequence verification, the verified fragments combined into a larger gene circuit.

One of the approaches to combine several genetic fragments is the BioBrick strategy allowing assembling a standardized DNA fragment flanked with restriction sites for restriction/ligation reactions (Knight, 2003). Different BioBrick fragments represent various biologically valuable modular parts such as promoters, ribosome-binding sites, coding sequences and transcriptional terminators that can be assembled in a combinatorial way (Shetty et al., 2008). The approach allows assembling several genes and regulatory elements. The method's drawback is the "scars" between modules caused by restriction/ligation assembly.

Utilization of the type II restriction enzymes can be a solution for scarless assembly of gene clusters with the length of about 32 kb (Kodumal et al., 2004). This family of enzymes cut outside of a non-palindromic recognition site has a 4nt overhang, which can be specially chosen for assembling of neighboring fragments without scars (Szybalski et al., 1991). This method is called Golden Gate and has been utilized for assembly of multiple inserts (Engler et al., 2008).

An alternative to restriction/ligation methods is a group of overlap assembly methods (Quan et al., 2009). Similarly to OE-PCR, the overlap methods calls for homologous DNA fragments between blocks to be joined by *in vitro* recombination. There are several commercial cloning kits such as Gateway (Thermo) using proprietary recombinase based on site-specific recombination system of bacteriophage lambda to shuttle sequences between plasmids bearing specific flanking compatible recombination attachment (att) sites (Liang et al., 2013). However, carrying out of a multiple assembly projects with this technique is not convenient because it requires specific sites to be present in each act of assembly. Another kit for cloning into plasmids is In-Fusion (Clontech) which is more promising since it only needs a 15nt sequence overlap between digested plasmid and gene fragment or between gene fragments, which significantly expands the possibility of its application for large circuits assembly (Sleight et al., 2010). Sequence and ligation independent cloning (SLIC) approach utilizes 3′-5′ proofreading exonuclease activity of the T4 DNA polymerase in the absence of dNTP to create overhangs of 30 nt after 30 min incubation with double stranded DNA fragments. DNA overhangs anneal with homologous excised fragments. Recombinase RecA is used to insert the fragment into a plasmid. As a demonstration of applicability of this approach for parallel assembly, 10 fragments with the size ranges between 275 and 980 bp with 40 bp overlaps were inserted into a 3.1 kb vector (Li, Elledge, 2007).

The recombinase-free approach named Gibson Assembly is quite useful for multiple assembly of dsDNA fragments with overlapping ends (Gibson et al., 2009). The approach utilizes a cocktail of three enzymes: (1) exonuclease that chews back the end of the fragments and exposes ssDNA overhangs that are specifically annealed; (2) olymerase that fills the gaps within the annealed products; (3) ligase that covalently sews the fragments together. The authors proposed one-step implementation of this technique using exonuclease III preserving its functional properties in the presence of dNTP, antibody-bound *Taq* polymerase preventing competition with exonuclease for 3′ ends of DNA and *Taq* ligase placed in a single tube in a thermocycler. Three 5-kb DNA fragments with 40 bp overlaps have been assembled and cloned into *E. coli* given 15 kb insert. Moreover, an assembly of 25 % *M. genitalium* genome using two over-100 kb DNA fragments with 257 bp overlap, and its further complete genome has made the Gibson assembly the most universal and convenient method for large scale gene assembly. It has also been adopted for oligonucleotide-scale to gene-scale assembly: the mouse mitochondrial genome was assembled from 60-mer oligonucleotides using this approach (Gibson et al., 2010).

## Oligonucleotides design software

One of the main limitations in gene synthesis is designing a set of oligonucleotides that should meet several criteria: (1) similar values of $T_m$ between overlapping fragments; (2) the length of oligonucleotides within the set should not exceed 50–60 nt; (3) oligonucleotides must be annealed only with their neighborhood oligonucleotides in the set and all abnormal intra and intermolecular annealing must be excluded by a splitting algorithm. Designing oligonucleotides for gene synthesis purposes manually is time-consuming because many factors such as GC content, restriction sites, overlapping fragments, codon frequency, and so on need to be taken into account. Various different software packages have been developed in order to optimize the gene design and synthesis process.

One of the most convenient programs used for these purposes is DNAWorks (https://hpcwebapps.cit.nih.gov/dnaworks/) (Hoover, Lubkowski, 2002). This program allows designing oligonucleotides for gapped and gapless overlap extension PCR assembly as well as TBIO assemblies with specified $T_m$ range and oligonucleotide length. It is also capable to transfer from amino acid to nucleotide sequence using about 10 different codon frequency tables as well as custom user defined table.

Gene2Oligo (http://berry.engin.umich.edu/gene2oligo/) (Rouillard et al., 2004) allows one to design oligonucleotides for LCA and gapless PCA assemblies. GeneDesign (http://54.235.254.95/gd/index.html) (Louw et al., 2011) provides oligos design for gapped PCA assembly.

GeneDesign (http://54.235.254.95/gd/index.html) (Villalobos et al., 2006; Richardson et al., 2010) is a program enabling one to divide long DNA sequence into ~500 bp fragments joined by restriction sites and to chop this fragments into sets of overlapping oligonucleotides. This approach can be successfully applied for very long DNA sequences up to several kb. GeneDesign also provides the "Codon Juggling" feature which is capable to produce sequence that is codon-optimized for expression and, as different as possible from the original sequence (and still coding the same protein). Either, addition and removing of restriction sites within sequence is possible.

TmPrime (http://prime.ibn.a-star.edu.sg) (Bode et al., 2009) and Genecomposer (http://www.genecomposer.net) (Lorimer et al., 2009) are other software solutions for gene design with expanded set of functions analogous to GeneDesign. Currently, not available on the Internet.

Another algorithm based on reliable gradient optimization and the derivative objective function approximated with a central difference operator was proposed for gene synthesis

application by Louw et al. (2011). The authors compared the results of their gradient optimization approach against the one using Gene2Oligo and concluded their approach produces higher purity and yield of assembled genes.

DNASynth (http://dnasynth.sourceforge.net/) (Nowak et al., 2015) is a relatively new algorithm that designs the whole artificial gene synthesis process, developing the optimal nucleotide sequence encoding a given peptide for a given host organism and determining the best long DNA LCA-based assembly protocol.

Microchip-based gene assembly protocols needs special software for oligonucleotide arrays design. PICKY (Birla, Chou, 2015) is software that can be applied for multiple one-pot gene assemblies from dsDNA fragments using the Gibson Assembly protocol. It provides thermodynamic analysis that identifies all unique junctions in gene, where consecutive DNA fragments are specially designed to connect to each other.

## Perspectives and applications

Current applications of synthetic DNA combine synthesis of genetic constructs and metabolic pathways as well as synthesis of artificial genomes and production of artificial genome-based microorganisms (Gibson et al., 2008). Last improvements of long DNA assembly approaches lead to the possibility of creation artificial living system based on *M. genitalium* with minimized functional genome (Hutchison et al., 2016). Recently, the Human Genome Wright project has been announced by a group of US scientists (Boeke et al., 2016). The goal of the project is to synthesize a human genome to improve understanding of the interconnections between genes. For this sake, the additional aim of the project is to catalyze the reduction in price per base, which is now floats around 0.2 $ per 1 bp, to as low as a few cents per bp. Further improvements in microchip-based DNA synthesis technologies will help to drop down the current price per 1 bp.

Another significant application of synthetic biology is metabolic engineering. By designing metabolic systems working in parallel with host metabolic pathways, scientists can program cells for practical applications such as bioremediation (Wasilkowski et al., 2012; Garbisu et al., 2017) and production of industrial compounds (Adrio, Demain, 2010). For example, engineering a pathway for biosynthesis of antimalarial compound Artemisinin in yeast has lowered its price by about 10 times if compared to the previously used method (Paddon, Keasling, 2014).

Application of next generation sequencing technologies (NGS) for synthesis and amplification of error-free oligonucleotides (Schwartz et al., 2012; Lim et al., 2018; Plesa et al., 2018) leads to expanding the application of chip-synthesized oligonucleotide pools into *in vitro* directed evolution methods. For instance, development of multiplex automated genome engineering (MAGE) platform for large-scale programming and evolution of cells and its application to optimization of 1-deoxy-D-xylulose-5-phosphate (DXP) biosynthesis pathway in *E. coli* for overproduction of the industrially important isoprenoid lycopene has allowed for fivefold increase in lycopene amounts which is quite remarkable achievement (Wang et al., 2009). Recent developments of several chip-based oligonucleotide synthesis platforms will enable us to amplify error-free subpools of oligos with further independent assembly of homologous gene libraries allowed authors to rationally explore sequence-function relationships at unprecedented scale (Plesa et al., 2018).

## Acknowledgments

This work has been supported by Russian Science Foundation (no. 17-74-10157).

## Conflict of interest

The authors declare to have no conflict of interest.

## References

Adrio J.-L., Demain A.L. Recombinant organisms for production of industrial products. Bioeng. Bugs. 2010;1(2):116-131. DOI 10.4161/bbug.1.2.10484.

Agarwal K.L., Büchi H., Caruthers M.H., Gupta N., Khorana H.G., Kleppe K., Kumar A., Ohtsuka E., Rajbhandary U.L., Van De Sande J.H., Sgaramella V., Weber H., Yamada T. Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. Nature. 1970;227:27-34. DOI 10.1038/227027a0.

Andrus A., Kuimelis G.R. Analysis and purification of synthetic nucleic acids using HPLC. Curr. Protoc. Nucleic Acid Chem. 2001;1(1):unit 10.5. DOI 10.1002/0471142700.nc1005s01.

Bang D., Church G.M. Gene synthesis by circular assembly amplification. Nat. Methods. 2008;5(1):37-39. DOI 10.1038/nmeth1136.

Beaucage S.L., Caruthers M.H. Deoxynucleoside phosphoramidites – a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. 1981;22(20):1859-1862. DOI 10.1016/S0040-4039(01)90461-7.

Binkowski B.F., Richmond K.E., Kaysen J., Sussman R.M., Belshaw P.J. Correcting errors in synthetic DNA through consensus shuffling. Nucleic Acids Res. 2005;33(6):1-8. DOI 10.1093/nar/gni053.

Birla B.S., Chou H.H. Rational design of high-number dsDNA fragments based on thermodynamics for the construction of full-length genes in a single reaction. PLoS One. 2015;10(12):e0145682. DOI 10.1371/journal.pone.0145682.

Blair S., Richmond K., Rodesch M., Bassetti M., Cerrina F. A scalable method for multiplex LED-controlled synthesis of DNA in capillaries. Nucleic Acids Res. 2006;34(16). DOI 10.1093/nar/gkl641.

Bode M., Khor S., Ye H., Li H.M., Ying J.Y. TmPrime: Fast, flexible oligonucleotide design software for gene synthesis. Nucleic Acids Res. 2009;37(Web Server issue):W214-W221. DOI 10.1093/nar/gkp461.

Boeke J.D., Church G., Hessel A., Kelley N.J., Arkin A., Cai Y., Carlson R., Chakravarti A., Cornish V.W., Holt L., Isaacs F.J., Lajoie M., Lessor T., Lunshof J., Maurano M.T., Mitchell L.A., Rine J., Rosser S., Sanjana N.E., Silver P.A., Valle D., Wang H., Way J.C., Yang L. The Genome Project-Write. Science. 2016;353(6295):126-127. DOI 10.1126/science.aaf6850.

Bosch J.R., Grody W.W. Keeping up with the next generation: massively parallel sequencing in clinical diagnostics. J. Mol. Diagn. 2008;10(6):484-492. DOI 10.2353/jmoldx.2008.080027.

Böhlke K.F., Pisani M., Vorgias C.E., Frey B., Sobek H., Rossi M., Antranikian G. PCR performance of the B-type DNA polymerase from the thermophilic euryarchaeon *Thermococcus aggregans* improved by mutations in the Y-GG/A motif. Nucleic Acids Res. 2000;28(20):3910-3917. DOI 10.1093/nar/28.20.3910.

Carr P.A., Park J.S., Lee Y., Yu T., Zhang S., Jacobson J.M. Protein-mediated error correction for *de novo* DNA synthesis. Nucleic Acids Res. 2004;32(20):e162. DOI 10.1093/nar/gnh160.

Cheong W.C., Lim L.S., Huang M.C., Bode M., Li M.H. New insights into the *de novo* gene synthesis using the automatic kinetics switch approach. Anal. Biochem. 2010;406:51-60. DOI 10.1016/j.ab.2010.06.036.

Cherry J., Nieuwenhuijsen B.W., Kaftan E.J., Kennedy J.D., Chanda P.K. A modified method for PCR-directed gene synthesis from

large number of overlapping oligodeoxyribonucleotides. J. Biochem. Biophys. Methods. 2008;70:820-822. DOI 10.1016/j.jprot.2007.12.009.

Chow B.Y. Emig C.J. Jacobson J.M. Photoelectrochemical synthesis of DNA microarrays. Proc. Natl. Acad. Sci. USA. 2009;106(36):15219-15224. DOI 10.1073/pnas.0813011106.

Church G.M. Genomes for all. Sci. Am. 2006;294(1):47-54.

Currin A., Swainston N., Day P.J., Kell D.B. SpeedyGenes: An improved gene synthesis method for the efficient production of error-corrected, synthetic protein libraries for directed evolution. Protein Eng. Des. Sel. 2014;27(9):273-280. DOI 10.1093/protein/gzu029.

Desai N.A., Vepatu S. Single-strand-specific nucleases. FEMS Microbiol. Rev. 2003;26:457-91.

Dietrich R., Wirsching F., Opitz T., Schwienhorst A. Gene assembly based on blunt-ended double-stranded DNA-modules. Biotechnol. Tech. 1998;12(1):49-54. DOI 10.1023/A:1008855526226.

Egeland R.D., Southern E.M. Electrochemically directed synthesis of oligonucleotides for DNA microarray fabrication. Nucleic Acids Res. 2005;33(14):1-7. DOI 10.1093/nar/gni117.

Ellington A., Pollard J.D., Jr. Introduction to the synthesis and purification of oligonucleotides. Curr. Protoc. Nucleic Acid Chem. 2000; App.3:A.3C.1-A.3C.22. DOI 10.1002/0471142700.nca03cs00.

Engler C., Romy K., Sylvestre M. A one pot, one step, precision cloning method with high throughput capability. PLoS One. 2008;3(11): e3647. DOI 10.1371/journal.pone.0003647.

Gao X., Yo P., Keith A., Ragan T.J., Harris T.K. Thermodynamically balanced inside-out (TBIO) PCR-based gene synthesis: a novel method of primer design for high-fidelity assembly of longer gene sequences. Nucleic Acids Res. 2003;31(22):2-11. DOI 10.1093/nar/gng143.

Garbisu C., Olatz G., Epelde L., Grohmann E., Alkorta I. Plasmid-mediated bioaugmentation for the bioremediation of contaminated soils. Front. Microbiol. 2017;8:1966. DOI 10.3389/fmicb.2017.01966.

Gibson D.G. Enzymatic Assembly of Overlapping DNA Fragments. In: Voigt C. (Ed.). Synthetic Biology. Pt. B: Computer Aided Design and DNA Assembly. (Ser. Methods in Enzymology. Vol. 498). Acad. Press, 2011;349-361. DOI 10.1016/b978-0-12-385120-8.00015-2.

Gibson D.G., Benders G.A., Andrews-Pfannkoch C., Denisova E.A., Baden-Tillson H., Zaveri J., Stockwell T.B., Brownley A., Thomas D.W., Algire M.A., Merryman C., Young L., Noskov V.N., Glass J.I., Venter J.C., Hutchison III C.A., Smith H.O. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science. 2008;319:1215-1220. DOI 10.1126/science.1151721.

Gibson D.G., Smith H.O., Hutchison III C.A., Venter J.C., Merryman C. Chemical synthesis of the mouse mitochondrial genome. Nat. Methods. 2010;7(11):901-903. DOI 10.1038/nmeth.1515.

Gibson D.G., Young L., Chuang R.-Y., Venter J.C., Hutchison C.A., Smith H.O., Hutchison III C.A. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods. 2009;6(5):343-345. DOI 10.1038/nmeth.1318.

Hall B.H., Micheletti J.M., Satya P., Ogle K., Pollard J., Ellington A.D. Design, synthesis, and amplification of DNA pools for *in vitro* selection. Curr. Protoc. Mol. Biol. 2009;88(1):24.2.1-24.2.27. DOI 10.1002/0471142727.mb2402s88.

Hall N. Advanced sequencing technologies and their wider impact in microbiology. J. Exp. Biol. 2007;210:1518-1525. DOI 10.1242/jeb.001370.

Hoover D.M., Lubkowski J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res. 2002;30(10):e43. DOI 10.1093/nar/30.10.e43.

Huang M.C., Ye H., Kuan Y.K., Li M.-H., Ying J.Y. Integrated two-step gene synthesis in a microfluidic device. Lab. Chip. 2009;9:276-285. DOI 10.1039/b807688j.

Hutchison C.A., Chuang R.-Y., Noskov V.N., Nacyra A.-G., Deerinck T.J., Ellisman M.H., Gill J., Kannan K., Karas B.J., Ma L., Pelletier J.F., Qi Z.-Q., Richter R.A., Strychalski E.A., Sun L., Suzuki Y., Tsvetanova B., Wise K.S., Smith H.O., Glass J.I., Merryman C., Gibson D.G., Venter J.C. Design and synthesis of a minimal bacterial genome. Science. 2016;351(6280):aad6253. DOI 10.1126/science.aad6253.

Knight T. Idempotent Vector Design for Standard Assembly of Biobricks. MIT Artificial Intelligence Laboratory, 2003;1-11.

Kodumal S.J., Patel K.G., Reid R., Menzella H.G., Welch M., Santi D.V. Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. Proc. Natl. Acad. Sci. USA. 2004;101(44):15573-15578. DOI 10.1073/pnas.0406911101.

Kosuri S., Church G. Large-scale *de novo* DNA synthesis: technologies and applications. Nat. Methods. 2014;11(5):499-507. DOI 10.1038/nmeth.2918.

Kosuri S., Eroshenko N., Leproust E.M., Super M., Way J., Li B.L., Church G.M. Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. Nat. Biotechnol. 2010; 28(12):1295-1299. DOI 10.1038/nbt.1716.

Lausted C., Dahl T., Warren C., King K., Smith K., Johnson M., Saleem R., Aitchison J., Hood L., Lasky S.R. POSaM: a fast, flexible, open-source, inkjet oligonucleotide synthesizer and microarrayer. Genome Biol. 2004;5(8):R58.1-R58.17. DOI 10.1186/gb-2004-5-8-r58.

Lee C.C., Snyder T.M., Quake S.R. A microfluidic oligonucleotide synthesizer. Nucleic Acids Res. 2010;38(8):2514-2521. DOI 10.1093/nar/gkq092.

LeProust E.M., Peck B.J., Spirin K., McCuen H.B., Moore B., Namsaraev E., Caruthers M.H. Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. Nucleic Acids Res. 2010;38(8):2522-2540. DOI 10.1093/nar/gkq163.

Li M.Z., Elledge S.J. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. Nat. Methods. 2007;4(3):251-256. DOI 10.1038/nmeth1010.

Liang X., Peng L., Baek C.-H., Katzen F. Single step BP/LR combined Gateway reactions. BioTechniques. 2013;55(5):265-268. DOI 10.2144/000114101.

Lim H., Cho N., Ahn J., Park S., Jang H., Kim H., Han H., Ji H.L., Bang D. Highly selective retrieval of accurate DNA utilizing a pool of *in situ*-replicated DNA from multiple next-generation sequencing platforms. Nucleic Acids Res. 2018;46(7):e40. DOI 10.1093/nar/gky016.

Lorimer D., Raymond A., Walchli J., Mixon M., Barrow A., Wallace E., Grice R., Burgin A., Stewart L. Gene composer: database software for protein construct design, codon engineering, and gene synthesis. BMC Biotechnol. 2009;9:36. DOI 10.1186/1472-6750-9-36.

Louw T.M., Whitney S.E., Termaat J.R., Pienaar E., Viljoen H.J. Oligonucleotide optimization for DNA synthesis. AIChE J. 2011;57(7): 1912-1918. DOI 10.1002/aic.12410.

Ma S., Saaem I., Tian J. Error correction in gene synthesis technology. Trends Biotechnol. 2012a;30(3):147-154. DOI 10.1016/j.tibtech.2011.10.002.

Ma S., Tang N., Tian J. DNA synthesis, assembly and applications in synthetic biology. Curr. Opin. Chem. Biol. 2012b;16:260-267. DOI 10.1016/j.cbpa.2012.05.001.

Matteucci M.D., Caruthers M.H. Synthesis of deoxyoligonucleotides on a polymer support. J. Am. Chem. Soc. 1981;103(11):3185-3191. DOI 10.1021/ja00401a041.

Mullis K., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. 1986;51:263-273. DOI 10.1101/SQB.1986.051.01.032.

Nowak R.M., Wojtowicz-Krawiec A., Pluciennicz A. DNASynth: A computer program for assembly of artificial gene parts in decreasing temperature. BioMed. Res. Int. 2015;Article ID:413262. DOI 10.1155/2015/413262.

Paddon C.J., Keasling J.D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nat. Rev. Microbiol. 2014;12:355-367. DOI 10.1038/nrmicro3240.

Pengpumkiat S., Koesdjojo M., Rowley E.R., Mockler T.C., Remcho V.T. Rapid synthesis of a long double-stranded oligonucleotide from a single-stranded nucleotide using magnetic beads and an oligo library. PLoS One. 2016;11(3):e0149774. DOI 10.1371/journal.pone.0149774.

Plesa C., Sidore A.M., Lubock N.B., Zhang D., Kosuri S. Multiplexed gene synthesis in emulsions for exploring protein functional landscapes. Science. 2018;369(6373):343-347. DOI 10.1126/science.aao5167.

Prodromou C., Pearl L.H. Recursive PCR: A novel technique for total gene synthesis. Protein Eng. Des. Sel. 1992;5(8):827-829. DOI 10.1093/protein/5.8.827.

Quan J., Tian J. Circular polymerase extension cloning of complex gene libraries and pathways. PLoS One. 2009;4(7):e6441. DOI 10.1371/journal.pone.0006441.

Richardson S.M., Nunley P., Yarrington R.M., Boeke J.D., Bader J.S. GeneDesign 3.0 is an updated synthetic biology toolkit. Nucleic Acids Res. 2010;38(8):2603-2606. DOI 10.1093/nar/gkq143.

Richmond K.E., Li M.H., Rodesche M.J., Patel M., Lowe A.M., Kim C., Chu L.L., Venkataramaian N., Flickinger S.F., Kaysen J., Belshaw P.J., Sussman M.R., Cerrina F. Amplification and assembly of chip-eluted DNA (AACED): A method for high-throughput gene synthesis. Nucleic Acids Res. 2004;32(17):5011-5018. DOI 10.1093/nar/gkh793.

Rouillard J., Lee W., Truan G., Gao X., Zhou X., Gulari E. Gene2Oligo: Oligonucleotide design for *in vitro* gene synthesis. Nucleic Acids Res. 2004;32(Web Server issue):W176-W180. DOI 10.1093/nar/gkh401.

Ryan J., Brown E.L., Sekiya T., Kiipper H., Khorana H.G. Total synthesis of a tyrosine suppressor tRNA gene. XVIII. Biological activity and transcription, *in vitro*, of the cloned gene. J. Biol. Chem. 1979; 254(13):5817-5826.

Saaem I., Ma S., Quan J., Tian J. Error correction of microchip synthesized genes using Surveyor nuclease. Nucleic Acids Res. 2012; 40(3):e23. DOI 10.1093/nar/gkr887.

Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 1985;230:1350-1354. DOI 10.1126/science.2999980.

Sanger F., Coulson A.R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. 1975; 94:441-448. DOI 10.1016/0022-2836(75)90213-2.

Schmidt T.L., Beliveau B.J., Uca Y.O., Theilmann M., Cruz F.D., Wu C.T., Shih W.M. Scalable amplification of strand subsets from chip-synthesized oligonucleotide libraries. Nat. Commun. 2015;6: 8634. DOI 10.1038/ncomms9634.

Schmitz C., Reetz M.T. Solid-phase enzymatic synthesis of oligonucleotides. Org. Lett. 1999;1(11):1729-1731.

Schuster S.C. Next-generation sequencing transforms today's biology. Nat. Methods. 2008;5(1):16-18. DOI 10.1038/nmeth1156.

Schwartz J.J., Lee C., Shendure J. Accurate gene synthesis with tag-directed retrieval of sequence-verified DNA molecules. Nat. Methods. 2012;9(9):913-915. DOI 10.1038/nmeth.2137.

Sequeira A.F., Guerreiro C.I., Vincentelli R., Fontes C.M. T7 endonuclease I mediates error correction in artificial gene synthesis. Mol. Biotechnol. 2016;58(8-9):573-584. DOI 10.1007/s12033-016-9957-7.

Shendure J., Ji H. Next generation DNA sequencing. Eng. Life Sci. 2008;26(10):1135-1345. DOI 10.1002/elsc.201600121.

Shetty R.P., Drew E., Knight T.F. Engineering BioBrick vectors from BioBrick parts. J. Biol. Eng. 2008;2:5. DOI 10.1186/1754-1611-2-5.

Sleight S.C., Bartley B.A., Lieviant J.A., Sauro H.M. In-Fusion BioBrick assembly and re-engineering. Nucleic Acids Res. 2010;38(8): 2624-2636. DOI 10.1093/nar/gkq179.

Smith J., Modrich P. Removal of polymerase-produced mutant sequences from PCR products. Proc. Natl. Acad. Sci. USA. 1997;94: 6847-6850.

Smith L.M., Sanders J.Z., Kaiser R.J., Hughes P., Dodd C., Connell C.R., Heiner C., Kent S.B.H., Hood L.E. Fluorescence detection in automated DNA sequence analysis. Nature. 1986;321:674-679. DOI 10.1038/321674a0.

Stemmer W.P.C., Crameri A., Ha K.D., Brennan T.M., Heyneker H.L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene. 1995;164:49-53. DOI 10.1016/0378-1119(95)00511-4.

Szybalski W., Kim S.C., Hasan N., Podhajska A.J. Class-IIS restriction enzymes – a review. Gene. 1991;100:13-26. DOI 10.1016/0378-1119(91)90345-C.

TerMaat J.R., Pienaar E., Whitney S.E., Mamedov T.G., Subramanian A. Gene synthesis by integrated polymerase chain assembly and PCR amplification using a high-speed thermocycler. J. Microbiol. Meth. 2009;79(3):295-300. DOI 10.1016/j.mimet.2009.09.015.

Tucker T., Marra M., Friedman J.M. Massively parallel sequencing: the next big thing in genetic medicine. Am. J. Hum. Genet. 2009;85:142-154. DOI 10.1016/j.ajhg.2009.06.022.

Villalobos A., Ness J.E., Gustafsson C., Minshull J., Govindarajan S. Gene Designer: A synthetic biology tool for constructuring artificial DNA segments. BMC Bioinformatics. 2006;7:285. DOI 10.1186/1471-2105-7-285.

Wang H.H., Isaacs F.J., Carr P.A., Sun Z.Z., Xu G., Forest C.R., Church G.M. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009;460:894-898. DOI 10.1038/nature08187.

Wasilkowski D., Swedziol Z., Mrozik A. The applicability of genetically modified microorganisms in bioremediation of contaminated environments. Chemik. 2012;8(66):817-826.

Wu G., Wolf J.B., Ibrahim A.F., Vadasz S., Gunasinghe M., Freeland S.J. Simplified gene synthesis: A one-step approach to PCR-based gene construction. J. Biotechnol. 2006;124:496-503. DOI 10.1016/j.jbiotec.2006.01.015.

Xiong A.-S., Yao Q.-H., Peng R.-H., Duan H., Li X., Fan H.-Q., Cheng Z.-M., Li Y. PCR-based accurate synthesis of long DNA sequences. Nat. Protoc. 2006;1(2):791-797. DOI 10.1038/nprot.2006.103.

Yarimizu T., Nakamura M., Hoshida H., Akada R. Screening of accurate clones for gene synthesis in yeast. J. Biosci. Bioeng. 2015; 119(3):251-259. DOI 10.1016/j.jbiosc.2014.08.006.

Ye H., Huang M.C., Li M.H., Ying J.Y. Experimental analysis of gene assembly with TopDown one-step real-time gene synthesis. Nucleic Acids Res. 2009;37(7):e51. DOI 10.1093/nar/gkp118.

Young L., Dong Q. Two-step total gene synthesis method. Nucleic Acids Res. 2004;32(7):e59. DOI 10.1093/nar/gnh058.

Zhou X., Cai S., Hong A., You Q., Yu P., Sheng N., Srivannavit O., Muranjan S., Rouillard J.M., Xia Y., Zhang X., Xiang Q., Ganesh R., Zhu Q., Matejko A., Gulari E., Gao X. Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences. Nucleic Acids Res. 2004;2(18):5409-5417. DOI 10.1093/nar/gkh879.

Вавиловский журнал генетики и селекции. 2018;22(5):507-514
DOI 10.18699/VJ18.388



# Alien introgressions and chromosomal rearrangements do not affect the activity of gliadin-coding genes in hybrid lines of *Triticum aestivum* L. × *Aegilops columnaris* Zhuk.

A.Yu. Novoselskaya-Dragovich ✉, A.A. Yankovskaya, E.D. Badaeva

Vavilov Institute of General Genetics, RAS, Moscow, Russia

Using chromosome C-banding and electrophoresis of grain storage proteins, gliadins, 17 *Triticum aestivum-Aegilops columnaris* lines with substitutions of chromosomes of homoeologous groups 1 and 6 were examined. Based on their high polymorphism, gliadins were used to identify alien genetic material. For all of the lines examined, electrophoretic analysis of gliadin spectra confirmed substitution of wheat chromosomes 6A, 6D or 1D for the homoeologous *Aegilops* chromosomes of genomes $U^c$ or $X^c$. The substitution manifested in the disappearance of the products of gliadin-coding genes on chromosomes 6A, 6D or 1D with the simultaneous appearance of the products of genes localized on alien chromosomes of genomes $U^c$ or $X^c$. Thus, *Aegilops* chromosomes were shown to be functionally active in the alien wheat genome. The absence of alien genes expression in the lines carrying a long arm deletion in chromosome $6X^c$ suggested that the gliadin-coding locus moved from the short chromosome arm (its characteristic position in all known wheat species) to the long one. This is probably associated with a large species-specific pericentric inversion. In spite of losing a part of its long arm and combination with a non-homologous chromosome of a different genome (4BL), chromosome 1D was fully functioning. For *Aegilops*, the block type of gliadin components inheritance was shown, indicating similarity in the structural organization of gliadin-coding loci in these genera. Based on determining genetic control of various polypeptides in the electrophoretic aegilops spectrum, markers to identify *Ae. columnaris* chromosomes $1X^c$, $6X^c$ and $6U^c$ were constructed.

Key words: electrophoresis of gliadin seed storage proteins; gliadins; C-banding; substitution; translocation; *Aegilops columnaris*; wheat; introgression.

## Интрогрессии и хромосомные перестройки не влияют на активность глиадинкодирующих генов в линиях гибридов *Triticum aestivum* L. × *Aegilops columnaris* Zhuk.

А.Ю. Новосельская-Драгович ✉, А.А. Янковская, Е.Д. Бадаева

Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия

Вид дикорастущей пшеницы *Aegilops columnaris* Zhuk. представляет собой потенциальный источник новых генов, важных для улучшения хозяйственно ценных признаков пшеницы. До настоящего времени он не использовался в селекционных программах. В работе методами С-окрашивания хромосом и электрофореза запасных белков зерна мягкой пшеницы – глиадинов – проанализированы 17 линий *Triticum aestivum* L. × *Aegilops columnaris* Zhuk. с замещениями по хромосомам 1-й и 6-й гомеологических групп. Глиадин за счет высокого полиморфизма позволил идентифицировать чужеродный генетический материал. Для всех исследованных линий анализ электрофоретических спектров глиадина подтвердил замещение хромосом 6A, 6D или 1D мягкой пшеницы на гомеологические хромосомы эгилопса *Ae. columnaris* Zhuk., относящиеся к $U^c$ или $X^c$-геномам. Замещение проявлялось в исчезновении продуктов экспрессии глиадинкодирующих генов на хромосомах 6A, 6D или 1D с одновременным появлением продуктов экспрессии генов, локализованных на чужеродных для пшеницы хромосомах $U^c$ или $X^c$-геномов. Таким образом, показана функциональная активность эгилопсных хромосом в чужеродном для них пшеничном геноме. Отсутствие экспрессии чужеродных глиадинкодирующих генов у линий с делецией длинного плеча хромосомы $6X^c$ позволило выдвинуть гипотезу о перемещении глиадинкодирующего локуса из короткого плеча (что характерно для всех известных видов пшеницы) в длинное. Перемещение глиадинкодирующего локуса, вероятно, связано с крупной видоспецифической перицентрической инверсией – цитогенетический анализ показал существенные различия ортологичных хромосом 6-й группы X-генома по морфологии. В то же время хромосома 1D, независимо от потери части длинного плеча и объединения с негомологичной хромосомой другого генома (4BL), полноценно функционирует. Для эгилопса показан «блочный» характер наследования компонентов глиадина, что свидетельствует о сходстве

Received 30.03.2018
Accepted for publication 30.05.2018


✉ e-mail: dragovich@vigg.ru

организационной структуры глиадинкодирующих локусов у представителей этих родов. Определение генетического контроля разных полипептидов электрофоретического спектра эгилопса позволило разработать маркеры для идентификации хромосом 1X$^c$, 6X$^c$ и 6U$^c$ *Ae. columnaris.*

Ключевые слова: электрофорез запасных белков; глиадины; дифференциальное окрашивание; замещения; транслокации; *Aegilops columnaris*; пшеница; интрогрессия.

**HOW TO CITE THIS ARTICLE:**
Novoselskaya-Dragovich A.Yu., Yankovskaya A.A., Badaeva E.D. Alien introgressions and chromosomal rearrangements do not affect the activity of gliadin-coding genes in hybrid lines of *Triticum aestivum* L. × *Aegilops columnaris* Zhuk. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):507-514. DOI 10.18699/VJ18.388

To improve common wheat, allied wild species, particularly those representing *Aegilops* L. genus, are used as donors of agriculturally valuable traits. A significant number of genes, especially the ones responsible for resistance to diseases, pests, drought, and other stress factors, were introduced into the common wheat genome from various species of this genus (Damania et al., 1992; Friebe et al., 1996; Monneveux et al., 2000; Schneider et al., 2008; McIntosh et al., 2013; Molnár-Láng et al., 2014, 2015). Moreover, various *Aegilops* species are used to improve grain quality of wheat cultivars (Garg et al., 2008, 2016; Tiwari et al., 2010; Rawat et al., 2011; Rakszegi et al., 2017).

In recent decades a full or nearly full spectrum of *Triticum-Aegilops* addition or substitution lines has been derived for more than ten *Aegilops* species (see surveys in Schneider et al., 2008; Molnár-Láng et al., 2015). However, it was not until recently that lines of wheat with introgression from *Ae. columnaris* Zhuk. have been successfully developed. *Aegilops columnaris* is a tetraploid species with U$^c$X$^c$ genomic formula (Dvořák, 1998; Badaeva et al., 2004), which possesses a number of valuable breeding traits, particularly drought tolerance and resistance to diseases and pests (Gill et al., 1985; Warham et al., 1986; Damania et al., 1992). A set of *T. aestivum* × *Ae. columnaris* lines has been derived by researchers at Genetics and Cytology Laboratory at Agricultural Research Institute for South-East Region (Saratov, Russia). Earlier, analysis of 84 lines using differential chromosome banding showed that hybrids differ in introgression spectra (Shishkina et al., 2017; Badaeva et al., 2018). The material studied included lines with additions and/or substitutions of chromosomes of homoeologous groups 1 and 6.

It was shown that genes coding grain storage proteins, gliadins, in grasses are localized in short arms of chromosomes of homoeologous groups 1 and 6 (Shepherd, 1968; Kasarda et al., 1976; Singh, Shepherd, 1988). Gliadins visualized by PAGE electrophoresis are highly polymorphic proteins controlled by multiple alleles of six non-linked gliadin-coding (GC) loci (*Gli-A1*, *Gli-B1*, *Gli-D1*, *Gli-A2*, *Gli-B2*, *Gli-D2*). As each GC locus is a cluster of tightly linked genes, the polypeptides controlled by these clusters appear on electrophoregrams as blocks of components inherited as a single Mendelian character. Multiple allelism in GC loci permits not only reliable identification of wheat cultivars, but also determination of their internal structure, i. e. biotype/haplotype composition. Thus, EP analysis of grain storage proteins is widely used in genetic research of wheat for cultivar certification purposes, identification of population structure, biodiversity investigation, gene mapping, and solving many other problems.

Gliadin genetics is studied quite thoroughly, and GC loci alleles are identified for many domestic and foreign cultivars (see the survey in Novoselskaya-Dragovich, 2015). However, genetic control of gliadins in *Aegilops* species, and *Ae. columnaris* in particular, is scarcely studied. Analysis of *Triticum-Ae. columnaris* lines with chromosome introgression of groups 1 and 6 permits not only to identify GC locus alleles controlled by *Aegilops* chromosomes, but also to develop identification markers for *Ae. columnaris* genetic material and evaluate effects of *Aegilops* chromosome introgression on gene expression in wheat.

Thus, the objective of our study was clarify the genome structure in introgressive lines and analyze alien gene expression in wheat genome by EP analysis of storage proteins.

## Materials and methods

In the present study, 17 *Triticum-Ae. columnaris* introgressive lines with chromosome substitution in homoeologous groups 1 and 6 were examined (see the Table). The Dobrynya and L-503 wheat cultivars ($2n = 42$) with T7DL-7Ai Triticum-Agropyrim translocation, as well as the Saratovskaya 68 cultivar with a standard set of wheat chromosomes, were taken as female parents (Badaeva et al., 2018). K-1193 *Ae. columnaris* line ($2n = 28$) acted as a male parent.

Gliadin extraction and electrophoresis in polyacrylamide gel were performed using standard techniques (Upelniek et al., 2013). GC locus alleles were classified according to the catalog (Metakovsky, Novoselskaya, 1991).

Differential chromosome banding (C-banding) was performed as in Badaeva et al., 1994. Wheat chromosomes were classified based on the genetic nomenclature (Gill et al., 1991); to classify *Ae. columnaris* chromosomes, the nomenclature developed at our laboratory was used (Badaeva et al., 2018).

## Results

Using gliadin EP spectra analysis, we identified GC locus alleles in all common wheat introgressive lines and parental cultivars (see the Table). The hybrids were shown to inherite alleles only from one of two previously established haplotypes of Dobrynya and L-503 cultivars (Novoselskaya-Dragovich et al., 2003) (see the Table). EP spectra of the lines match those of the parental cultivars. For instance, line1794/1 inherited *Gli-A1m* and *Gli-B2d* from the Saratovskaya 68 cultivar and *Gli-B1e*, *Gli-D1a*, *Gli-A2s* – from Dobrynya. *Gli-A1f* allele uncharacteristic for the parental wheat cultivar was found only in line 2304/1. Probably, it was introduced as a result of random cross-pollination. All the *Ae. columnaris*

Cytogenetic and genetic characteristics of *T. aestivum – Ae. columanris* introgressive lines

| № | Cultivar/Line | 2*n* | Cytological characteristics of lines | Genetic characteristics of lines | Gliadin-coding locus alleles *Gli-* | | | | | | Biotypes/ Chromosome substitutions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | *A1* | *B1* | *D1* | *A2* | *B2* | *D2* | |
| 1 | Dobrynya | 42 | T7DL-7Ai | Cultivar consists of two biotypes, *Gli-B2q/Gli-B2s* | *i* | *e* | *a* | *s* | ***q*** | *e* | Biotype 1 |
| | | | | | *i* | *e* | *a* | *s* | ***s*** | *e* | Biotype 2 |
| 2 | L-503 | 42 | T7DL-7Ai | Cultivar consists of two biotype, *Gli-A1i/Gli-A1m* | ***i*** | *e* | *a* | *q* | *qx* | *e* | Biotype 1 |
| | | | | | ***m*** | *e* | *a* | *q* | *qx* | *e* | Biotype 2 |
| 3 | Saratovskaya 68 | 42 | – | Monomorphic cultivar | *m* | *e* | *f* | *o* | *d* | *e* | |
| 4 | 1721/1 | 40 | Disomic substitution of 6A(6U$^c$). Terminal deletion of 1BL. 6A and 1D absent; disomic T4BL-1D and T7DL-7Ai | 1D functional; 6A absent, 6U$^c$ expressed | *i* | ***e*** | *a* | **6U$^c$** | *s* | *e* | **6A(6U$^c$)** |
| 5 | 1776/1 | 40 | id. | id. | *i* | *e* | *a* | **6U$^c$** | *s* | *e* | **6A(6U$^c$)** |
| 6 | 1777/4 | 40 | » | » | *i* | *e* | *a* | **6U$^c$** | ***s*** | ***e*** | **6A(6U$^c$)** |
| 7 | 2307/1W | 40 | » | » | *i* | *e* | *a* | **6U$^c$** | *s* | *e* | **6A(6U$^c$)** |
| 8 | 2307/1w | 43 | Disomic addition of 6U$^c$; 4B+1D/ T4BL-1D or 4B(T4B-1D). T7DL-7Ai | » | *i* | *e* | *a* | *s*+**6U$^c$** | ***s*** | ***e*** | **6A+6U$^c$** |
| 9 | 2308/5 | 40 | » | 1D functional; 6A absent, 6U$^c$ expressed (in one dose) | *i* | *e* | *a* | **6U$^c$** | ***s*** | ***e*** | **6A(6U$^c$)** |
| 10 | 1777/1 | 42 | Double disomic substitution of 5D(5X$^c$) 6A(6X$^c$). Terminal deletion of 1BL; splitting by chromosomes 7B and 7D/ T7DL-7A | 6A absent, 6U$^c$ expressed | *m* | *e* | *a* | *s* | *s* | **6X$^c$** | **6A(6X$^c$)** |
| 11 | 1813 | 42 | Disomic substitution of 6D (6U$^c$); T7DL-7Ai | 6D not expressed, 6U$^c$ components appear | *i* | *e* | *a* | *q* | *s* | **6U$^c$** | **6D(6U$^c$)** |
| 12 | 2054/3 | 42 | id. | 6D not expressed 6U$^c$ components appear | *i* | *e* | *a* | *q* | *s* | **6U$^c$** | **6D(6U$^c$)** |
| 13 | 2015/2 | 42 | Monosomic substitution of 6D/ 6U$^c$. T7DL-7Ai | 6D not expressed, 6U$^c$ components appear | *i* | *e* | *a* | *q* | *q* | **6U$^c$** | **6D(6U$^c$)** |
| 14 | 2306/3 | 42 | id. | id. | *i* | *e* | *a* | *q* | *q*+*s* | **6U$^c$** | **6D(6U$^c$)** |
| 15 | 2310/1 | 42 | Monosomic substitution of 6D/ 6U$^c$, 1DL with terminal translocation. T7DL-7Ai | » | *i* | *e* | *a* | *q* | *q* | **6U$^c$** | **6D(6U$^c$)** |
| 16 | 2304/1 | 42 | Double disomic substitution of 5D(5X$^c$) 6B(6X$^c$). T7DL-7Ai | 6D not expressed, 6X$^c$ expressed | *f* | *e* | *a* | *q* | *q* | **6X$^c$** | **6D(6X$^c$)** |
| 17 | 2033/1 | 40 | Hybrid; splitting of 6D/ 6Ai-2 + + 5D/5X+?tel6X$^c$S + 3DL with terminal translocation, 6D absent | id. | *i* | *e* | *a* | *s* | *s* | – | – |
| 18 | 2034/3-1 | 42 | 6D absent. 6X$^c$ with large terminal long arm deletion | » | *m*+*i* (4:1) | *e* | *a* | *s* | *s* | **6X$^c$** | **6D(6X$^c$)** |

**End of table**

| № | Cultivar/ Line | 2*n* | Cytological characteristics of lines | Genetic characteristics of lines | Gliadin-coding locus alleles *Gli-* | | | | | | Biotypes/ Chromosome substitutions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | *A1* | *B1* | *D1* | *A2* | *B2* | *D2* | |
| 19 | 1930 | 42 | Disomic substitution of 1D(1X$^c$). Nulli-6A, tetra-6D; one 6D pair with terminal long arm translocation. T7DL-7Ai | 1D and 6A not expressed, 1X$^c$ expressed, tetra-6D | *i* | *e* | **1X$^c$** | – | *q* | *e* | **1D(1X$^c$)** |
| 20 | 1794/1 | 42 | 6D substituted T7U$^c$S:6X$^c$S; T7DL-7Ai | 6D and 6X$^c$ not expressed | *m* | *e* | *a* | *s* | *s* | – | – |

Notes: 2*n* –chromosomes number, Gliadin-coding loci *Gli-A1*, *Gli-B1*, *Gli-D1*, *Gli-A2*, *Gli-B2*, *Gli-D2* controlled by chromosomes 1A, 1B, 1D, 6A, 6B и 6D, respectively.

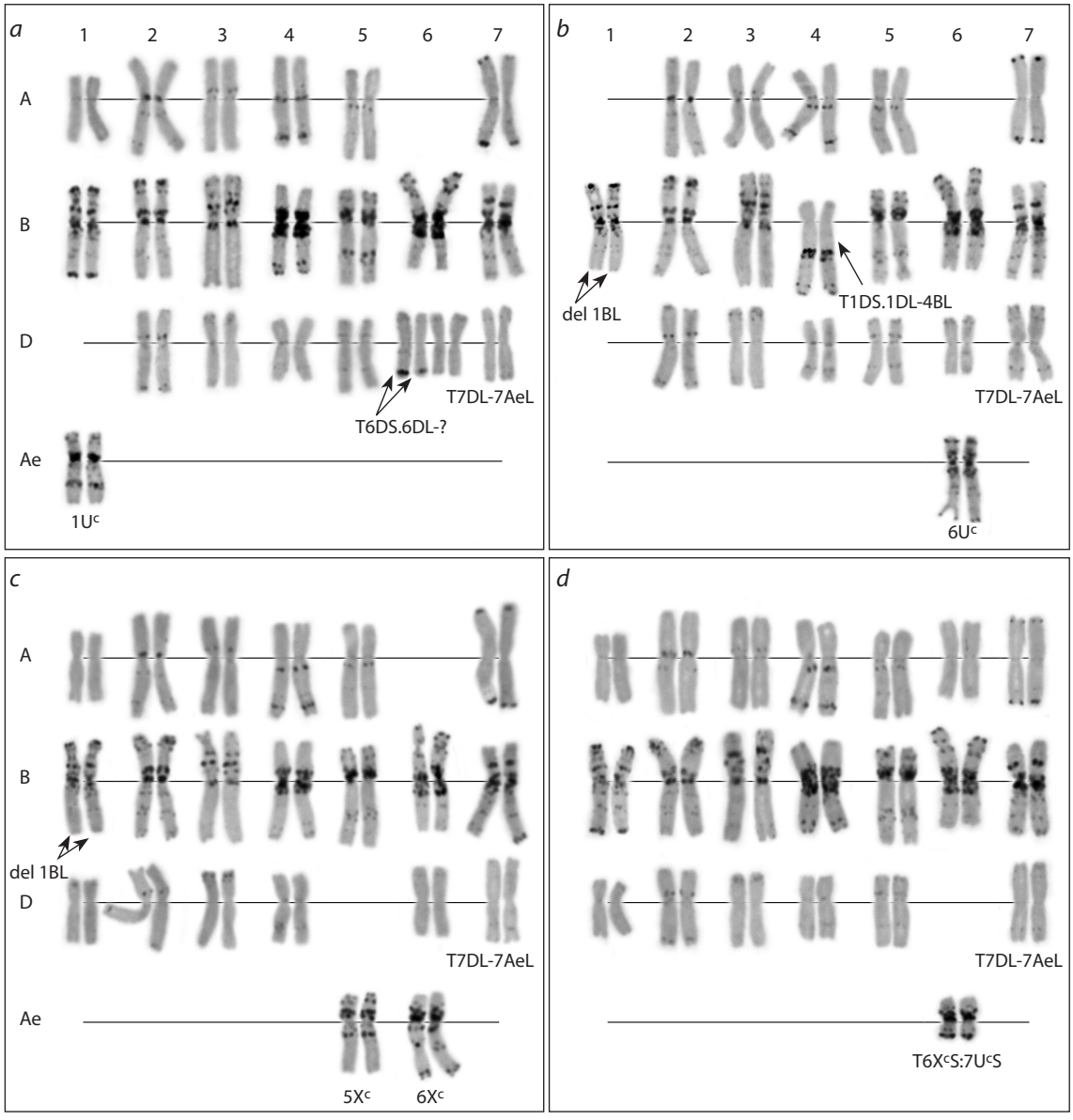



**Fig. 1.** Karyotypes of introgressive lines *T. aestivum* × *Ae. columnaris*, K-1193. *a* – 1930; *b* – 2308; *c* – 1771/1; *d* – 1794/1.

A, B, D – Common wheat genomes, Ae – genome of *Ae. columnaris*, 1–7 – homoeologous groups. Numbers of alien chromosomes are shown at the bottom of the figure. Chromosomes rearrangements are indicated with arrows.

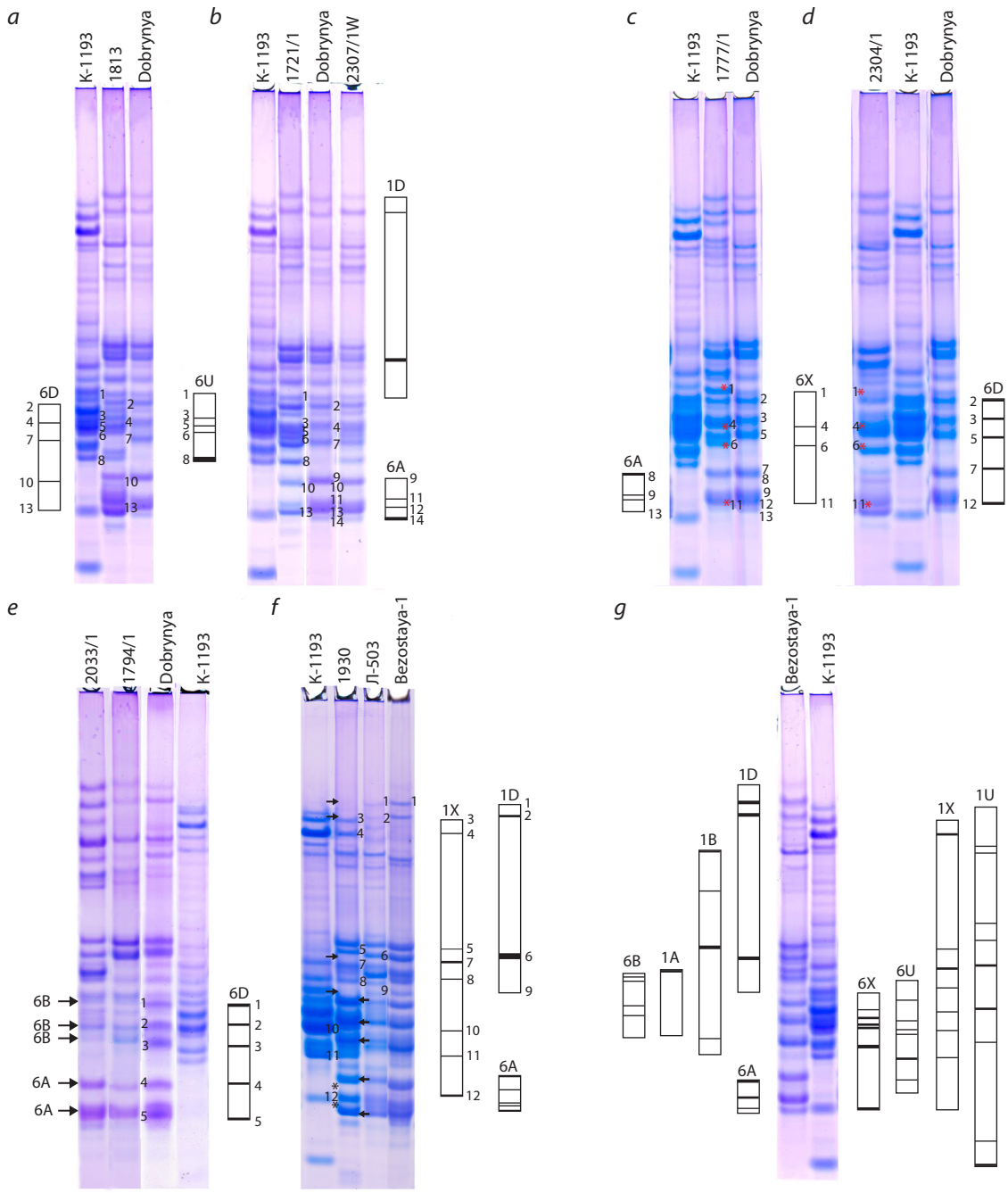



**Fig. 2.** Gliadin electrophoretic spectra of introgressive lines with chromosome rearrangements of chromosome homoeologous groups 1 and 6.

K-1193 – the parental line of *Ae. columnaris*. Blocks of gliadin components controlled by wheat and *Aegilops* chromosomes are shown in the schemes. Individual components of blocks are shown by numerals. Chromosome substitution lines: *a* – 6A(6U$^c$); *b* – 6D(6U$^c$); *c* – 6A(6X$^c$); *d* – 6D(6X$^c$); *e* – lines without gliadin components controlled by chromosomes 6D, 6U and 6X. The components controlled by chromosomes 6A and 6B close in mobility to 6D-controlled components are indicated with arrows; *f* – 1D(1X$^c$), nulli-6A tetra-6D, absence the components controlled by chromosome 1D are indicated with arrows (left), components controlled by tetra 6D are indicated with arrows (right). The asterisks denote an absence of the gliadin components controlled by chromosome 6A; *g* – genetic control of *Ae. columnaris* gliadin, blocks of gliadin components controlled by homoeologous chromosome groups 1 and 6 are schematically represented for *Aegilops* (right) and wheat (left).

K-1193 parental line kernels were characterized by similar EP spectra.

Lines 1721/1, 1776/1, 1777/4, 2307/1W, 2307/1w, 2308/5, and 1777/1 (see the Table) were developed from the same crossing combination (Dobrynya × K-1193). Their karyotypes were characterized by monosomic or disomic substitution of wheat chromosome 6A with *Aegilops* chromosomes 6U$^c$ or 6X$^c$ (Fig. 1, *b, c*). According to cytological data, chromosomes 4B and 1D were absent in lines1721/1, 1776/1, 1777/4, and 2307/1w, where a pair of acrocentric chromosomes appeared instead (Fig. 1, *b*). C-banding and in situ hybridization with pSc119.2 and pAs1 DNA-probes demonstrated that thise chromosome forms resulted from translocation of the short arm and major part of the long arm of 1D to the 4B long arm

with a breakpoint at the pericentromeric region of 4B and distal section of 1DL (Badaeva et al., 2018).

All the four lines listed had identical gliadin EP spectra coded by the same set of GC locus alleles (see the Table). Gliadin components controlled by chromosome 6A were absent, but polypeptides appear in the spectrum, whose EP mobility corresponded to that of a number of *Aegilops* components (Fig. 2, *b*). This confirms the 6A wheat chromosome substitution with *Aegilops* chromosome and demonstrate the *Aegilops* gene expression.

All gliadin polypeptides controlled by 1D chromosome were identified in EP spectra for all the four lines, hence the translocated part of chromosome1D included GC locus as well. The GC genes in chromosomes1A, 1B, 6B, and 6D were fully expressed, similarly to the Dobrynya parent wheat cultivar (see Fig. 2, *b*).

The disomic addition of chromosome 6U$^c$ was identified in line karyotype2307/1W, in contrast to the other lines in the group. In this line, the EP spectrum included the components coded by both 6A and 6U$^c$ chromosomes (see Fig. 2, *b*). Lines with 6A(6U$^c$) substitution and disomic addition of 6U$^c$ had the same block of EP components, which indicates that the same *Aegilops* chromosome, i. e. 6U$^c$, was involved in the introgression.

The block of gliadin components controlled by *Gli-A1m* allele, which could only be introduced from the Saratovskaya 68 cultivar, was present in the 1777/1 line EP spectrum. Cytogenetic analysis showed that line 1777/1 carried substitution by two *Aegilops* chromosomes, i. e. 5X$^c$ and 6X$^c$ (see Fig. 1, *c*), uncharacteristic for the rest of the lines from the group (Badaeva et al., 2018). These results indicate random cross-pollination with the Saratovskaya 68 cultivar that occurred when line 1777/1 was synthesized or reproduced. A 6A a(6X$^c$) substitution was cytologically identified in the line, which was confirmed by EP analysis (see Fig. 2, *c*).

The next, i. e. the second, series of lines with numbers 11–16 (see the Table) was obtained involving L-503 common wheat cultivar. The allele composition of GC loci in these lines mostly matched the first parental cultivar biotype, apart from lines1813 and 2054/3, in which *Gli-Bs* allele characteristic of the Dobrynya cultivar appeared instead of *Gli-B2q*, and line 2304/1 with *Gli-A1f* allele. Karyotypes of lines from this series did not include chromosome 6D, substituted with 6U$^c$ or in one case with 6X$^c$. These substitutions were confirmed by EP analysis of the lines. Disappearance of the components coded by chromosome 6D and appearance of a group of components close in mobility to Aegilops parental line components was observed in lines 1813, 2054/3, 2306/3, and 2310/1 (see Fig. 2, *a*). The only line, where 6D chromosome was substituted with *Aegilops* chromosome 6X$^c$, i. e. 2304/1, possessed a unique composition of gliadin components controlled by *Aegilops* chromosome. It proves that a chromosome from another *Aegilops* genome, i. e. 6X$^c$, was involved in the substitution (see Fig. 2, *d*).

The next two lines, i. e. 2033/1 and 2034/3-1, were derived from crossings involving Dobrynya and Saratovskaya 68 cultivars. The allele composition of GC loci in these lines was overall consistent with the parental one. Segregating by *Gli*-A1 locus was observed in line 2034/3-1 (see the Table). The karyotypes of both lines did not include chromosome 6D,

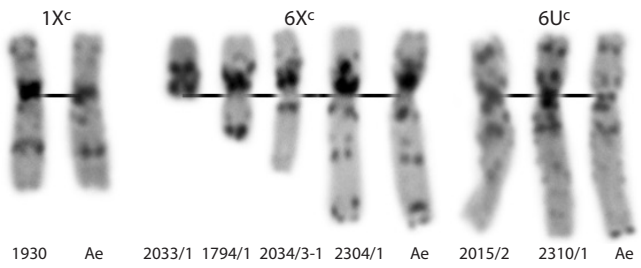



**Fig. 3.** Comparison of chromosomes 1X$^c$, 6U$^c$ and 6X$^c$ detected in introgressive lines with chromosomes of the parental *Ae. columnaris* accessions K-1193 (Ae).

as they were substituted with the deletion derivatives of chromosome 6X$^c$. Thus, chromosome 6X$^c$ in 2033/1 line was represented by a short arm telesomic (Fig. 3) and in line 2034/3-1 it had terminal deletion affecting about a half of the long arm. Components controlled by chromosome 6D are present in EP spectra for both lines. Their substitution with components similar to the ones identified in line 2304/1 and coded by *Aegilops* chromosome 6X$^c$ was visible only in line 2034/3-1 (see Fig. 2, *d*). No components coded by chromosome 6X$^c$ were identified in the EP spectrum of line 2033/1 (see Fig. 2, *e*).

Substitution of chromosomes of homoeologous group 1, i. e. 1D(1X$^c$), was only identified in line 1930 (Badaeva et al., 2018). It originated from cultivar Л-503 and thus falls into the second group of lines. Cytogenetic analysis showed that chromosomes1D and 6A were absent in the line karyotype (see the Table, Fig. 1, *a*). It was further confirmed by the EP analysis results: no components controlled by chromosome1D were identified in the gliadin spectrum, while the components close in mobility to Aegilops components were present (see Fig. 2, *f* ). Polypeptides controlled by wheat chromosomes 1A and 1B matched the parent cultivar spectrum. Absence of 6A chromosomes was compensated by an additional pair of chromosomes 6D, one of which carried unidentifiable terminal translocation in the long arm that did not affect the GC locus, whose genes were expressed. It was manifested in the form of clear color intensification in components controlled by chromosome 6D (see Fig. 2, *f* ), which implied a multiple increase in gene dose, and therefore, gene expression in all the four 6D chromosomes. No components controlled by chromosome 6A were visible in the EP image.

Line 1794/1 was not only the derivative of the Dobrynya cultivar but of the Saratovskaya 68 cultivar as well. Consequently, it carried the *Gli-A1m* allele. The karyotype of line 1794/1 did not include chromosome 6D, as it was substituted with a pair of translocated T7U$^c$S:6X$^c$S chromosomes (see Figs. 1, *d*, 3). The absence of 6D chromosome was further confirmed by EP analysis (see Fig. 2, *e*). However, no gliadin components controlled by 6X$^c$ *Aegilops* chromosome were identified in the EP spectrum.

## Discussion

Before using introgressive lines derived by remote hybridization in breeding practice, they need to be tested for alien genetic material, and this genetic material is to be characterized (Friebe et al., 1996; Molnár-Láng et al., 2014). Cytological approaches (such as differential chromosome banding or *in*

*situ* hybridization) permits identifying substituted/added lines and classifying them. However, these methods are rather labor intensive and not suited for massive analysis. Introduction of molecular genetic markers as a supplement for cytological analysis to characterize introgressive lines enhances the capability for accurate alien material identification, reduces the cost of testing and simplifies the procedure. It also permits evaluating the effect of alien chromosome introgression on gene expression.

EP analysis of gliadin spectra confirmed 6A, 6D, and 1D wheat chromosome substitution with *Aegilops* homoeologous chromosomes associated with $U^c$ or $X^c$ genomes in all the lines studied. The substitution manifested in the disappearance of GC gene expression products in chromosomes 6A, 6D or 1D with simultaneous appearance of the gene expression products localized in respective alien chromosomes. Comparison of EP spectra for lines with various wheat gene pools, which carried identical *Aegilops* chromosomes, showed that neither EP mobility, nor intensity of components within alien chromosome gliadin blocks change, i. e. they do not depend on line genotype. It makes it possible to use these markers to identify individual *Ae. columnaris* chromosomes in genetic and breeding lines.

The appearance in EP spectra of introgression lines, polypeptides controlled by *Aegilops* chromosome indicates the gene expression of this chromosome, as well as its functional activity within the alien wheat genome. In turn, alien chromosome does not affect the behavior of fully expressed wheat *Gli* genes, that is confirmed by the match between hybrids and parental cultivars in terms of gliadin spectrum. Such a "good neighborhood" between wheat and alien genetic material in the same nucleus may be considered as a relative phylogenetic affinity of these two species (or parts of their genetic material), which is determined by reticulate evolution, which is generally typical for the Triticeae tribe.

However, not all substitutions involving *Aegilops* chromosomes of homoeologous group 6 resulted in appearance of new gliadin components in EP spectra for hybrids. For instance, the analysis of lines with short arm of $6X^c$ *Aegilops* chromosome in the form of translocation (1794/1) or telocentric (2033/1) showed absence of the gliadin components present in lines with full $6X^c$ chromosome substitution. GC genes in wheat are known to be localized in terminal short arm sections of chromosomes from homoeologous groups 1 and 6 (Singh, Shepherd, 1988). Nevertheless, GC genes are not expressed in both these lines, in spite of chromosome short arms not being structurally rearranged. On the other hand, all gliadin components controlled by $6X^c$ *Aegilops* chromosome were present in the EP spectrum of 2034/3-1 line with large terminal long arm deletion in this chromosome. Based on these facts, we may infer that genes controlling gliadin synthesis are localized in the proximal part of the $6X^c$ *Ae. columnaris* chromosome long arm. It is possible that GC locus relocation from $6X^c$ chromosome short arm to the long one is associated with large species-specific pericentric inversion, accompanying divergence between *Ae. columnaris* and phylogenetically affine *Ae. triaristata.* This assumption is supported by cytogenetic analysis results, which identified significant morphological differences in 6 group orthologic chromosomes of X-genome in these species (Badaeva et al., 2004).

Four copies of chromosome 6D (tetra-6D) were identified using cytological methods in line 1930, with terminal long arm translocation identified in one pair. Enhanced intensity of EP components controlled by chromosome 6D indicates that all tetra-6D loci were expressed in line 1930. Independence of gliadin gene expression on the presence of chromosome rearrangements may be observed in the other lines as well. Particularly, all six introgressive lines, in which a major part of the 1D chromosome, including the short arm, was translocated to the 4B chromosome long arm, show fully expressed GC genes in the 1D chromosome. It shows that the 1D chromosome functioning does not depend on the loss of a part of its long arm and combination with nonhomologous chromosome from another genome (4BL).

Introgressions of homoeologous group 6 chromosomes were encountered in the sample studied significantly more frequently than those of group 1: $1D(1X^c)$ substitution was only identified in one of 18 lines. The absence of substitutions involving chromosome $1U^c$ may be due to its low compensatory capability, since the parent K-1193 *Ae. columnaris* sample had this chromosome rearranged as a result of major translocation. Due to lack of samples carrying substitutions or translocations associated with $1U^c$ chromosome, we have only been able to identify gliadin EP component blocks controlled by *Aegilops* chromosomes $1X^c$, $6U^c$, and $6X^c$. Spectrum components, whose genetic control remained unidentified, were hypothetically attributed to polypeptides controlled by $1U^c$ chromosome (see Fig. 2, *g*). The data obtained may be used in future research to test breeding material for presence of these chromosomes.

EP spectra analysis of the two species belonging to *Aegilops* and *Triticum* genera indicates that there are no individual species-specific EP components. At the same time, we may define specific groups of components inherited as the blocks controlled by linked gene clusters. Component blocks differ in their structure or their pattern. Based on these differences, intraspecific and interspecific diversity can be estimated. The block nature of gliadin component inheritance in *Aegilops* indicates the similarity between the specimens of these genera in their GC locus organizational structure.

## Acknowledgments

The authors thank S.N. Sibikeev, Head of Genetics and Cytology Laboratory at Agricultural Research Institute for South-East Region, and A.E. Druzhinina for the provided materials. The work was supported in part by the Russian Foundation for Basic Research, project No. 17-04-00087.

## Conflict of Interest

The authors declare they have no conflict of interest.

## References

Badaeva E.D., Amosova A.V., Samatadze T.E., Zoshchuk S.A., Shostak N.G., Chikida N.N., Zelenin A.V., Raupp W.J., Friebe B., Gill B.S. Genome differentiation in *Aegilops*. 4. Evolution of the U-genome cluster. Plant. Syst. Evol. 2004;246(1-2):45-76.

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

Badaeva E.D., Ruban A.S., Shishkina A.A., Sibikeev S.N., Druzhin A.E., Surzhikov S.A., Dragovich A.Y. Genetic classification of

*Aegilops columnaris* Zhuk. ($2n = 4x = 28$, $U^cU^cX^cX^c$) chromosomes based on FISH analysis and substitution patterns in common wheat × *Ae. columnaris* introgressive lines. Genome. 2018;61(2):131-143. DOI 10.1139/gen-2017-0186.

Damania A.B., Altunji H., Dhaliwal H.S. Evaluation of *Aegilops* spp. for drought and frost tolerance. Genetic Resources Unit Annual Report, ICARDA. 1992;45-46.

Dvořák J. Genome analysis in the *Triticum-Aegilops* alliance. In: Slinkard A.E. (Ed.) Proc. of the 9th Int. Wheat Genetics Symp., 2–7 Aug. 1998. Saskatoon, Saskatchewan: Printcrafters Inc., 1998; 8-11.

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

Garg M., Tanaka H., Ishikawa N., Takata K., Yanaka M., Tsujimoto H. A novel pair of HMW glutenin subunits from *Aegilops searsii* improves quality of hexaploid wheat. Cereal Chem. J. 2008;86:26-32.

Garg M., Tsujimoto H., Gupta R.K., Kumar A., Kaur N., Kumar R., Chunduri V., Sharma N.K., Chawla M., Sharma S., Mundey J.K. Chromosome specific substitution lines of *Aegilops geniculata* alter parameters of bread making quality of wheat. PLoS One. 2016;11: e0162350.

Gill B.S., Friebe B., Endo T.R. Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). Genome. 1991;34:830-839.

Gill B.S., Sharma C., Raupp W.J., Browder L.E., Heachett J.H., Harvey T.L., Moseman J.G., Waines J.G. Evaluation of *Aegilops* species for resistance to wheat powdery mildew, wheat leaf rust, Hessian fly, and greenbug. Plant Dis. 1985;69:314-316.

Kasarda D.D., Bernardin J.E., Qualset C.O. Relationship of gliadin protein components to chromosomes in hexaploid wheats (*Triticum aestivum* L.). Proc. Natl. Acad. Sci. USA. 1976;73:3646-3650.

McIntosh R.A., Yamazaki Y., Dubkovsky G., Rogers J., Morris C.F., Appels R., Xia X.C. Catalogue of Gene Symbols for Wheat. The 12th Int. Wheat Genetics Symp., 8–13 Sept. 2013. Yokohama, Japan, 2013;395.

Metakovsky E.V., Novoselskaya A.Yu. Gliadin allele identification in common wheat. 1. Methodological aspects. J. Genet. Breed. 1991; 45:319-323.

Molnár-Láng M., Ceoloni C., Doležel J. (Eds.). Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and Genomics. Switzerland: Springer Int. Publ., 2015.

Molnár-Láng M., Molnár I., Szakács É., Linc G., Bedö Z. Production and Molecular Cytogenetic Identification of Wheat-Alien Hybrids and Introgression Lines. In: Tuberosa R., Graner A., Frison E. (Eds.). Genomics of Plant Genetic Resources. Vol. 1. Managing, Sequencing and Mining Genetic Tesources. Springer, 2014;255-284.

Monneveux P., Zaharieva M., Rekika D. The utilisation of *Triticum* and *Aegilops* species for the improvement of durum wheat. In: Royo C., Nachit M., Di Fonzo N., Araus J.L. (Eds.). Durum Wheat Improvement in the Mediterranean Region: New Challenges. Zaragoza: CIHEAM, 2000;71-81.

Novoselskaya-Dragovich A.Yu. Genetics and genomics of wheat: storage proteins, ecological plasticity, and immunity. Russ. J. Genet. 2015;51(5):476-490. DOI 10.1134/S102279541505004X.

Novoselskaya-Dragovich A.Yu., Krupnov V.A., Saifulin R.A., Puhalskiy V.A. Dynamics of genetic variation at gliadin coding loci in Saratov cultivars of common wheat *Triticum aestivum* L. for over eight decades of scientific breeding. Russ. J. Genet. 2003;39(10):1130-1137.

Rakszegi M., Molnár I., Lovegrove A., Darkó É., Farkas A., Láng L., Bedo Z., Doležel J., Molnár-Láng M., Shewry P. Addition of *Aegilops* U and M chromosomes affects protein and dietary fiber content of wholemeal wheat flour. Front. Plant Sci. 2017;8:1529. DOI 10.3389/fpls.2017.01529.

Rawat N., Neelam K., Tiwari V.K., Randhawa G.S., Friebe B., Gill B.S., Dhaliwal H.S. Development and molecular characterization of wheat-*Aegilops kotschyi* addition and substitution lines with high grain protein, iron, and zinc. Genome. 2011;54:943-953.

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

Shepherd K.W. Chromosomal control of endosperm proteins in wheat and rye. In: Finlay K.W., Sherherd K.W. (Eds.). Proc. of the 3rd Int. Wheat Genetics Symp. Canberra: Austral. Acad. Sci., 1968;86-96.

Shishkina A.A., Dragovich A.Yu., Rouban A.S., Sibikeev S.N., Druzhin A.E., Badaeva E.D. Development of the genetic classification of *Aegilops columnaris* Zhuk. chromosomes based on the analysis of introgression lines *Triticum aestivum* × *Ae. columnaris*. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2017;21(2):241-249. DOI 10.18699/VJ17.243 (in Russian)

Singh N.K., Shepherd K.W. Linkage mapping of genes controlling endosperm storage proteins in wheat: 1. Genes on the short arms of group 1 chromosomes. Theor. Appl. Genet. 1988;75:628-641.

Tiwari V., Rawat N., Neelam K., Kumar S., Randhawa G., Dhaliwal H. Substitutions of 2S and 7U chromosomes of *Aegilops kotschyi* in wheat enhance grain iron and zinc concentration. Theor. Appl. Genet. 2010;121(2):259-269.

Upelniek V.P., Novoselskaya-Dragovich A.Yu., Shishkina A.A., Mel'nik V.A., Dedova L.V., Kudrjavtsev A.M. The laboratory analysis of wheat seed proteins. Technological instruction. Diagnostics of varietal identity and purity of seed wheat. Moscow: Vavilov Institute of General Genetics Publ., 2013. (in Russian)

Warham E.J., Mujeeb-Kazi A., Rosas V. Karnal bunt (*Tilletia indica*) resistance screening of *Aegilops* species and their practical utilization for *Triticum aestivum* improvement. Can. J. Plant Pathol. 1986; 8:65-70.

# Изучение генетического полиморфизма диплоидной пшеницы *Triticum boeoticum* Boiss. с использованием SSR-маркеров

М.А. Аббасов

Институт генетических ресурсов Национальной академии наук Азербайджана, Баку, Азербайджан

Диплоидная пшеница *Triticum boeoticum* Boiss. (геном AA) – потенциальный источник новых ценных аллелей для улучшения возделываемых видов пшеницы. В связи с этим оценка внутривидового разнообразия *T. boeoticum* и ДНК-паспортизация образцов этого вида является актуальной задачей. В настоящей работе исследовано генетическое разнообразие более 60 образцов *T. boeoticum* с использованием 11 микросателлитных маркеров. По данным SSR-анализа было идентифицировано 83 аллеля, в среднем наблюдалось по 7.5 аллелей на локус. Величины ожидаемой ($H_E$) и наблюдаемой ($H_O$) гетерозиготности варьировали в пределах 0.17–0.89 и 0.00–0.74 при среднем показателе $H_E = 0.52$ и $H_O = 0.13$ соответственно. Значение PIC для каждого локуса находилось в пределах 0.17–0.88 и в среднем равнялось 0.49. Для всех изученных локусов были обнаружены уникальные аллели. Кластерный анализ позволил объединить изученные образцы в пять основных групп, расстояния между группами варьировали от 0 до 1, что указывает на высокий уровень генетических различий в исследуемой коллекции. Согласно анализу PCoA, было образовано пять основных групп и выявлены некоторые соответствия с дендрограммой. При обобщении полученных данных PCoA и кластерного анализа отмечена слабая генетическая дифференциация изученной коллекции *T. boeoticum*. Корреляция генетического расстояния с географическим происхождением выявлена лишь для образцов диплоидной пшеницы *T. boeoticum* из Ирана. Анализ образцов показывает широкое разнообразие *T. boeoticum* по микросателлитным локусам. Полученные нами данные расширяют представления и дают дополнительную информацию о генетической структуре коллекции и разнообразии изученных образцов *T. boeoticum*.

Ключевые слова: *T. boeoticum*; генетический полиморфизм; SSR-маркеры; кластерный анализ; анализ PCoA.

**КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:**
Аббасов М.А. Изучение генетического полиморфизма диплоидной пшеницы *Triticum boeoticum* Boiss. с использованием SSR-маркеров. Вавиловский журнал генетики и селекции. 2018;22(5):515-523. DOI 10.18699/VJ18.389

**HOW TO CITE THIS ARTICLE:**
Abbasov M.A. Study of the genetic polymorphism of diploid wheat *Triticum boeoticum* Boiss. using SSR markers. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):515-523. DOI 10.18699/VJ18.389 (in Russian)

Received 18.01.2018
Accepted for publication 04.06.2018


e-mail: mehraj_genetic@yahoo.com

## Study of the genetic polymorphism of diploid wheat *Triticum boeoticum* Boiss. using SSR markers

M.A. Abbasov

Genetic Resources Institute of Azerbaijan National Academy of Sciences, Baku, Azerbaijan

Diploid wheat *Triticum boeoticum* Boiss. (genome constitution AA) is a promising source of new valuable alleles for improving cultivated wheat species. Therefore, the evaluation of the intraspecies diversity of *T. boeoticum* and DNA fingerprinting of accessions of this species are topical tasks. In this paper, the genetic diversity of over 60 *T. boeoticum* accessions was studied using 11 SSR markers. The analysis revealed 83 alleles, 7.5 alleles per locus on the average. The values of expected ($H_E$) and observed ($H_O$) heterozygosity varied within 0.00–0.74 and 0.17–0.89, respectively, the average indices being $H_O = 0.13$ and $H_E = 0.52$. The PIC value for each locus was within 0.17–0.88, 0.49 on the average. Unique alleles were found in all loci studied. Cluster analysis allowed the accessions studied to be combined into five major groups. The distances between the groups varied from 0 to 1, pointing to a high level of genetic differences in the collection under study. On the base of PCoA, five major groups were formed and some correspondence with the dendrogram was detected. Summarizing the data of PCoA and cluster analysis, we noted a weak genetic differentiation in the studied collection of *T. boeoticum*. A correlation between the genetic distance and geographic origin was revealed only for accessions of diploid wheat *T. boeoticum* from Iran. The analysis of the *T. boeoticum* accessions studied showed a wide diversity for SSR loci. The results expand our knowledge and provide additional information on the genetic structure of the collection and on the genetic diversity of *T. boeoticum* accessions studied.

Key words: *T. boeoticum*; genetic polymorphism; SSR markers; cluster analysis; principal coordinates analysis.

Пшеница – одна из самых важных и агрономически значимых сельскохозяйственных культур, которая широко возделывается по всему миру (Hajiyev et al., 2015). Род *Triticum* L. состоит из четырех групп, в том числе: einkorn ($2n = 2x = 14$, AA), emmer ($2n = 4x = 28$, AABB), timopheevi ($2n = 4x = 28$, AAGG) и обычной пшеницы ($2n = 6x = 42$, AABBDD). Три вида – *T. monococcum* L., *T. boeoticum* Boiss. и *T. urartu* Thum. ex Gandil. – относятся к группе пшеницы einkorn (Mizumoto et al., 2002). Как известно, дикие виды пшеницы являются важным источником для улучшения генетических признаков пшеницы. *T. boeoticum* и *T. urartu* – два основных вида-кандидата для донора генома A для мягкой и твердой пшеницы (Farouji et al., 2015). Показано, что вид *T. monococcum* произошел от *T. boeoticum* и *T. urartu* был донором генома для полиплоидных видов пшеницы (Конарев и др., 1976; Dvorak et al., 1993; Takumi et al., 1993). Первичное место произрастания *T. boeoticum* – центральная и восточная часть Плодородного Полумесяца (Zohary, Hopf, 1993). В результате анализа 288 AFLP локусов *T. monococcum* и его диких сородичей M. Heun с сотрудниками (1997) установили, что *T. boeoticum* из Каракадагских гор (Юго-Восточная Турция) оказался вероятным предшественником культурной однозернянки. *T. boeoticum* Boiss. с геномом $A^bA^b$ – источником ценных генов для улучшения современных сортов пшеницы (Bahrai et al., 1998; Harjit-Singh et al., 2000; Anker, Niks, 2001). Несмотря на низкую потребительскую ценность по сравнению с *T. durum* и *T. aestivum*, *T. boeoticum* содержит богатое аллельное разнообразие для адаптивных признаков. Как известно, окультивирование и различные стратегии селекции привели к огромной потере аллелей и ограничению генетического разнообразия современных сортов пшеницы – это проблема в селекции пшеницы на устойчивость к биотическим и абиотическим стрессорам (Aliyev et al., 2007; Wang et al., 2017). Поэтому необходимо проводить исследования гермоплазмы пшеницы с целью расширения генетического разнообразия для селекционных программ. В ближайшем будущем ожидается, что *T. boeoticum* будет играть важную роль в генетических и геномных исследованиях пшеницы. Для эффективного использования ресурсов *T. boeoticum* в программах генетического улучшения пшеницы необходимо оценить разнообразие этого вида на уровне генома. При изучении генетического разнообразия можно использовать различные подходы, такие как анализ родословных, морфологические признаки или молекулярные маркеры. Использование систем молекулярных маркеров позволяет изучать полиморфизм ДНК, устанавливать генетические взаимоотношения, выявлять гены хозяйственно ценных признаков, что представляет собой актуальное направление в селекции сельскохозяйственных культур (Cox et al., 1985; Motawei et al., 2007; Cifci, Yagdi, 2012; Abouzied et al., 2013; Malik et al., 2013).

Генетическое разнообразие диплоидной пшеницы оценивали с использованием маркеров: RFLP (Figliuolo, Perrino, 2004), RAPD (Ovesna et al., 2002), AFLP (Malaki et al., 2006), IRAP (Farouji et al., 2015) и ISSR (Kojima et al., 1998). Наиболее широкое применение в генетическом анализе получили микросателлитные маркеры (Chen et al., 1994). Благодаря таким качествам, как воспроизводимость, мультиаллельная природа, кодоминантный характер наследования и хромосомоспецифичность, они считаются высокоинформативными генетическими маркерами (Babayeva et al., 2009). Доказано, что SSR-маркеры – эффективный инструмент для анализа генетического разнообразия различных видов пшеницы (McLauchlan et al., 2001).

Несмотря на большой перечень публикаций по оценке генетического разнообразия популяций *T. boeoticum* с применением различных маркерных систем (Korzun et al., 1998; Hammer et al., 2000; Mousavifard et al., 2014), это направление исследований все еще остается актуальным. Нами поставлена задача – изучить генетическое разнообразие 63 образцов коллекции дикой однозернянки и определить степень родства и характер различий между генотипами с использованием SSR-маркеров.

## Материал и методы

Объектом для молекулярных исследований служили 63 образца диплоидной пшеницы *T. boeoticim* различного происхождения из коллекции Национального генетического банка Института генетических ресурсов Национальной академии наук Азербайджана и генбанка Международного центра сельскохозяйственных исследований в засушливых районах (ИКАРДА). Названия и происхождение изученных генотипов приведены в табл. 1.

**Выделение ДНК и ПЦР.** Выделение геномной ДНК и ПЦР проводили согласно ранее описанному протоколу (Zhang et al., 2010). ПЦР-фрагменты анализировали ДНК-анализатором ABI PRISM 3730 (Applied Biosystems, Fargo, ND, США) и оценивали с помощью Gene Marker версии 1.6 (Soft Genetics LLC, State College, PA, США).

**Анализ данных.** В статистический анализ были включены четкие и воспроизводимые результаты. Продукты амплификации документировали в виде фрагментов, SSR-профили были составлены на основе их размера (bp). Количество аллелей (Na), частота встречаемости основных аллелей, наблюдаемая ($H_O$) и ожидаемая ($H_E$) гетерозиготность, величина информационного полиморфизма (PIC) были рассчитаны для каждого локуса с использованием программного обеспечения PowerMarker, версия 3.25 (Liu, Muse, 2005). В работе было использовано девять микросателлитных маркеров *Xbarc* (the USDA-ARS Beltsville Agricultural Research Center, США) (Song et al., 2002, 2005) и два маркера *Xgwm* (Gatersleben Wheat Microsatellites, IPK-Gatersleben, Германия) (Röder et al., 1998; Pestsova et al., 2000). Кластерный анализ и анализ главных компонент (PCoA) осуществлены с помощью программного пакета DARwin 6.0 (Perrier, Jacquemoud-Collet, 2006) с использованием алгоритма кластеризации UNJ (unweighted Neighbor-joining), предложенного O. Gascuel (1997).

## Результаты

**Генетический полиморфизм.** Для оценки генетического полиморфизма 63 образцов *T. boeoticum* использован SSR-метод и проанализировано 11 микросателлитных локусов (табл. 2). В целом идентифицировано 83 аллеля, в среднем наблюдалось по 7.5 аллелей на локус. Количество синтезированных аллелей варьировало от 4 (маркеры *Xbarc200*, *Xbarc101*, *Xbarc142* и *Xgwm219*) до 15 (маркеры *Xbarc213*

**Table 1.** Names and geographical origins of *T. boeoticum* accessions used in the study

| No. | Accession ID | Origin | Region | Altitude | Longitude | Latitude |
|---|---|---|---|---|---|---|
| 1 | – | AZE | – | – | – | – |
| 2 | – | AZE | – | – | – | – |
| 3 | – | AZE | – | – | – | – |
| 4 | – | AZE | – | – | – | – |
| 5 | – | AZE | – | – | – | – |
| 6 | 44821 | TUR | Kayseri | 1888 | E36 47 | N38 51 |
| 7 | 137335 | ARM | Kotayk | 1282 | E044 36 17 | N40 09 04 |
| 8 | 44833 | TUR | Kırşehir | 1086 | E33 43 | N39 22 |
| 9 | 137409 | ARM | Vayots Dzor | 1852 | E045 07 16 | N39 49 18 |
| 10 | 116146 | TUR | Gaziantep | 796 | E37 17 32 | N37 16 27 |
| 11 | 113283 | IRN | Kordestan | 1302 | E46 08 | N35 30 |
| 12 | 44861 | TUR | Çanakkale | 126 | E26 34 | N40 03 |
| 13 | 113291 | IRN | Khuzestan | 706 | E48 52 | N32 33 |
| 14 | 113311 | IRN | East Azerb. | 1127 | E47 45 | N37 23 |
| 15 | 44937 | SYR | Damascus | 1134 | E36 03 49 | N33 35 48 |
| 16 | 44813 | TUR | Kırşehir | 1086 | E33 43 | N39 22 |
| 17 | 113261 | IRN | Kermanshah | 1433 | E46 25 | N34 20 |
| 18 | 109086 | IRG | Ninawa | 847 | E41 51 | N36 21 |
| 19 | 113254 | IRN | Lorestan | 1554 | E48 45 | N33 53 |
| 20 | 109080 | IRG | Ninawa | 1216 | E41 47 | N36 23 |
| 21 | 113266 | IRN | Kermanshah | 1433 | E46 25 | N34 20 |
| 22 | 116138 | TUR | Gaziantep | 678 | E37 28 15 | N36 55 32 |
| 23 | 110826 | LBN | Rachaiya | 1224 | E35 54 | N33 38 |
| 24 | 44818 | TUR | Kayseri | 1656 | E36 30 | N38 48 |
| 25 | 113282 | IRN | West Azerb. | 1819 | E47 06 | N36 24 |
| 26 | 113264 | IRN | Ilam | 1558 | E46 27 | N33 37 |
| 27 | 44941 | ARM | Ararat | 1668 | E44 50 | N39 50 |
| 28 | 113275 | IRN | Kordestan | 1302 | E46 08 | N35 30 |
| 29 | 109081 | IRG | Ninawa | 1216 | E41 47 | N36 23 |
| 30 | 45103 | ARM | – | – | – | – |
| 31 | 44820 | TUR | Kayseri | 1656 | E36 30 | N38 48 |
| 32 | 113309 | IRN | East Azerb. | 1127 | E47 45 | N37 23 |
| 33 | 113260 | IRN | Kermanshah | 1433 | E46 25 | N34 20 |
| 34 | 135341 | SYR | Al Hasakah | 470 | E41 34.571 | N37 04.220 |
| 35 | 44949 | IRN | East Azerb. | 1321 | E46 10 | N37 20 |
| 36 | 131176 | TUR | Ankara | 904 | E32 45 | N40 00 |
| 37 | 113280 | IRN | West Azerb. | 1407 | E45 44 | N36 44 |
| 38 | 137477 | ARM | Ararat | 1694 | E045 00 03 | N39 49 29 |
| 39 | 137341 | ARM | Kotayk | 1282 | E044 36 17 | N40 09 04 |
| 40 | 44868 | TUR | Bursa | 567 | E29 35 | N40 20 |
| 41 | 113288 | IRN | Khuzestan | 706 | E48 52 | N32 33 |
| 42 | 113284 | IRN | Kordestan | 2032 | E45 54 | N35 58 |
| 43 | 110789 | SYR | Al Hasakah | 337 | E42 12 43 | N37 17 35 |
| 44 | 44822 | TUR | Kayseri | 1888 | E36 47 | N38 51 |
| 45 | 116134 | TUR | Gaziantep | 692 | E37 21 04 | N36 52 55 |
| 46 | 44955 | LEB | Baalbek | 1083 | E36 07 | N34 05 |
| 47 | 113286 | IRN | Khuzestan | 706 | E48 52 | N32 33 |
| 48 | 113277 | IRN | Kordestan | 1311 | E46 09 | N35 30 |

**End of Table 1**

| No. | Accession ID | Origin | Region | Altitude | Longitude | Latitude |
|---|---|---|---|---|---|---|
| 49 | 44856 | BGR | Burgas | 166 | E26 21 | N42 11 |
| 50 | 116153 | TUR | Gaziantep | 838 | E37 11 57 | N36 53 02 |
| 51 | 116139 | TUR | Gaziantep | 625 | E37 31 06 | N36 52 51 |
| 52 | 44857 | BGR | Burgas | 192 | E26 20 | N42 09 |
| 53 | 139313 | SYR | Sweida | 1399 | E36 50.388 | N32 41.993 |
| 54 | 110749 | SYR | Aleppo | 541 | E37 35 00 | N36 39 00 |
| 55 | 113273 | IRN | Kordestan | 1614 | E46 59 | N35 19 |
| 56 | 44953 | SYR | Aleppo | 485 | E36 58 00 | N36 29 00 |
| 57 | 131180 | IRG | Sulaymaniyah | 941 | E45 24 | N35 23 |
| 161 | 44904 | SSR | – | 25 | E27 09 | N38 26 |
| 162 | 44900 | SSR | – | 25 | E27 09 | N38 26 |
| 163 | 113258 | IRN | – | – | E48 21 | N33 29 |
| 164 | 44936 | SAR | – | – | E36 01 40 | N33 38 10 |
| 165 | 110820 | LEB | – | – | E35 49 | N33 30 |
| 166 | 44948 | LEB | – | – | E35 56 | N33 35 |

AZE, Azerbaijan; TUR, Turkey; ARM, Armenia; IRN, Iran; SYR, Syria; IRG, Iraq; LEB, Lebanon; BGR, Bulgaria; SAR, Saudi Arabia; SSR, Serbia.

**Table 2.** Genetic diversity parameters of 63 *T. boeoticum* accessions according to SSR markers

| Locus | Chromosome | Number of alleles | Major allele frequency | $H_O$ | $H_E$ | PIC |
|---|---|---|---|---|---|---|
| Xbarc213 | 1AL | 15 | 0.18 | 0.02 | 0.89 | 0.88 |
| Xbarc15 | 2AL | 8 | 0.46 | 0.74 | 0.62 | 0.54 |
| Xbarc1021 | 3AL | 9 | 0.29 | 0.11 | 0.82 | 0.79 |
| Xbarc206 | 4AS | 8 | 0.82 | 0.03 | 0.32 | 0.31 |
| Xbarc117 | 5AS | 7 | 0.42 | 0.06 | 0.73 | 0.69 |
| Xbarc174 | 1BL | 15 | 0.38 | 0.24 | 0.81 | 0.79 |
| Xbarc200 | 2BS | 4 | 0.70 | 0.10 | 0.44 | 0.37 |
| Xbarc101 | 2BL | 4 | 0.70 | 0.00 | 0.45 | 0.38 |
| Xbarc142 | 5BL | 4 | 0.91 | 0.00 | 0.17 | 0.17 |
| Xgwm361 | 6BS | 5 | 0.89 | 0.14 | 0.21 | 0.21 |
| Xgwm219 | 6BL | 4 | 0.84 | 0.02 | 0.29 | 0.27 |
| Mean | | 7.5 | 0.60 | 0.13 | 0.52 | 0.49 |
| Total | | 83 | | | | |

$H_O$, observed heterozygosity; $H_E$, expected heterozygosity; PIC – polymorphism information content.

и *Xbarc174*). Для всех изученных локусов были обнаружены уникальные аллели, общее число которых составило 27 (диапазон 1–5). Наибольшее количество уникальных аллелей было выявлено праймерами *Xbarc206*, *Xbarc174* и *Xbarc15*. Нами рассчитаны частоты встречаемости выявленных аллелей, которые послужили основой для расчета индекса информативности (PIC) каждого маркера. Частота основных аллелей варьировала от 0.18 (*Xbarc213*) до 0.91 (*Xbarc142*) и в среднем составила 0.60. Анализ изменчивости исследованных локусов показал, что по значениям наблюдаемой и ожидаемой гетерозиготности прослеживаются существенные различия. Величины ожидаемой ($H_E$) и наблюдаемой ($H_O$) гетерозиготности варьировали в пределах 0.17–0.89 и 0.00–0.74 при среднем показателе $H_E = 0.52$ и $H_O = 0.13$ соответственно. Среднее значение PIC составило 0.49 и изменялось от 0.17 в локусе *Xbarc142* до 0.88 в локусе *Xbarc213*. В локусах *Xbarc213*, *Xbarc1021*, *Xbarc174* и *Xbarc15* выявлено значение PIC более 0.5 (см. табл. 2 и 3).

**Кластерный анализ и метод главных компонент.** Проведен кластерный анализ и построена дендрограмма, отображающая генетические различия между изученными образцами пшеницы (рис. 1). Генотипы были сгруппированы в пять различных кластеров, где индекс генетического

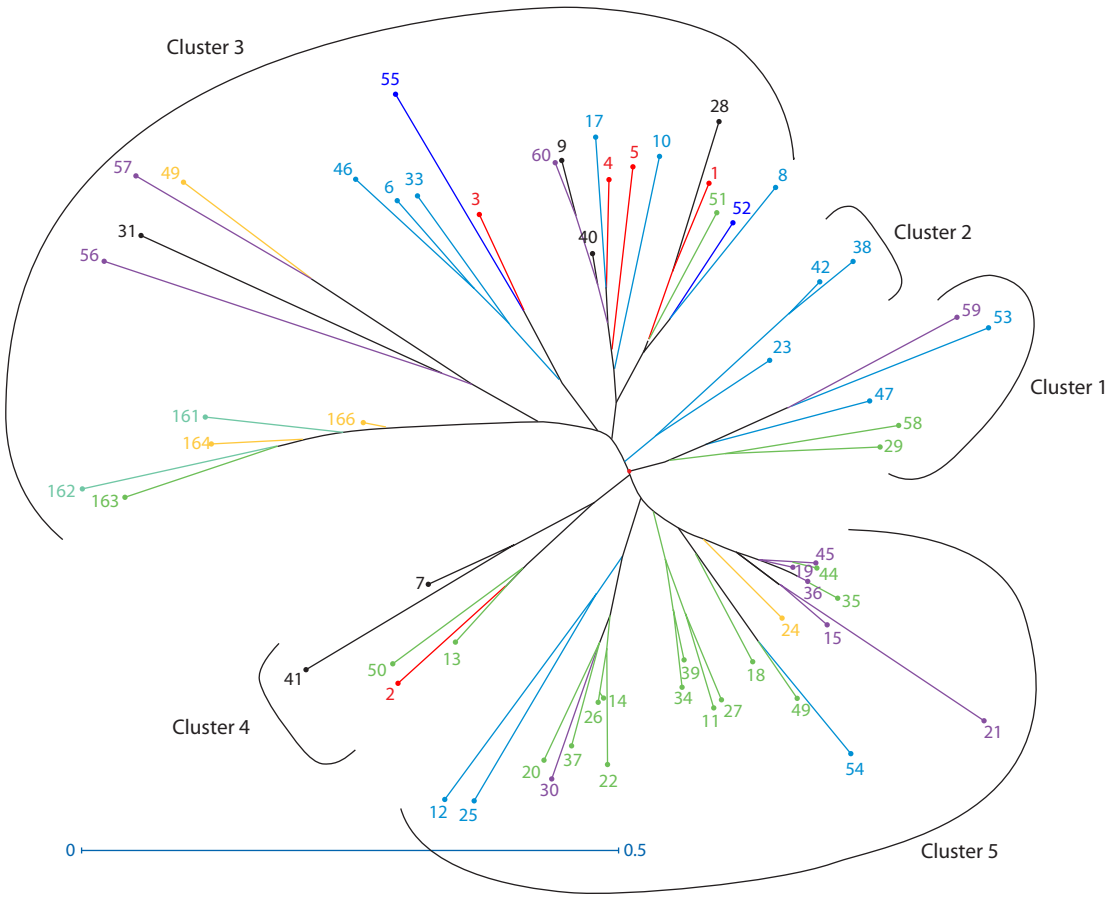



**Fig. 1.** Genetic distances deduced from the results of the SSR analysis. Different colors indicate accessions from different geographical regions: blue, Bulgaria; green, Iran; red, Azerbaijan; violet, Syria; yellow, Lebanon; black, Armenia; blue, Turkey; turquoise, Serbia.

расстояния (ИГР) составил от 0 до 1. Количество сгруппированных образцов варьировало от 3 до 27. Отмечена преимущественная локализация образцов в кластерах 3 и 5. Самый обширный – третий кластер, состоящий из 27 образцов, в свою очередь подразделяется на два больших субкластера. Первый и четвертый кластеры включают в себя по пять образцов, а наименьшим числом генотипов представлен кластер 2. При анализе дендрограммы (см. рис. 1) можно отметить, что установлена некоторая корреляция между географическими ареалами и генетическим расстоянием. Наименьшее генетическое расстояние было выявлено между генотипами под номерами 45 и 44, 22 и 14 (ИГР = 0.06) и между 35 и 36 (ИГР = 0.05). Между генотипами 164 и 165 отмечено полное генетическое сходство (ИГР = 0). Наибольшее генетическое расстояние установлено между генотипами под номерами 22 и 57, 30 и 162, 37 и 57, 46 и 57, где ИГР был равен 1.

Генетические взаимоотношения у образцов *T. boeoticum* были дополнительно исследованы с использованием анализа главных компонент (PCoA). Согласно анализу PCoA, было образовано пять основных групп и показано соответствие с дендрограммой. Первые две основные оси дифференциации (PCo1 и PCo2) объясняют 24.96 % от общей вариации (рис. 2).

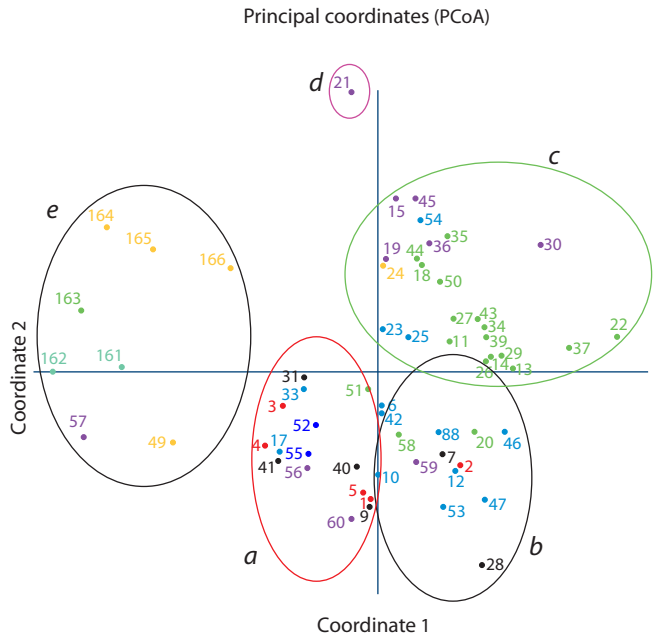



**Fig. 2.** PCoA analysis of SSR data showing the grouping of accessions of the *T. boeoticum* species. .

Notation, see in the signature to the fig. 1.

## Обсуждение

Исследования, направленные на выявление генетического разнообразия пшеницы с помощью SSR-маркеров, ведутся во всем мире. M.R. Naghavi с коллегами (2004) провели сравнительный анализ генетического разнообразия генотипов пшеницы на основе 17 RAPD- и 35 SSR-маркеров. В работе S. Sud с коллегами (2005), посвященной необходимости создания новых сортов пшеницы с разнообразным генетическим фоном и включения новой изменчивости в существующий генофонд пшеницы, осуществлен анализ родословных у 20 элитных сортов пшеницы с использованием 25 микросателлитных маркеров. Наконец, B. Zeb с коллегами (2009) изучили на 10 генотипах различные разновидности пшеницы с помощью 14 SSR-маркеров, которые могут быть использованы в будущих селекционных программах.

Настоящее исследование направлено на оценку генетической вариабельности у 63 образцов диплоидной пшеницы вида *T. boeoticum* с использованием 11 SSR-маркеров. Во многих работах доказана эффективность маркеров *Xbarc* и *Xgwm* для изучения генетического разнообразия пшеницы (Dresigacker et al., 2004; Drikvand et al., 2012; Spanic et al., 2012; Kumar et al., 2016). В общей сложности определено 83 аллеля. Число аллелей на локус варьировало от 4 до 15 и в среднем составило 7.5 аллелей. Диапазон длины полученных фрагментов находился в пределах 100–323 п.н. Наши результаты согласуются с данными других авторов. Так, M.R. Naghavi с коллегами (2004) для 36 популяций *T. boeoticum* с использованием 17 SSR-маркеров выявил 147 аллелей, в среднем 8.5 аллеля на локус. Рядом исследователей показан высокий уровень полиморфизма микросателлитных локусов генома диплоидных пшениц по сравнению с другими маркерными системами (Medini et al., 2005; Naghavi et al., 2011). Следует подчеркнуть, что информативность SSR-маркеров может быть связана с уникальным механизмом, ответственным за генерацию аллельного разнообразия посредством проскальзывания репликации.

Выявление уникальных и разнообразных аллелей дает возможность исследовать генетическое разнообразие и специфичность сортов, их идентификацию и регистрацию, улучшение культурных растений, включая пшеницу (Abouzied et al., 2013). Кроме того, наличие уникальных аллелей определяет индивидуальность популяций, что подразумевает наличие генетической вариации, необходимой для адаптации к экологическим условиям. В наших исследованиях обнаружено 19 генотипов, для которых идентифицированы специфичные аллели. Каждый из них содержал от одного до четырех таких аллелей. У образца под номером 56 выявлены специфичные аллели в большинстве локусов (*Xbarc15*, *Xbarc174*, *Xbarc101* и *Xbarc142*), что указывает на уникальность этого генотипа. Чаще такие аллели были обнаружены в локусах *Xbarc206* у генотипов 31, 41, 49, 57 и 161 (хромосома 4AS, 5 аллелей), *Xbarc174* – у генотипов 1, 5, 53, 55 и 56 (хромосома 1BL, 5 аллелей) и *Xbarc15* – у генотипов 12, 31, 41 и 56 (хромосома 2AL, 4 аллеля), т.е. маркеры *Xbarc206*, *Xbarc174* и *Xbarc15* наиболее эффективны, их можно рекомендовать для идентификации образцов *T. boeoticum*.

Показатель PIC характеризует дискриминационную силу локуса не только по количеству выявленных аллелей, но и по их относительным частотам. Для отобранных нами локусов значение PIC варьировало от 0.17 для *Xbarc142* до 0.88 для *Xbarc213* и в среднем было 0.49. Значения PIC для локусов *Xbarc206*, *Xbarc200*, *Xbarc101*, *Xbarc142*, *Xgwm361* и *Xgwm219* находились в пределах 0.17–038, что достаточно для идентификации изучаемых образцов. Остальные пять локусов с показателями PIC более 0.5 особенно эффективны для дифференциации изученной группы генотипов. Высокое значение PIC свидетельствовало о широком генетическом разнообразии в изученной коллекции диплоидной пшеницы. Аналогичные результаты для разных коллекций пшеницы представлены и другими исследователями (Prasad et al., 2000; Bossolini et al., 2006; Zeshan et al., 2016). Исходя из рассмотренных данных по оценке эффективности апробированных маркеров, можно отметить информативность праймеров barc-213, barc-1021 и barc-174, которые выделились наибольшим числом общих аллелей, высокими показателями гетерозиготности и PIC, а также наименьшей частотой основных аллелей. Полученные показатели разнообразия были ожидаемы, так как образцы, включенные в настоящее исследование, представляют различные регионы Плодородного Полумесяца и соседних стран, считающиеся первичным очагом произрастания диких пшениц, а также *T. monococcum*. Эти результаты согласуются с ранее полученными данными (Farouji et al., 2015; Wang et al., 2017).

Значение PIC рассчитывали для каждой страны, оно не зависело от размера выборки (табл. 3). Например, при сравнении исследуемых образцов различного происхождения установлено, что наибольший показатель выявлен среди генотипов из Сирии (PIC = 0.49), представленных шестью образцами; второе по величине значение PIC обнаружено у генотипов из Турции (14 образцов, PIC = 44). В результате анализа полученных данных было также установлено, что образцы из Сирии (3.5 аллелей на локус, $H_E = 0.54$) и Турции (4 аллеля на локус, $H_E = 0.48$) более полиморфны, чем остальные генотипы. Относительно высокое разнообразие среди генотипов сирийского и турецкого происхождения по сравнению с другими регионами отмечалось и в других исследованиях (Hammer et al., 2000; Moghaddam et al., 2000; Wang et al., 2017). Для остальных изученных регионов выявлены средние значения PIC, находящиеся в пределах 0.26–0.399. Наименьший показатель разнообразия (PIC = 0.26) установлен для четырех образцов из Ирака. Следует отметить, что все эти генотипы были из одного и того же региона Нинава. Несмотря на большой объем выборки и широкий диапазон мест сбора, генотипы из Ирана (19 образцов) демонстрируют относительно узкое генетическое разнообразие (PIC = 0.33), что указывает на сходный генетический фон этих образцов. Для генотипов, происходящих из Азербайджана, несмотря на небольшое число изученного материала (пять образцов), значение PIC составило 0.399. Различия между образцами иранского происхождения в основном были связаны с первыми тремя локусами, а остальные восемь локусов показали высокую долю общих аллелей. Генетическое разнообразие популяций диплоидной пшеницы из Ирана изучали многие исследователи. Полученные

**Table 3.** Genetic diversity parameters within particular geographic regions

| Subsets | Sample size | Number of alleles | Major allele frequency | $H_O$ | $H_E$ | PIC |
|---|---|---|---|---|---|---|
| *T. boeoticum* | 63 | 7.5 | 0.60 | 0.13 | 052 | 0.49 |
| SYR | 6 | 3.5 | 0.60 | 0.20 | 0.54 | 0.49 |
| TUR | 14 | 4 | 0.62 | 0.13 | 0.48 | 0.44 |
| IRN | 19 | 3.4 | 0.71 | 0.11 | 0.37 | 0.33 |
| LEB | 4 | 2.3 | 0.64 | 0.16 | 0.40 | 0.35 |
| ARM | 6 | 2.7 | 0.67 | 0.10 | 0.44 | 0.395 |
| BGR | 2 | 1.5 | 0.70 | 0.0 | 0.31 | 0.27 |
| IRG | 4 | 1.9 | 0.76 | 0.12 | 0.32 | 0.26 |
| AZE | 5 | 2.2 | 0.63 | 0.0 | 0.45 | 0.399 |

$H_O$, observed heterozygosity; $H_E$, expected heterozygosity; PIC – polymorphism information content.

в нашей работе результаты не согласуются с данными других авторов, обнаруживших более высокий уровень генетического разнообразия среди диплоидной пшеницы *T. boeoticum* иранского происхождения. Так, M.R. Naghavi с коллегами (2009) с использованием SSR-маркеров выявили высокий уровень полиморфизма у видов *T. boeoticum* (PIC = 0.81), собранных из разных агросистем, а также с помощью RAPD-, AFLP- и SSR-маркеров установили, что наиболее высокий генетический полиморфизм наблюдается среди популяций *T. boeoticum* из Ирана.

**Генетические взаимосвязи.** Как известно, эффективность гибридизации зависит от выбора генетически различимых генотипов (Burkhamer et al., 1998; Bohn et al.,1999). В связи с этим нами предприняты усилия для прогнозирования наиболее различимых образцов путем определения степени сходства или расстояния между ними.

В анализируемой коллекции собраны образцы диплоидной пшеницы *T. boeoticum* различного происхождения: отечественные и зарубежные генотипы. На основании данных о генетическом полиморфизме различных сортов пшеницы была построена дендрограмма (см. рис. 1), отражающая сходство изучаемых генотипов. Индекс генетического расстояния среди генотипов *T. boeoticum* варьировал от 0 до 1, что указывает на высокий уровень генетических различий в изучаемой коллекции. Было также выявлено 100 % сходство между образцами под номером 164 (Саудовская Аравия) и 165 (Ливан). Образцы под номерами 21 (Ирак) и 56 (Сирия), а также 22 (Иран) и 57 (Сирия) оказались самыми отдаленными, у них индекс генетического расстояния составил 0.94 и 1 соответственно. Генотипы, содержащие различные комбинации аллелей, могут послужить ценным источником для будущих селекционных программ, поскольку чем больше различия между родительскими формами, тем больше число желаемых аллелей (Ghaderi et al., 1984). Выявлена некоторая дифференциация генотипов в зависимости от географического региона. В частности, в кластере 2 локализовались исключительно образцы из Турции, а в кластере 5 превалировали генотипы иранского происхождения. Несмотря на малую выборку, генотипы из Азербайджана объединились в кластере 3. Однако в большинстве случаев можно наблюдать локализацию изученных образцов одинакового происхождения в разных кластерах.

Если сопоставить данные кластерного анализа и метода главных компонент, можно выявить некоторые сходства и различия групп образцов.

Группа A содержит 15 генотипов, которые главным образом происходили из Азербайджана и Армении (см. рис. 2). Схожую группировку можно отметить на дендрограмме кластерного анализа (см. рис. 1), где четыре из пяти и четыре из шести генотипов из Азербайджана и Армении соответственно объединились в кластере 3. В группе B превалируют генотипы из Турции и соседних с ней регионов. На дендрограмме кластерного анализа кластер 2 представлен также исключительно генотипами из Турции. Группировка генотипов под номерами 6 (Турция) и 42 (Турция), 12 (Турция) и 2 (Азербайджан) указывает на наличие общих аллелей (см. рис. 2). В группе C, аналогично кластерному анализу, количественно преобладают образцы из Ирана, в ней также присутствуют генотипы из Турции и Сирии. Связь между географическим происхождением и генетической организацией для образцов из Ирана (см. рис. 2) может быть обусловлена общим генетическим фоном среди гермоплазмы, ограничением генетического потока, что также подтверждено сравнительно низким показателем PIC. Схожие результаты ранее были получены на других культурах (Izzatullayeva et al., 2014). В самостоятельную группу D обособился генотип под номером 21 из Ирана.

При обобщении полученных данных PCoA и кластерного анализа можно сделать заключение о слабой генетической дифференциации образцов изученной коллекции. Корреляция генетического расстояния с географическим происхождением выявлена лишь для образцов диплоидной пшеницы *T. boeoticum* из Ирана. В ранних исследованиях показано отсутствие корреляции между генетическим расстоянием и географическим происхождением у популяций *T. boeoticum* из Ирана (Ovesna et al., 2002; Malaki et al., 2006).

Таким образом, анализ нашей коллекции подтверждает высокое разнообразие *T. boeoticum* по микросателлитным локусам, эти маркеры вполне могут быть использованы

для оценки генетического полиморфизма коллекции пшеницы в целом. С помощью этих маркеров в коллекции были установлены уникальные генотипы, несущие редкие аллели по микросателлитным локусам, и, соответственно, эти праймеры могут оказаться перспективными для идентификации и генетической паспортизации образцов *T. boeoticum*. Знание генетического разнообразия образцов дикой однозернянки, сохраняемых в коллекции, необходимо для их использования при проведении научных исследований. Полученные нами результаты дают дополнительную информацию о генетической структуре и разнообразии изученных образцов *T. boeoticum*.

## Conflict of interest

The authors declare no conflict of interest.

## References

Abouzied H.M., Eldemery S.M.M., Abdellatif K.F. SSR-based genetic diversity assessement in tetraploid and hexaploid wheat populations. Br. Biotechnol. J. 2013;3:390-404.

Aliyev R.T., Abbasov M.A., Mammadov A.C. Genetic identification of diploid and tetraploid wheat species with RAPD markers. Turk. J. Biol. 2007;31(3):173-180.

Anker C.C., Niks R.E. Prehaustorial resistance to the wheat leaf rust fungus, *Puccinia triticina*, in *Triticum monococcum* (s.s.). Euphytica. 2001;117:3-12.

Babayeva S., Akparov Z., Abbasov M., Mammadov A., Zaifizadeh M., Street K. Diversity analysis of Central Asia and Caucasian lentil (*Lens culinaris* Medikus) germplasm using SSR fingerprinting. Genet. Res. Crop. Evol. 2009;56:293-298.

Bahrai S., Jaradat A.A., Jaradat A.A. Diversity in seed storage proteins of *T. boeoticum* and *T. urartu*. Triticeae III. Proc. 3rd Int. Triticeae Symp. Aleppo, Syria, 4–8 May. 1998:237-243.

Bohn M., Utz H.F., Melchinger A.E. Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs and SSRs and their use for predicting progeny variance. Crop Sci. 1999;39: 228-237.

Bossolini E., Krattinger S.G., Keller B. Development of simple sequence repeat markers specific for the *Lr*34 resistance region of wheat using sequence information from rice and *Aegilops tauschii*. Theor. Appl. Genet. 2006;113:1049-1062.

Burkhamer R.L., Lanning S.P., Martens R.J., Martin J.N., Talbert L.E. Predicting progeny variance from parental divergence in hard red spring wheat. Crop Sci. 1998;38:234-248.

Chen H.B., Martin J.M., Lavin M., Talbert L.E. Genetic diversity in hard red spring wheat based on sequence-tagged-site PCR markers. Crop Sci. 1994;34:1629-1632.

Cifci E.A., Yagdi K. Study of genetic diversity in wheat (*Triticum aestivum*) varieties using random amplified polymorphic DNA (RAPD) analysis. Turk. J. Field Crops. 2012;17:91-95.

Cox T.S., Lookhart G.L., Walker D.E., Harrell L.G., Albers L.D., Rodgers D.M. Genetic relationships among hard red winter wheat cultivars as evaluated by pedigree analysis and gliadin polyacrylamide-gel electrophoretic patterns. Crop Sci. 1985;25:1058-1063.

Dreisigacker S., Zhang P., Warburton M.L., Van Ginkel M., Hoisington D., Bohn M., Melchinger A.E. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments. Crop Sci. 2004;44(2):381-388.

Drikvand R., Bihamta M.R., Najafian G., Ebrahimi A. Investigation of genetic diversity among bread wheat cultivars (*Triticum aestivum* L.) using SSR markers. J. Agric. Sci. 2012;5(1):122.

Dvorak J., Terlizzi P.D., Zhang H.B., Resta P. The evolution of polyploid wheats: identification of the A genome donor species. Genome. 1993;36:21-31.

Farouji A., Khodayari H., Saeidi H., Rahiminejad M.R. Genetic diversity of diploid Triticum species in Iran assessed using inter-retroelement amplified polymorphisms (IRAP) markers. Biologia. 2015; 1:52-60.

Figliuolo G., Perrino P. Genetic diversity and intra-specific phylogeny of *Triticum turgidum* L. subsp. dicoccon (Schrank) Thell. revealed by RFLPs and SSRs. Genet. Resour. Crop Evol. 2004;51:519-527.

Gascuel O. Concerning the NJ algorithm and its unweighted version, UNJ. Mathematical Hierarchies and Biology. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. Providence: Am. Math. Soc. 1997;37:149-170.

Ghaderi A., Adams M.W., Nassib A.M. Relationship between genetic distance and heterosis for yield and morphological traits in dry edible bean and faba bean. Crop Sci. 1984;24:37-24.

Hajiyev E.S., Akparov Z.I., Aliyev R.T., Saidova S.V., Izzatullayeva V.I., Babayeva S.M., Abbasov M.A. Genetic polymorphism of durum wheat (*Triticum durum* Desf.) accessions of Azerbaijan. Russ. J. Genet. 2015;51:863-870.

Hammer K., Filatenko A.A., Korzun V. Microsatellite markers – a new tool for distinguishing diploid wheat species. Genet. Resour. Crop Evol. 2000;47(5):497-505.

Harjit-Singh X., Dhaliwal H.S., Yifru-Teklu Y., Singh H. Germplasm enhancement through wide hybridization and molecular breeding. 11th Regional Wheat Workshop Eastern Central and Southern Africa, Addis-Abeba, Ethiopia, September. 2000;18-22.

Heun M., Schafer-Pregl R., Klawan D., Castagna R., Accerbi M., Borghi B., Salamini F. Site of Einkorn wheat domestication identified by DNA fingerprinting. Science. 1997;278:1312-1314.

Izzatullayeva V.İ., Akparov Z.I., Babayeva S.M., Ojaghi J., Abbasov M.A. Efficiency of using RAPD and ISSR markers in evaluation of genetic diversity in sugar beet. Turk. J. Biol. 2014;38:429-438.

Kojima T., Nagaoka T., Noda K., Ogihara Y. Genetic linkage map of ISSR and RAPD markers in Einkorn wheat in relation to that of RFLP markers. Theor. Appl. Genet. 1998;96:37-45.

Konarev V.G., Gavrilyuk I.P., Peneva T.I., Konarev A.V., Khakimova A.G., Migushova E.F. On the nature and origin of wheat genomes with regard to the data on the biochemistry and immunochemistry of grain proteins. Selskokhozyaystvennaya Biologiya = Agricultural Biology. 1976;11(5):656-665. (in Russian)

Korzun V., Röder M., Ganal M., Filatenko A., Hammer K. Genetic diversity and evolution of the diploid wheat *T. urartu*, *T. boeoticum* and *T. monococcum* revealed by microsatellite markers. Schr. Genet. Ressour. 1998;8:244-247.

Kumar S., Kumar V., Kumari P., Singh A.K., Singh R. DNA fingerprinting and genetic diversity studies in wheat genotypes using SSR markers. J. Environ. Biol. 2016;37(2):319.

Liu K., Muse S.V. PowerMarker: Integrated analysis enviornment for genetic marker data. Bioinformatic. 2005;21:2128-2129.

Malaki M., Naghavi M.R., Alizadeh H., Potki P., Kazemi M., Pirseyedi S.M., Mardi M., Tabatabaei F. Study of genetic variation in wild diploid wheat (*Triticum boeoticum*) from Iran using AFLP markers. Iran. J. Biotech. 2006;4:269-274.

Malik R., Tiwari R., Arora A., Kumar P., Sheoran S., Sharma P., Singh R., Tiwari V., Sharma I. Genotypic characterization of elite Indian wheat genotypes using molecular markers and their pedigree analysis. Aust. J. Crop Sci. 2013;7:561-567.

McLauchlan A., Henry R.J., Issac P.G., Edwards K.J. Microsatellite analysis in cultivated hexaploid wheat and wild wheat relatives. Ed. R.J. Henry. Plant Genotyping: The DNA Fingerprinting of Plants. Wallingford, UK: CABI Publishing, CAB International, 2001;147-159.

Medini M., Hamza S., Rebai A., Baum M. Analysis of genetic diversity in Tunisian durum wheat cultivars and related wild species by SSR and AFLP markers. Genet. Resour. Crop Evol. 2005;52:21-31.

Mizumoto K., Hirosawa S., Nakamura C., Takumi S. Nuclear and chloroplast genome genetic diversity in the wild einkorn wheat, *Triticum urartu*, revealed by AFLP and SSLP analyses. Hereditas. 2002;137: 208-214.

Moghaddam M., Ehdaie B., Waines J.G. Genetic diversity in populations of wild diploid wheat *Triticum urartu* Tum. ex. Gandil. revealed by isozyme markers. Genet. Res. Crop Evol. 2000;47(3):323-334.

Motawei M.I., Al-Doss A.A., Moustafa K.A. Genetic diversity among selected wheat lines differing in heat tolerance using molecular markers. J. Food Agr. Environ. 2007;5:180-183.

Mousavifard S.S., Saeidi H., Rahiminejad M.R., Shamsadini M. Molecular analysis of diversity of diploid Triticum species in Iran using ISSR markers. Genet. Resour. Crop Evol. 2014. DOI 10.1007/s10722-014-0160-z.

Naghavi M., Ebrahimi A., Sabokdast M., Mardi M. Assessment of genetic variation among five hordeum species from Iran. Cer. Res. Commun. 2011;39(4):487-496. http://www.jstor.org/stable/23792314.

Naghavi M.R., Malaki M., Alizadeh H., Pirseiedi M., Mardi M. An assessment of genetic diversity in wild diploid wheat *Triticum boeoticum* from west of Iran using RAPD, AFLP and SSR markers. J. Agr. Sci. Tech. 2009;11:585-598.

Naghavi M.R., Mardi M., Ramshini H.A., Fazelinasab B. Comparative analyses of the genetic diversity among bread wheat genotypes based on RAPD and SSR markers. Iran. J. Biotechnol. 2004;2:195-202.

Ovesná J., Kučera L., Bocková R., Holubec V. Characterisation of powdery mildew resistance donors within *Triticum boeoticum* accessions using RAPDs. Czech. J. Genet. Plant Breed. 2002;38:117-124.

Perrier X., Jacquemoud-Collet J.P. DARwin software http://darwin.cirad.fr/darwin. 2006.

Pestsova E., Ganal M.W., Roder M.S. Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. Genome. 2000;43:689-697.

Prasad M., Varshney R.K., Roy J.K., Balyan H.S., Gupta P.K. The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theor. Appl. Genet. 2000;100:584-592.

Röder M.S., Korzun V., Wendehake K., Plaschke J., Tixier M.H., Leroy P., Ganal M.W. A microsatellite map of wheat. Genetics. 1998;149:2007-2023.

Song Q.J., Fickus E.W., Cregan P.B. Characterization of trinucleotide SSR motifs in wheat. Theor. Appl. Genet. 2002;104:286-293.

Song Q.J., Shi J.R., Singh S., Fickus E.W., Costa J.M., Lewis J., Gill B.S., Ward R., Cregan P.B. Development and mapping of microsatellite (SSR) markers in wheat. Theor. Appl. Genet. 2005;110:550-560.

Španić V., Buerstmayr H., Drezner G. Assessment of genetic diversity of wheat genotypes using microsatellite markers. Period. Biol. 2012;114(1):37-42.

Sud S., Bains N.S., Nanda G.S. Genetic relationships among wheat genotypes, as revealed by microsatellite markers and pedigree analysis. J. Appl. Genet. 2005;46:375-379.

Takumi S., Nasuda S., Liu Y.G., Tsunewaki K. Wheat phylogeny determined by RFLP analysis of nuclear DNA. Einkorn Wheat. Jpn. J. Genet. 1993;68:73-79.

Wang X., Luo G., Yang W., Li Y., Sun J., Zhan K., Liu D., Zhang A. Genetic diversity, population structure and marker-trait associations for agronomic and grain traits in wild diploid wheat *Triticum urartu*. BMC Plant. Biol. 2017;17:112.

Zeb B., Khan I.A., Ali S., Bacha S., Mumtaz S., Swati Z.A. Study on genetic diversity in Pakistani wheat varieties using simple sequence repeat (SSR) markers. Afr. J. Biotechnol. 2009;8:4016-4019.

Zeshan A., Afzal M., Alghamdi S., Kettener K., Mubashar A., Mubushar M., Shakeel A. Evaluation of genetic diversity amongthe Pakistani wheat (*Triticum aestivum* L.) lines through random molecular markers. Braz. Arch. Biol. Technol. 2016;59:e16160282.

Zhang D., Bai G., Zhu C., Yu J., Carver B.F. Genetic diversity, population structure, and linkage disequilibrium in U.S. elite winter wheat. Plant Genome. 2010;3:117-127.

Zohary D., Hopf M. Domestication of plants in the Old World: the Origin and Spread of Cultivated Plants in West Asia, Europe, and the Nile Valley. 2nd ed. Oxford: New York: Clarendon Press, 1993.

Вавиловский журнал генетики и селекции. 2018;22(5):524-530
DOI 10.18699/VJ18.390

# Сравнительный анализ полных последовательностей пластидных геномов чеснока *Allium sativum* и лука репчатого *Allium cepa*

М.А. Филюшин ✉, А.М. Мазур, А.В. Щенникова, Е.З. Кочиева

Федеральный исследовательский центр «Фундаментальные основы биотехнологии» Российской академии наук, Институт биоинженерии, Москва, Россия

Секвенирование и сравнительная характеристика пластидных геномов растений – актуальный инструмент для современных филогенетических и таксономических исследований. Род *Allium* L. (сем. Amaryllidaceae) объединяет более 900 видов растений, включая такие экономически значимые овощные культуры, как чеснок *Allium sativum*, лук репчатый *Allium cepa*, лук-порей *A. porrum* и др. В настоящей работе впервые определена полная нуклеотидная последовательность пластидного генома чеснока *A. sativum*. Пластом *A. sativum* имел размер 153172 п. н. и состоял из большой уникальной (LSC, 82035 п. н.) и малой уникальной (SSC, 18015 п. н.) копий, разделенных инвертированными повторами (IRa и IRb) размером 26561 п. н. каждый. В пластидном геноме чеснока аннотировано 134 гена, из них 82 белок-кодирующих, 38 тРНК, 8 рРНК и 6 псевдогенов. Сравнительный анализ пластидных геномов *A. sativum* и *A. cepa* выявил различия в размерах структурных элементов и спейсеров на границах инвертированных повторов. Общее число генов в пластомах *A. sativum* и *A. cepa* было одинаковым, однако генный состав различался: ген *rpl22* был функциональным у *A. sativum* и псевдогеном у *A. cepa*, а ген *rps16* – наоборот, псевдогеном у *A. sativum* и белок-кодирующим у *A. cepa*. В последовательностях пластидных геномов чеснока *A. sativum* и лука репчатого *A. cepa* были идентифицированы 32 микросателлитные последовательности, больше половины которых – динуклеотидные, остальные – тетра-, пента- и гексануклеотидные, а тринуклеотидные отсутствовали. Сравниваемые пластидные геномы отличались числом повторов ряда микросателлитов, а некоторые микросателлиты встречались только у одного из видов.

Ключевые слова: пластидный геном; чеснок; *Allium sativum*; *Allium cepa*.

**КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:**
Филюшин М.А., Мазур А.М., Щенникова А.В., Кочиева Е.З. Сравнительный анализ полных последовательностей пластидных геномов чеснока *Allium sativum* и лука репчатого *Allium cepa*. Вавиловский журнал генетики и селекции. 2018;22(5):524-530. DOI 10.18699/VJ18.390

**HOW TO CITE THIS ARTICLE:**
Filyushin M.A., Mazur A.M., Shchennikova A.V., Kochieva E.Z. Comparative analysis of the complete plastomes of garlic *Allium sativum* and bulb onion *Allium cepa*. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):524-530. DOI 10.18699/VJ18.390 (in Russian)

Received 20.03.2018
Accepted for publication 03.05.2018


✉ e-mail: michel7753@mail.ru

# Comparative analysis of the complete plastomes of garlic *Allium sativum* and bulb onion *Allium cepa*

M.A. Filyushin ✉, A.M. Mazur, A.V. Shchennikova, E.Z. Kochieva

Federal Research Centre "Fundamentals of Biotechnology", RAS, Institute of Bioengineering, Moscow, Russia

Sequencing and comparative characterization of plant plastid genomes, or plastomes, is an important tool for modern phylogenetic and taxonomic studies, as well as for understanding the plastome evolution. The genus *Allium* L. (family Amaryllidaceae) incorporates more than 900 species, includes economically significant vegetable crops such as garlic *A. sativum*, onion *A. cepa*, leek *A. porrum*, etc. In this work, the plastome of garlic *A. sativum* has been completely sequenced. The *A. sativum* plastome is 153172 bp in size. It consists of a large unique (LSC, 82035 bp) and small unique (SSC, 18015 bp) copies, separated by inverted repeats (IRa and IRb) of 26561 bp each. In the garlic plastome, 134 genes have been annotated: 82 protein-coding genes, 38 tRNA genes, 8 rRNA genes, and 6 pseudogenes. Comparative analysis of *A. sativum* and *A. cepa* plastomes reveals differences in the sizes of structural elements and spacers at the inverted repeat boundaries. The total numbers of genes in *A. sativum* and *A. cepa* are the same, but the gene composition is different: the *rpl22* gene is functional in *A. sativum*, being a pseudogene in *A. cepa*; conversely, the *rps16* gene is a pseudogene in *A. sativum* and a protein-coding gene in *A. cepa*. In the *A. sativum* and *A. cepa* plastomes, 32 SSR sequences have been identified. More than half of them are dinucleotides, and the remaining are tetra-, penta-, and hexanucleotides at the same time, trinucleotides were absent. The compared plastomes differ in the numbers of certain SSRs, and some are present in only one of the species.

Key word: plastid genome; garlic; *Allium sativum*; *Allium cepa*.

Крупнейший род однодольных растений *Allium* L. (сем. Amaryllidaceae, Asparagales) объединяет более 900 видов (Seregin et al., 2015), произрастающих преимущественно в Северном полушарии. К этому роду относятся овощные культуры, например чеснок, лук репчатый, лук-порей и другие, а ряд видов используется в качестве декоративных. Многие виды луков являются редкими и эндемичными (Friesen et al., 2006). Чеснок *Allium sativum* и лук репчатый *A. cepa* – наиболее экономически значимые виды рода, общемировой объем производства которых, по данным Продовольственной и сельскохозяйственной организации ООН (FAO) за 2016 г., составил 26.5 и 57.2 млн т соответственно.

Пищевая ценность как чеснока, так и лука репчатого определяется, прежде всего, углеводным составом листьев и луковиц, а также наличием специфических вторичных метаболитов (Jones et al., 2004). Образование первичных углеводов происходит в хлоропластах в процессе фотосинтеза. Хотя фотосинтез признается ключевой функцией пластид, они играют также важную роль в других аспектах физиологии и развития растений, включая синтез аминокислот, нуклеотидов, жирных кислот, фитогормонов, витаминов и множества метаболитов (Daniell et al., 2016).

Пластидный геном растений, за исключением нефотосинтезирующих видов (Bellot, Renner, 2016; Ravin et al., 2016) и видов семейства Fabaceae (Sveinsson, Cronk, 2014), имеет консервативную четырехчастную структуру – большая уникальная (large single copy, LSC) и малая уникальная (small single copy, SSC) копии, разделенные инвертированными повторами (inverted repeat, IR), а также сходный порядок генов (Wicke et al., 2011). По данным NCBI, в настоящее время секвенированы полные нуклеотидные последовательности пластидных геномов более чем у 2400 видов высших растений.

Для определения последовательностей пластомов существует три стратегии: 1) секвенирование тотальной ДНК и сборка пластома; 2) выделение и секвенирование хлоропластной ДНК; 3) амплификация с использованием в качестве матрицы тотальной ДНК перекрывающихся фрагментов пластома и их дальнейшее секвенирование и сборка (Dong et al., 2013). Геном представителей рода *Allium* имеет большой размер (Ricroch et al., 2005), что делает проблематичной сборку пластидного генома при секвенировании тотальной ДНК. Второй метод весьма трудозатратный и требует больших количеств растительного материала. С помощью этого метода была получена полная последовательность пластидного генома только у одного представителя рода *Allium* – лука репчатого *A. cepa* (von Kohn et al., 2013). Третьим методом определены последовательности пластома многих видов растений, например кувшинки *Nymphaea alba* (Goremykin et al., 2004), огурца *Cucumis sativus* (Chung et al., 2007), ряски *Lemna minor* (Mardanov et al., 2008) и др. Так как к настоящему времени секвенирована полная последовательность пластидного генома нескольких генотипов *Allium cepa*, то представлялось возможным разработать универсальную систему праймеров для амплификации перекрывающихся участков пластома у других видов *Allium*. Поэтому цель нашей работы – разработка системы праймеров и определение полной нуклеотидной последовательности пластидного генома чеснока *A. sativum* и ее последующее сравнение с пластомом *A. cepa*.

## Материалы и методы

Для секвенирования был взят образец *A. sativum* из Узбекистана (каталожный номер k31) из коллекции Всероссийского института генетических ресурсов растений им. Н.И. Вавилова. Тотальную ДНК выделяли из листьев методом CTAB (Doyle J.J., Doyle J.L., 1987) с дополнительной депротеинизацией фенол-хлороформом (1:1). На основе ранее определенных последовательностей пластома нескольких генотипов лука репчатого *A. cepa* (von Kohn et al., 2013; Kim et al., 2015) было разработано 11 пар праймеров (табл. 1) для амплификации пластидного генома у родственного ему чеснока. Размеры участков варьировали от 6 до 21 т. п. н., их концы перекрывались на 500–1000 п. н. Амплификацию фрагментов проводили с использованием набора LongAmp® Hot Start Taq DNA Polymerase (New England Biolabs, США). Эквимолярная смесь полученных амплификатов, содержащая по 100 нг каждого фрагмента, после пробоподготовки была секвенирована с помощью Illumina HiSeq 1500 Sequencing System (ЗАО Геноаналитика, Москва, Россия). Сборку полученного пула чтений в контиги проводили с помощью Spades v.3.8 (Bankevich et al., 2012). Получено 11 контигов, которые затем были объединены в единую последовательность, в качестве референса использовали пластом *A. cepa* (KF728079). Пробелы были закрыты в программе Bandage (Wick et al., 2015), после этого чтения были сопоставлены с полученным одиночным контигом для обеспечения правильности готовой сборки и определения величины покрытия. Покрытие участков составило от 900x до 8000x, длина контигов варьировала от 6 до 21 т. п. н. Полученный пластидный геном был аннотирован с помощью программ DOGMA (http://dogma.ccbb.utexas.edu) и tRNAscan-SE (Lowe, Eddy, 1997). Координаты всех генов проверены вручную. Физическая карта пластидного генома *A. sativum* была построена с помощью программы OGDRAW (рис. 1) (Lohse et al., 2013). Границы инвертированных повторов были перепроверены секвенированием с праймерами, специфичными к фланкирующим участкам: ycf1-ndhF (5′-ATTGGTATTATTCTGGATACGGC-3′; 5′-GATCTGTTTATTCATCTCTACAG-3′), rpl22-rpl2 (5′-GATGCTCGATAAGGCATAAG-3′; 5′-GTAAGCG TCCTGTAGTAAG-3′), rpl2-psbA (5′-CTGTAGTAAGAG GAGTAGT-3′; 5′-TTGATAGTCAAGGTCGTGT-3′).

Для проведения сравнительного анализа пластидных геномов использована последовательность стерильного генотипа *A. cepa* (KF728079). Последовательность *A. cepa* KF728079 была аннотирована нами *de novo*, так как в исходной аннотации (von Kohn et al., 2013) ряд генов, в частности тРНК, имеют неверные координаты, а, например, в гене *rps16* не определены границы экзонов, в результате чего он был признан псевдогеном, хотя таковым не является. Выравнивание последовательностей, определение GC-состава проводили в программе MEGA 7.0 (Kumar et al., 2016). Поиск микросателлитных последовательностей выполняли в программе SSRIT (Temnykh et al., 2001) со следующими параметрами: для ди- и тринуклеотидов – не

**Table 1.** Primers used for amplification of *A. sativum* plastome regions

| Plastid genome region | Sequences of primers | Annealing temperature, °C | Amplicon size, kb |
|---|---|---|---|
| *rpl2-atpA* | 5′-GGCTATAATTGGAGATACCATT-3′<br>5′-GAAATTAGTAATATTATCCGYGAA-3′ | 54 | 12 |
| *atpA-rpoC1* | 5′-TAATCGTTGACCTCTTGCCAATT-3′<br>5′-TCCCATCGAAGTTCACTATGAAT-3′ | 57 | 10 |
| *rpoC1-psaB* | 5′-TGTAKRGCTTCTTCKATTTCTCG-3′<br>5′-ATTTCCAAGGTTTAGCCAAG-3′ | 53 | 20 |
| *psaB-rbcL* | 5′-CATAAAGATTCCAYTGACC-3′<br>5′-GTTGTCCAKGTACCAGTAGAAG-3′ | 53 | 18 |
| *rbcL-psaJ* | 5′-ATGTCACCMCAAACAGARACT-3′<br>5′-TACTGGGTCATTCATAGCATTAC-3′ | 57 | 10 |
| *psaJ-psbB* | 5′-ACTAAATTCATCRAGTTGTTCC-3′<br>5′-GGAATGCTCCAAATTCSACTT-3′ | 54 | 10 |
| *psbB-ycf2* | 5′-GAATTAGAYCGKGCTACTTT-3′<br>5′-TCTCTGGATTGATCAGARGAT-3′ | 55 | 15 |
| *ycf2-ndhB* | 5′-TTAGACAATGTGTGGTTGKTA-3′<br>5′-ACGAAACCAAGAAATAACCC-3′ | 53 | 8 |
| *ndhB-rrn23* | 5′-AGGGTATCCTGAGCAATTKCAA-3′<br>5′-CAAACCTCCTGGCTGTCTCT-3′ | 58 | 11 |
| *rrn23-ndhF* | 5′-TGATCTATCCATGACCAGG-3′<br>5′-GATGAAATTCTTAATGATAGTTGG-3′ | 54 | 6 |
| *ndhF-trnN* | 5′-GTAAATAGATMCGAAACATATAA-3′<br>5′-CTCCCCAAGTAGGATTCGA-3′ | 52 | 21 |

менее пяти повторов, для тетра-, пента- и гексануклеотидов – не менее трех повторов.

## Результаты и обсуждение

Длина полной нуклеотидной последовательности пластидного генома чеснока *A. sativum* составила 153 172 п. н. (см. рис. 1). Последовательность пластома *A. sativum* была депонирована в базу данных NCBI (KX683282). Краткая аннотация пластидного генома чеснока *A. sativum* опубликована нами ранее (Filyushin et al., 2016). Пластидный геном *A. sativum* состоит из большой уникальной копии (82 035 п. н.) и малой уникальной копии (18 015 п. н.), разделенных инвертированными повторами (IRa и IRb) размером 26 561 п. н. каждый. В пластидном геноме аннотировано 134 гена, из них 82 белок-кодирующих, 38 тРНК, 8 рРНК и 6 псевдогенов (Filyushin et al., 2016).

Белок-кодирующие гены, аннотированные в пластидном геноме *A. sativum*, включали гены большой субъединицы рибосомы (гены *rpl2*, *rpl14*, *rpl16*, *rpl20*, *rpl22*, *rpl23*, *rpl32*, *rpl33* и *rpl36*), малой субъединицы рибосомы (гены *rps3*, *rps4*, *rps7*, *rps8*, *rps11*, *rps12*, *rps14*, *rps15*, *rps18* и *rps19*), фотосистемы I (*psaA*, *psaB*, *psaC*, *psaI* и *psaJ*), фотосистемы II (*psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT* и *psbZ*), субъединиц АТФ-синтазы (*atpA*, *atpB*, *atpE*, *atpF*, *atpH* и *atpI*) и других семейств. Большинство генов находится в пластидном геноме *A. sativum* в одной копии, только 18 генов представлены двумя копиями за счет нахождения в инвертированных повторах: шесть белок-кодирующих генов (*rps19*, *rpl2*, *rpl23*, *ycf2*, *ndhB* и *rps7*), восемь тРНК (*trnR-ACG*, *trnL-CAA*, *trnV-GAC*, *trnH-GUG*, *trnI-CAU*, *trnI-GAU*, *trnA-UGC* и *trnN-GUU*) и все четыре гена рРНК (*rrn4.5*, *rrn5*, *rrn16* и *rrn23*). Шесть генов были признаны псевдогенами, так как либо содержали внутренний стоп-кодон (*rps2*, *rps16*, *infA*, два *ycf15* в IRs), либо были образованы в результате неполной дупликации (*ycf1*). Семнадцать генов содержали интроны, из них пятнадцать – один интрон (*atpF*, *rpoC1*, *trnK-UUU*, *trnG-GCC*, *trnL-UAA*, *trnV-UAC*, *ndhA*; четыре гена в IRs: *rpl2*, *ndhB*, *trnI-GAU* и *trnA-UGC*), а два (*clpP* и *ycf3*) – два интрона (Filyushin et al., 2016).

Сравнение пластидных геномов чеснока *A. sativum* и лука репчатого *A. cepa* (KF728079) выявило различия как в размерах структурных элементов, так и в генном составе (табл. 2). В результате анализа нуклеотидных последовательностей сравниваемых пластомов идентифицировано 2 063 вариабельных сайта (1.33 % от выровненной длины). Пластомы *A. sativum* и *A. cepa* имели одинаковое количество белок-кодирующих генов, однако состав функциональных и псевдогенов был различен. Так, *rpl22* у *A. cepa* – псевдоген, а у *A. sativum* в этом гене не выявлено внутренних стоп-кодонов. В то же время ген *rps16* у *A. cepa* – белок-кодирующий, а у *A. sativum* этот ген считается псевдогеном, так как содержит стоп-кодон во втором экзоне.

Последовательности *rps2*, *infA*, *ycf15* являются псевдогенами и у *A. sativum* и у *A. cepa* (табл. 2). Ген *rps2*, как было показано в результате массивного параллельного секвенирования тотальной ДНК 48 представителей Asparagales (Steele et al., 2012), содержит преждевременный

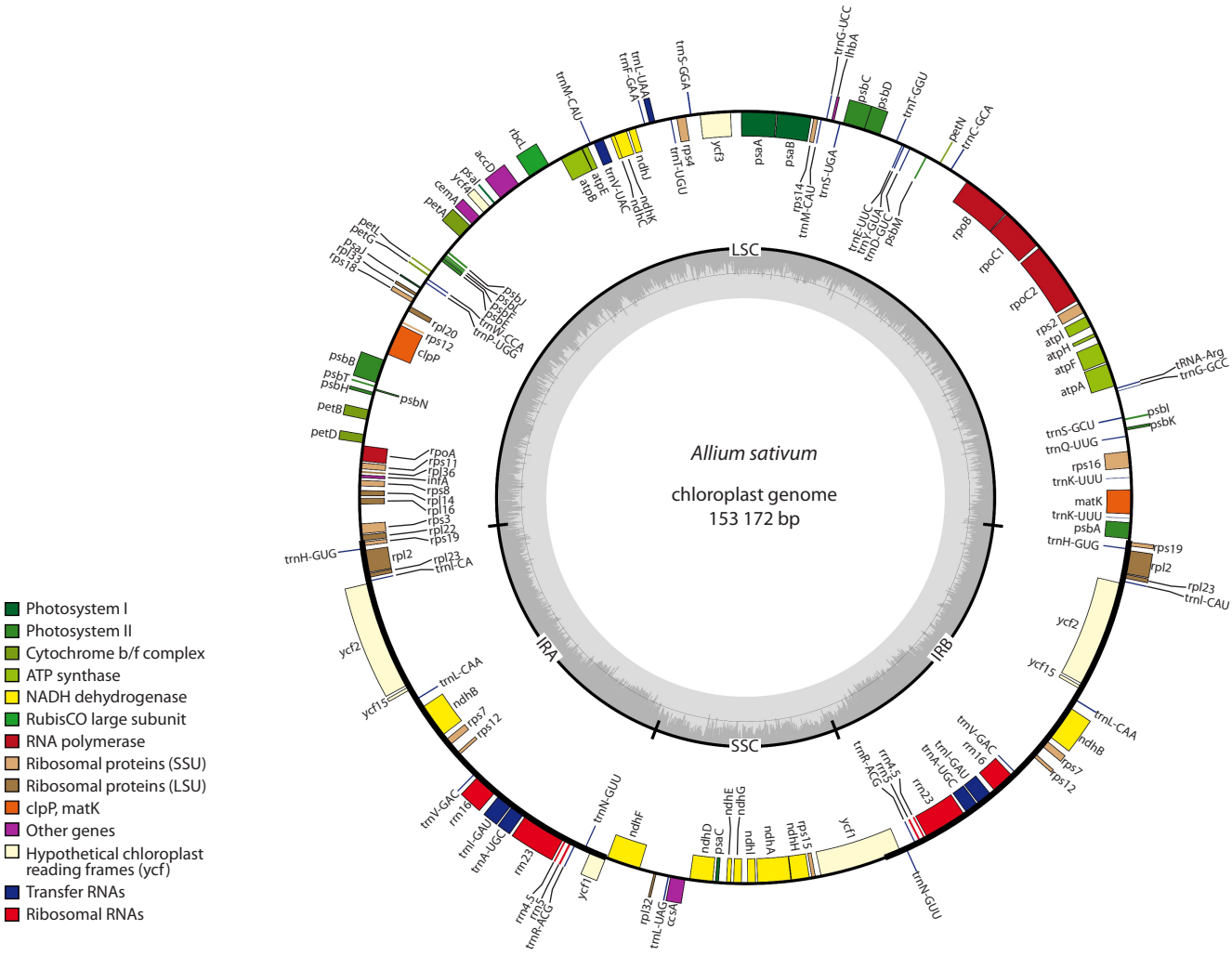



**Fig. 1.** Circular gene map of the *A. sativum* (KX683282) plastid genome.

Genes drawn inside the circle are transcribed clockwise, and those outside, counterclockwise. The darker gray color in the inner circle indicates the GC content.

стоп-кодон лишь у некоторых представителей этого порядка, включая виды *Allium*. Анализ доступных в NCBI последовательностей пластомов показал, что *ycf15* – псевдоген не только у видов *Allium*, но и у видов других родов порядка Asparagales, в то время как *infA* (пседоген у *A. sativum* и *A. cepa*) у видов *A. altaicum* (MH159130; собственные неопубликованные данные) и *A. obliquum* (NC_037199; Filyushin et al., 2018) – белок-кодирующий.

Для функционирования пластид необходимо более 2000 белков, при этом в самом пластидном геноме кодируются не более 90 белков (Синявская и др., 2015). Перенос генетического материала из цитоплазматических геномов в ядерный геном считается характерной чертой эволюции современных растений (Brandvain, Wade, 2009; Sloan, 2015). Для ряда видов растений ранее был показан перенос гена *rps16* в ядерный геном, при этом в пластидном геноме он либо сохраняется, либо отсутствует (полностью, частично или содержит стоп-кодон), а рибосомальный белок S16 транспортируется в пластиды из ядра (Keller et al., 2017). Наличие/отсутствие гена *rps16* у видов одного рода было описано ранее, например, для видов *Arabidopsis*

**Table 2.** Comparison of major structural features of complete plastid genomes of *A. sativum* and *A. cepa*

| Parameter | *A. sativum* (KX683282) | *A. cepa* (KF728079) |
|---|---|---|
| Genome size, bp | 153172 | 153355 |
| LSC | 82035 | 82467 |
| SSC | 18015 | 17918 |
| IRs | 26561 | 26485 |
| GC percentage | 36.7 | 36.8 |
| Total number of genes | 134 | 134 |
| Number | | |
| protein-coding genes | 82 | 82 |
| number of tRNA genes | 38 | 38 |
| number of rRNA genes | 8 | 8 |
| Pseudogenes | 6 (*rps2*, ***rps16***, *infA*, *ycf15* (IRs), *ycf1* (IRa)*) | 6 (*rps2*, ***rpl22***, *infA*, *ycf15* (IRs), *ycf1* (IRa)*) |

* *ycf1* rom IRa is considered a pseudogene because of incomplete duplication.

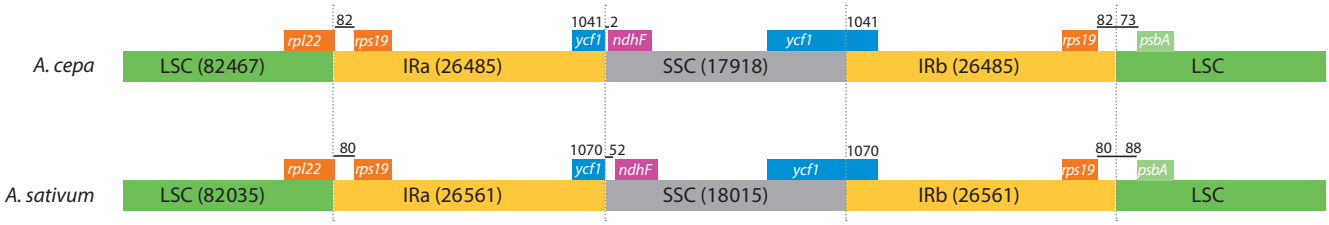



**Fig. 2.** Comparison of the inverted repeat border regions of *A. cepa* (KF728079) and *A. sativum* (KX683282).

**Table 3.** Characterization of microsatellite sequences in plastid genomes of *A. sativum* (KX683282) and *A. cepa* (KF728079)

| Location | Region | Motif | Number of repeats | | Location | Region | Motif | Number of repeats | |
|---|---|---|---|---|---|---|---|---|---|
| | | | *A. sativum* | *A. cepa* | | | | *A. sativum* | *A. cepa* |
| *atpH-atpI* | Intergenic | AT | 5 | 4 | *psaC-ndhE* | Intergenic | TA | 6 | – |
| *rpoC2* | Gene | AT | 5 | 5 | *trnK* | Intron | AATT | 3 | – |
| *trnT-psbD* | Intergenic | AT | 6 | 6 | *trnK-rps16* | Intergenic | ATAA | 3 | 3 |
| *trnS-psbZ* | » | TA | 6 | 6 | *ycf3* | Intron | TTTC | 3 | 3 |
| *rps4-trnT* | » | TA | 5 | 5 | *ndhF-rpl32* | Intergenic | ATTG | – | 3 |
| *rps4-trnT* | » | TA | 6 | 5 | *cemA* | Gene | AATG | 3 | 3 |
| *trnT-trnL* | » | AT | 8 | 5 | *petA-psbJ* | Intergenic | TAAA | 3 | 3 |
| *ycf4-cemA* | » | AT | 5 | 5 | *clpP* | Intron | TAAA | 3 | 3 |
| *cemA* | Gene | TC | 5 | 5 | *rpoC2-rpoC1* | Intergenic | ATTT | 3 | 3 |
| *psaJ-rpl33* | Intergenic | AT | 7 | 8 | *atpF* | Intron | AATA | 3 | 2 |
| *rps18-rpl20* | » | TA | 8 | 5 | *petB-petD* | Intergenic | GGAT | 3 | 3 |
| *ycf2* | Gene | GA | 5 | 5 | *trnT-trnL* | » | TTTTA | 3 | – |
| *ndhF-rpl32* | Intergenic | AT | 9 | – | *accD-psaI* | » | TATAA | 2 | 3 |
| *rpl32-trnL* | » | TA | 5 | – | *trnK* | Intron | TAAAA | – | 3 |
| *rpl32-trnL* | » | TA | 5 | – | *petN-psbM* | Intergenic | ATTGA | 3 | 3 |
| *psaC-ndhE* | » | TA | 7 | 5 | *rpl16-rps3* | » | TTTATT | – | 3 |

(Roy et al., 2010). Можно предположить, что гены рибосомальных белков *rps16* (у *A. sativum*), *rpl22* (у *A. cepa*), *rps2* и ген *infA*, кодирующий фактор инициации трансляции, были перенесены в ядерный геном, так как без этих белков был бы невозможен процесс трансляции в пластидах. Что касается гена *ycf15*, то, во-первых, достоверно не известно о его функциях в пластидах, во-вторых, не известно его происхождение, так как ортологи этого гена не идентифицированы у эубактерий (Shi et al., 2013).

В пластидном геноме *A. sativum* в сравнении с пластомом *A. cepa* было обнаружено семь протяженных делеций, локализованных в межгенных спейсерах *trnC-petN* (221 п. н.), *rps18-rpl20* (62 п. н.), *rpoB-trnC* (56 п. н.), *cemA-petA* (53 п. н.) и *trnT-trnL* (три делеции 38, 23 и 18 п. н.).

Анализ нуклеотидных последовательностей границ инвертированных повторов в пластидных геномах *A. sativum* и *A. cepa* выявил различия в размерах фланкирующих спейсеров (рис. 2). Так, размер спейсера между *ycf1* и *ndhF* у *A. sativum* составляет 52 п. н., а у *A. cepa* – всего 2 п. н. У *A. sativum* и *A. cepa* различаются также длины спейсера LSC-IRb. Следует отметить, что в пластомах *A. sativum* и *A. cepa* гены *rps19* и *trnH* так же, как и у другого представителя порядка Asparagales, *Asparagus officinalis* (Steele et al., 2012), полностью расположены в инвертированных повторах. В то же время у представителей других родов, например *Solanum*, в IRs находится только фрагмент гена *rps19*. Однако присутствие генов *rps19* и *trnH* в IRs не специфично для видов Asparagales, а также встречается у других порядков однодольных, например у Poales (Redwan et al., 2015).

Поиск микросателлитных последовательностей в пластомах *A. sativum* и *A. cepa* выявил всего 32 SSRs, из которых 17 были динуклеотидными (табл. 3). Интересно отметить отсутствие тринуклеотидных микросателлитов. Сравниваемые пластомы отличались числом повторов некоторых микросателлитов, при этом для девяти микросателлитов было показано их отсутствие у одного из исследуемых видов. Большая часть SSRs была локализована в межгенных спейсерах, в то время как в генах и интронах обнаружены только четыре и пять микросателлитов соответственно. У *A. sativum* в межгенных спейсерах *ndhF-rpl32* и *rpl32-trnL*, локализованных в SSC, было идентифицировано три микросателлита ($(AT)_9$ и два повтора $(TA)_5$), которые полностью отсутствуют у *A. cepa*. Нужно отметить, что спейсер *rpl32-trnL* высоко вариабельный у разных групп растений (Shaw et al., 2007), в связи с этим

он наиболее часто используется для филогенетических исследований, в том числе и для рода *Allium* (Wheeler et al., 2013; Herden et al., 2016; Sinitsyna et al., 2016).

Таким образом, впервые определена и охарактеризована полная нуклеотидная последовательность пластидного генома *A. sativum* и проведено ее сравнение с пластомом родственного вида *A. cepa*. Полученную последовательность пластома *A. sativum* можно использовать в дальнейшем для филогенетических исследований рода *Allium*.

## Acknowledgments

This work was supported by the Federal Agency for Scientific Organizations (Russia).

## Conflict of interest

The authors declare no conflict of interest.

## References

Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., Nikolenko S.I., Pham S., Prjibelski A.D., Pyshkin A.V., Sirotkin A.V., Vyahhi N., Tesler G., Alekseyev M.A., Pevzner P.A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 2012;19(5): 455-477. DOI 10.1089/cmb.2012.0021.

Bellot S., Renner S.S. The plastomes of two species in the endoparasite genus *Pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. Genome Biol. Evol. 2016;8(1):189-201. DOI 10.1093/gbe/evv251.

Brandvain Y., Wade M.J. The functional transfer of genes from the mitochondria to the nucleus: the effects of selection, mutation, population size and rate of self-fertilization. Genetics. 2009;182(4):1129-1139. DOI 10.1534/genetics.108.100024.

Chung S.M., Gordon V.S., Staub J.E. Sequencing cucumber (*Cucumis sativus* L.) chloroplast genomes identifies differences between chilling-tolerant and -susceptible cucumber lines. Genome. 2007;50: 215-225.

Daniell H., Lin C.-S., Yu M., Chang W.-J. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 2016;17(1):134. DOI 10.1186/s13059-016-1004-2.

Dong W., Xu C., Cheng T., Lin K., Zhou S. Sequencing angiosperm plastid genomes made easy: a complete set of universal primers and a case study on the phylogeny of saxifragales. Genome Biol. Evol. 2013;5(5):989-997. DOI 10.1093/gbe/evt063.

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

Filyushin M.A., Beletsky A.V., Mazur A.M., Kochieva E.Z. The complete plastid genome sequence of garlic *Allium sativum* L. Mitochondrial DNA Part B: Resources. 2016;1(1):831-832. DOI 10.1080/23802359.2016.1247669.

Filyushin M.A., Beletsky A.V., Mazur A.M., Kochieva E.Z. Characterization of the complete plastid genome of lop-sided onion *Allium obliquum* L. (Amaryllidaceae). Mitochondrial DNA Part B: Resources. 2018;3(1):393-394. DOI 10.1080/23802359.2018.1456369.

Friesen N., Fritsch R.M., Blattner F.R. Phylogeny and new intrageneric classification of *Allium* (Alliaceae) based on nuclear ribosomal DNA ITS sequences. Aliso. 2006;22:372-395.

Goremykin V.V., Hirsch-Ernst K.I., Wolfl S., Hellwig F.H. The chloroplast genome of *Nymphaea alba*: whole-genome analyses and the problem of identifying the most basal angiosperm. Mol. Biol. Evol. 2004;21:1445-1454.

Herden T., Hanelt P., Friesen N. Phylogeny of *Allium* L. subgenus Anguinum (G. Don. ex W.D.J. Koch) N. Friesen (Amaryllidaceae). Mol. Phylogenet. Evol. 2016;95:79-93. DOI 10.1016/j.ympev.2015.11.004.

Jones M.G., Hughes J., Tregova A., Milne J., Tomsett A.B., Collin H.A. Biosynthesis of the flavour precursors of onion and garlic. J. Exp. Bot. 2004;55(404):1903-1918.

Keller J., Rousseau-Gueutin M., Martin G.E., Morice J., Boutte J., Coissac E., Ourari M., Aïnouche M., Salmon A., Cabello-Hurtado F., Aïnouche A. The evolutionary fate of the chloroplast and nuclear rps16 genes as revealed through the sequencing and comparative analyses of four novel legume chloroplast genomes from *Lupinus*. DNA Res. 2017;24(4):343-358. DOI 10.1093/dnares/dsx006.

Kim S., Park J.Y., Yang T. Comparative analysis of the complete chloroplast genome sequences of a normal male-fertile cytoplasm and two different cytoplasms conferring cytoplasmic male sterility in onion (*Allium cepa* L.). J. Hortic. Sci. Biotechnol. 2015;90(4):459-468. DOI 10.1080/14620316.2015.11513210.

Kumar S., Stecher G., Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2016;33(7):1870-1874. DOI 10.1093/molbev/msw054.

Lohse M., Drechsel O., Kahlau S., Bock R. OrganellarGenome DRAW-a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 2013;41:575-581. DOI 10.1093/nar/gkt289.

Lowe T.M., Eddy S.R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955-964.

Mardanov A.V., Ravin N.V., Kuznetsov B.B., Samigullin T.H., Antonov A.S., Kolganova T.V., Skyabin K.G. Complete sequence of the duckweed (*Lemna minor*) chloroplast genome: structural organization and phylogenetic relationships to other angiosperms. J. Mol. Evol. 2008;66(6):555-564. DOI 10.1007/s00239-008-9091-7.

Ravin N.V., Gruzdev E.V., Beletsky A.V., Mazur A.M., Prokhortchouk E.B., Filyushin M.A., Kochieva E.Z., Kadnikov V.V., Mardanov A.V., Skryabin K.G. The loss of photosynthetic pathways in the plastid and nuclear genomes of the non-photosynthetic mycoheterotrophic eudicot *Monotropa hypopitys*. BMC Plant Biol. 2016; 16(Suppl. 3):153-161. DOI 10.1186/s12870-016-0929-7.

Redwan R.M., Saidin A., Kumar S.V. Complete chloroplast genome sequence of MD-2 pineapple and its comparative analysis among nine other plants from the subclass Commelinidae. BMC Plant Biol. 2015;15:196. DOI 10.1186/s12870-015-0587-1.

Ricroch A., Yockteng R., Brown S.C., Nadot S. Evolution of genome size across some cultivated *Allium* species. Genome. 2005;48(3): 511-520.

Roy S., Ueda M., Kadowaki K., Tsutsumi N. Different status of the gene for ribosomal protein S16 in the chloroplast genome during evolution of the genus *Arabidopsis* and closely related species. Genes Genet. Syst. 2010;85(5):319-326.

Seregin A.P., Anačkov G., Friesen N. Molecular and morphological revision of the *Allium saxatile* group (Amaryllidaceae): geographical isolation as the driving force of underestimated speciation. Bot. J. Linn. Soc. 2015;178(1):67-101. DOI 10.1111/boj.12269.

Shaw J., Lickey E.B., Schilling E.E., Small R.L. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 2007;94(3):275-288. DOI 10.3732/ajb.94.3.275.

Shi C., Liu Y., Huang H., Xia E.-H., Zhang H.-B., Gao L.-Z. Contradiction between plastid gene transcription and function due to complex posttranscriptional splicing: an exemplary study of *ycf15* function and evolution in angiosperms. PLoS ONE. 2013;8(3):e59620. DOI 10.1371/journal.pone.0059620.

Sinitsyna T.A., Herden T., Friese N. Dated phylogeny and biogeography of the Eurasian *Allium* section Rhizirideum (Amaryllidaceae). Plant Syst. Evol. 2016;302:1311-1328. DOI 10.1007/s00606-016-1333-3.

Siniauskaya M.G., Danilenko N.G., Lukhanina N.V., Shymkevich A.M., Davydenko O.G. Expression of the chloroplast genome: modern concepts and experimental approaches. Russ. J. Genet.: Appl. Res. 2016;6(5):491-509. DOI 10.1134/S2079059716050117.

Sloan D.B. Using plants to elucidate the mechanisms of cytonuclear co-evolution. New Phytol. 2015;205(3):1040-1046. DOI 10.1111/nph.12835.

Steele P.R., Hertweck K.L., Mayfield D., McKain M.R., Leebens-Mack J., Pires J.C. Quality and quantity of data recovered from massively parallel sequencing: Examples in Asparagales and Poaceae. Am. J. Bot. 2012;99(2):330-348. DOI 10.3732/ajb.1100491.

Sveinsson S., Cronk Q. Evolutionary origin of highly repetitive plastid genomes within the clover genus (*Trifolium*). BMC Evol. Biol. 2014;14:228. DOI 10.1186/s12862-014-0228-6.

Temnykh S., DeClerck G., Lukashova A., Lipovich L., Cartinhour S., McCouch S. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res. 2001; 11(8):1441-1452.

von Kohn C.M., Kielkowska A., Havey M.J. Sequencing and annotation of the chloroplast DNAs of normal (N) male-fertile and male-sterile (S) cytoplasms of onion and single nucleotide polymorphisms distinguishing these cytoplasms. Genome. 2013;56(12):737-742. DOI 10.1139/gen-2013-0182.

Wheeler E.J., Mashayekhi S., McNeal D.W., Columbus J.T., Pires J.C. Molecular systematics of *Allium* subgenus *Amerallium* (Amaryllidaceae) in North America. Am. J. Bot. 2013;100(4):701-711. DOI 10.3732/ajb.1200641.

Wick R.R., Schultz M.B., Zobel J., Holt K.E. Bandage: interactive visualisation of *de novo* genome assemblies. Bioinformatics. 2015; 31(20):3350-3352. DOI 10.1093/bioinformatics/btv383.

Wicke S., Schneeweiss G.M., dePamphilis C.W., Müller K.F., Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol. Biol. 2011; 76(3-5):273-297. DOI 10.1007/s11103-011-9762-4.

Вавиловский журнал генетики и селекции. 2018;22(5):531-535
DOI 10.18699/VJ18.391



# Genetics polymorphism of poplars from Moscow region based on high-throughput sequencing of ITS

E.V. Borkhert[1], G.S. Krasnov[1], N.L. Bolsheva[1], P. Kezimana[1, 2], O.Yu. Yurkevich[1], O.V. Muravenko[1], A.V. Kudryavtseva[1], N.V. Melnikova[1] ✉

[1] Engelhardt Institute of Molecular Biology RAS, Moscow, Russia
[2] Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

Poplars are widely used in landscaping of Moscow due to the ability to effectively purify the air from harmful impurities and to release a large amount of oxygen. The genus *Populus* is characterized by a high level of intraspecies polymorphism, as well as the presence of natural interspecies hybrids. The aim of our work was to evaluate the genetic diversity of poplars, which are growing on the territory of Moscow city by high-throughput sequencing of internal transcribed spacers of 45S rRNA genes (ITS sequences). Sequencing of ITS of 40 poplar plants was performed on Illumina platform (MiSeq) and about 3 000 reads were obtained for each sample in average. Bioinformatics analysis was performed using CLC Genomics Workbench tool. The involved set of poplars had a high level of genetic diversity – the number of single nucleotide polymorphisms (SNPs) detected in each genotype relative to the reference ITS1 and ITS2 sequences of *P. trichocarpa* varying from 4 to 44. We showed that even trees which had been planted on the same territory and, probably, at the same time had significant genetic differences. It can be speculated that highly polymorphic plant material was used for planting poplars in Moscow. For some sites with SNPs, several variants of nucleotides were found in the same individual and the ratio of SNPs was different. We assume that close to 50/50 ratio is observed in interspecific hybrids due to genetic differences in the ITS sequences between maternal and paternal genotypes. For SNPs with a predominance of one of the variants, the presence of paralogues among numerous genomic copies of ITS sequences is more likely. The results of our work can provide a framework for molecular genetic markers application with the purpose of *Populus* species and interspecific hybrids identification, determination the origin of a number of natural hybrids, and monitoring the diversity of genus *Populus* in the Moscow city.

Key words: *Populus*; poplar; Moscow; high-throughput sequencing; ITS; polymorphism; genetic diversity.

# Генетический полиморфизм тополей Московского региона на основе высокопроизводительного секвенирования ITS-последовательностей

Е.В. Борхерт[1], Г.С. Краснов[1], Н.Л. Большева[1], П. Кезимана[1, 2], О.Ю. Юркевич[1], О.В. Муравенко[1], А.В. Кудрявцева[1], Н.В. Мельникова[1] ✉

[1] Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия
[2] Российский университет дружбы народов, Москва, Россия

Тополь широко используется в озеленении Москвы благодаря способности эффективно очищать воздух от вредных примесей и выделять большое количество кислорода. Роду Тополь (*Populus*) свойствен высокий уровень внутривидового полиморфизма, а также наличие естественных межвидовых гибридов. Целью настоящей работы была оценка генетического разнообразия тополей, растущих на территории города Москвы, с использованием высокопроизводительного секвенирования внутренних транскрибируемых спейсеров генов 45S рРНК (ITS-последовательностей). На платформе Illumina (MiSeq) проведено секвенирование ITS-последовательностей 40 растений тополя и в среднем получено около 3000 прочтений для каждого образца. Биоинформатическая обработка данных проведена с использованием программы CLC Genomics Workbench. Исследованная выборка тополей имела высокий уровень генетического разнообразия: число выявленных в каждом генотипе однонуклеотидных полиморфизмов (SNP) относительно референсных последовательностей ITS1 и ITS2 *P. trichocarpa* варьировало от 4 до 44. Показано, что даже деревья, посаженные на одной территории и, вероятно, в одно время, значительно различаются генетически. Можно предположить, что при посадке тополей в Москве использовался крайне полиморфный растительный материал. Для некоторых сайтов с SNP у одного и того же индивидуума выявлено несколько вариантов нуклеотидов, соотношение которых было различным. Мы предполагаем, что соотношение, близкое к 50/50, наблюдается в межвидовых гибридах и является следствием генетических различий в ITS-последовательностях между материнским и отцовским

Received 11.09.2017
Accepted for publication 06.05.2018


✉ e-mail: mnv-4529264@yandex.ru

генотипами. Для SNP с преобладанием одного из вариантов вероятнее наличие паралогов среди многочисленных геномных копий ITS-последовательностей. Результаты работы закладывают основу для применения молекулярно-генетических маркеров с целью идентификации видов и межвидовых гибридов тополя, определения происхождения ряда естественных гибридов, а также мониторинга разнообразия представителей рода *Populus*, растущих на территории города Москвы.

Ключевые слова: *Populus*; тополь; Москва; высокопроизводительное секвенирование; ITS; полиморфизм; генетическое разнообразие.

**HOW TO CITE THIS ARTICLE:**
Borkhert E.V., Krasnov G.S., Bolsheva N.L., Kezimana P., Yurkevich O.Yu., Muravenko O.V., Kudryavtseva A.V., Melnikova N.V. Genetics polymorphism of poplars from Moscow region based on high-throughput sequencing of ITS. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):531-535. DOI 10.18699/VJ18.391

Moscow is one of the largest megalopolises of the world with a developed infrastructure, in which there are more than 12 million inhabitants, that is associated with an unfavorable ecology in the city. To improve the situation, effective landscaping of the city is necessary. Poplar is actively used in the landscaping of Moscow due to the ability to purify the air from pollutants, and release a large amount of oxygen.

Genus *Populus*, according to the Eckenwalder classification (Eckenwalder, 1996), includes 29 species predominantly distributed in the Northern hemisphere. Poplars are dioecious wind-pollinated plants that leads to high intraspecies diversity (Rae et al., 2007). It is known that various species of poplar are easily crossed forming natural interspecific hybrids (Roe et al., 2014; Jiang et al., 2016) that poses difficulties in identifying their taxonomic status. Genome of *P. trichocarpa* was sequenced in 2006 being the first genome of a tree (Tuskan et al., 2006). It is shown that the use of nucleotide sequences of internal transcribed spacers (ITS) of 45S ribosomal RNA (rRNA) genes (Hamzeh, Dayanandan, 2004) is efficient for genetic polymorphism evaluation, taxonomic classification, and determination of phylogenetic relationships in poplars. The ITS region includes highly variable ITS1 and ITS2 sequences located on both sides of highly conserved sequence encoding 5.8S rRNA. ITS sequences, unlike chloroplast and mitochondrial markers, are inherited from both parents and have high variability, while the procedure for their amplification is standardized (Poczai, Hyvonen, 2010). All of the above promotes the active use of ITS sequences for plant barcoding (Li et al., 2011).

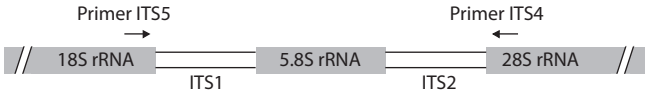
ITS sequences are represented by many copies in a genome and different ITS paralogs may be present in one individual that requires special attention in data analysis and may even hinder obtaining of reliable data by Sanger sequencing (Hollingsworth et al., 2011). High-throughput sequencing can overcome the mentioned above difficulties because hundreds of ITS are sequenced for one individual and sample preparation does not require cloning. In the present work, high-throughput sequencing of ITS was performed and genetic polymorphism of poplars growing on the territory of Moscow city was evaluated.

## Materials and methods

Plant material was collected during the poplar flowering in the south and north of Moscow city. Young leaves were frozen in liquid nitrogen and stored at –70 °C. DNA isolation was performed as described previously (Melnikova et al., 2014). The DNA quality was evaluated by electrophoresis on 1 % agarose gel. DNA concentration was measured on Qubit 2.0 fluorometer (Life Technologies, USA). For further work, a test set of DNA from 40 poplar plants was used.

Two-stage polymerase chain reaction (PCR) was used to prepare DNA libraries for high-throughput sequencing: the first stage included amplification of selected regions of the genome and the addition of universal sequences to the amplicons; at the second stage, the addition of sequences necessary for high-throughput sequencing and dual indexes for sample identification was performed. To amplify the ITS region, we used the primers proposed by Hsiao and White (White et al., 1990; Hsiao et al., 1995) (see Figure) with the universal adapters added. For the second PCR, Nextera XT v2 primers were used (Table 1). Primer design was proceeded according to the recommendations of the Illumina protocol (https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).

PCR for DNA library preparation was performed on Geneamp 9700 (Applied Biosystems, USA). The reaction volume was 25 μl and included 1x Tersus polymerase (Evrogen, Russia), 1x buffer for Tersus polymerase, 0.5 μM of each primer (Evrogen), 200 μM dNTP (Thermo Fisher Scientific, USA), and 40 ng of total DNA. The following reaction conditions were used: 3 min – 95 °C; 10 cycles for the first PCR and 35 cycles for the second PCR: 15 s – 95 °C, 30 s – 58 °C, and 1 min – 72 °C; 3 min – 72 °C. The quality and concentration of 40 DNA libraries were evaluated on Agilent 2100 bioanalyzer (Agilent Technologies) and




Primer localization for amplification of ITS sequences of 45S rRNA gene.

Qubit 2.0 fluorometer (Life Technologies). DNA libraries were sequenced on a MiSeq sequencer (Illumina, USA) using MiSeq Reagent Kit v2 (500 cycles). The reading length was 250 nucleotides from each side for the target sequence; double indexes (Multiplexed Paired-End Run) were also sequenced.

The CLC Genomics Workbench software package (Qiagen, USA) was used for bioinformatics analysis of the data. The reads were mapped to ITS sequence of *P. trichocarpa* (GenBank: AJ006440.1), the genome of which is the reference one for *Populus*. The parameters were as follows: window length – 11, maximum number of gaps and mismatches – 2, minimum average quality of surrounding bases – 15, minimum quality of central base – 20, minimum coverage – 500, minimum paired-end coverage – 0, maximum coverage – 20 000, minimum variant frequency – 20 % or 50 reads.

## Results

We performed high-throughput sequencing of ITS of 40 poplar plants growing on the territory of Moscow city. Sequence length was 250 nucleotides (paired-end reads), and, on average, about 3 000 reads were obtained for each sample. A bioinformatics analysis of the ITS sequences was carried out. The results are presented in Table 2 and Supplementary Materials[1].

The investigated set of trees was characterized by a high level of genetic diversity, the number of detected single nucleotide polymorphisms (SNPs) varied from 4 to 44 relative to ITS sequences of *P. trichocarpa* (GenBank: AJ006440.1). One of the subgroups of trees (numbers 17–28) had been planted in one territory and, probably, at the same time on both sides of the pedestrian road. Table 2 shows even this group of plants to be extremely heterogeneous – the number of detected SNPs varied from 6 to 44.

For some sites with SNPs, more than one nucleotide variant was detected. For these SNPs, in some cases, the ratio of allelic variants was close to 50/50, while in other cases, the distribution was unequal (Supplementary materials). It might be assumed that the 50/50 ratio is observed in hybrids and is a result of genetic differences in ITS sequences between paternal and maternal plants. For SNPs with a significant prevalence of one nucleotide variant over another, polymorphism within numerous copies of ITS in the genome is more likely.

[1] Supplementary Materials are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2018-22/appx8.pdf

**Table 1.** Primer sequences for DNA library preparation

| Primer | Nucleotide sequence |
|---|---|
| ITS5_Illu_F | TCGTCGGCAGCGTCAGATGTGTATAAGAGA CAG*TCGTAACAAGGTTTCCGTAGGTG* |
| ITS4_Illu_R | GTCTCGTGGGCTCGGAGATGTGTATAAGAG ACAG*TCCTCCGCTTATTGATATGC* |
| Nextera XT v2 (i7) | CAAGCAGAAGACGGCATACGAGAT[i7] GTCTCGTGGGCTCGG |
| Nextera XT v2 (i5) | AATGATACGGCGACCACCGAGATCTACAC [i5]TCGTCGGCAGCGTC |

Note: Primer sequences, which are necessary for amplification of ITS, are italicized. i7 and i5 are the sequences for dual indexing of samples.

## Discussion

Poplar is a model object for biological research in trees (Jansson, Douglas, 2007). Over the last decades, numerous approaches have been developed and applied for the analysis of poplar genome, the study of interaction between genotype and environment, and the identification of inter- and intraspecific polymorphism in *Populus* (Jansson et al., 2010; Melnikova et al., 2017). Morphology analysis is actively used in studies of poplars growing on the territory of Moscow city and the Moscow region. High heterogeneity of poplar populations in Moscow and widespread distribution of interspecific hybrids were shown (Kostina, Nasimovich, 2014; Kostina et al., 2017).

In addition to morphological features, the use of molecular markers is effective for plant diversity evaluation (Melnikova et al., 2009–2011; Khadeeva et al., 2011; Bolsheva et al., 2015). In our work, we first applied high-throughput sequencing of ITS to assess the genetic diversity of poplars in Moscow. ITS sequences were already used to study the polymorphism and barcoding of poplar species growing in western China and the number of detected SNPs (38) was high (Feng et al., 2013); that is comparable to the obtained by us data. It should be noted that high-throughput sequencing of ITS performed in the present work allowed us to obtain a much more complete picture of the genetic polymorphism of poplars growing in Moscow by contrast to Sanger sequencing. Thus a high level of genetic diversity of the studied plants was revealed. It can be assumed that such heterogeneous populations of poplars are highly adaptive and have the advantage of surviving in ecologically unfavorable urban conditions. We also showed that while planting poplars in Moscow, an extremely polymorphic plant material was probably used, and even trees planted at the same time in one limited territory were considerably genetically different.

The results of our work lay the foundation for the development of molecular markers for poplars species and interspecific hybrids of poplars growing on the territory of Moscow identification, as well as for determination of the origin of a number of natural hybrids. In addition, recent studies showed that poplar sex is genetically determined

**Table 2.** Single nucleotide polymorphisms of poplars growing in Moscow based on ITS sequence

| No. | Number | | | No. | Number | | |
|---|---|---|---|---|---|---|---|
| | paired-end reads | SNPs* | SNPs with several nucleotide variants | | paired-end reads | SNPs* | SNPs with several nucleotide variants |
| 1 | 4886 | 11 | 0 | 21 | 2530 | 7 | 2 |
| 2 | 4447 | 16 | 11 | 22 | 2169 | 10 | 8 |
| 3 | 3274 | 8 | 3 | 23 | 1606 | 6 | 2 |
| 4 | 4465 | 4 | 0 | 24 | 2740 | 16 | 12 |
| 5 | 7812 | 30 | 13 | 25 | 4463 | 44 | 37 |
| 6 | 5920 | 29 | 13 | 26 | 2586 | 9 | 7 |
| 7 | 1329 | 6 | 2 | 27 | 2619 | 10 | 8 |
| 8 | 1685 | 7 | 5 | 28 | 3086 | 39 | 37 |
| 9 | 4863 | 16 | 2 | 29 | 2727 | 10 | 8 |
| 10 | 321 | 10 | 10 | 30 | 2443 | 10 | 8 |
| 11 | 3171 | 13 | 9 | 31 | 1482 | 8 | 6 |
| 12 | 2597 | 18 | 6 | 32 | 4008 | 39 | 31 |
| 13 | 2829 | 10 | 8 | 33 | 2722 | 25 | 21 |
| 14 | 2740 | 9 | 6 | 34 | 3017 | 9 | 5 |
| 15 | 2567 | 9 | 7 | 35 | 1784 | 9 | 6 |
| 16 | 3908 | 25 | 21 | 36 | 2916 | 18 | 15 |
| 17 | 6592 | 30 | 14 | 37 | 2944 | 12 | 10 |
| 18 | 2150 | 8 | 6 | 38 | 1117 | 8 | 6 |
| 19 | 2373 | 9 | 7 | 39 | 1186 | 8 | 6 |
| 20 | 2163 | 10 | 8 | 40 | 1217 | 22 | 20 |

* ITS sequences of *P. trichocarpa* (GenBank: AJ006440.1) were used as references.

and only a small percentage of trees with recombination in the sex-associated genome region could change the sex (Geraldes et al., 2015; Borkhert et al., 2017; McKown et al., 2017). These data open up new opportunities for molecular marker development so as to use in the landscaping only male poplars, which do not produce fluff, while barcoding using ITS will allow evaluation of polymorphism and maintenance the diversity of populations adaptive to unfavorable urban conditions.

## Acknowledgments

The reported study was funded by RFBR and the Moscow City Government, Research Project No. 15-34-70054 mol_a_mos. This work was performed using the equipment of "Genome" center of Engelhardt Institute of Molecular Biology (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

## Conflict of interest

The authors declare no conflict of interest.

## References

Bolsheva N.L., Zelenin A.V., Nosova I.V., Amosova A.V., Samatadze T.E., Yurkevich O.Y., Melnikova N.V., Zelenina D.A., Volkov A.A., Muravenko O.V. The diversity of karyotypes and genomes within section *Syllinum* of the genus *Linum* (Linaceae) revealed by molecular cytogenetic markers and RAPD analysis. PLoS One. 2015;10(4):e0122015. DOI 10.1371/journal.pone.0122015.

Borkhert E.V., Dmitriev A.A., Krasnov G.S., Bolsheva N.L., Kudryavtseva A.V., Melnikova N.V. Sex associated single nucleotide polymorphisms in the poplar genome. Acta Naturae. 2017;Spec. Iss.1(9):61. (in Russian)

Eckenwalder J.E. Systematics and evolution of *Populus*. In: Stettler R.F., Bradshaw H.D., Heilman P.E., Hinckler T.M. (Eds.). Biology of *Populus* and its Implications for Management and Conservation. Ottawa: Can. Government Publ., 1996;7-32.

Feng J., Jiang D., Shang H., Dong M., Wang G., He X., Zhao C., Mao K. Barcoding poplars (*Populus* L.) from western China. PLoS One. 2013;8(8):e71710. DOI 10.1371/journal.pone.0071710.

Geraldes A., Hefer C.A., Capron A., Kolosova N., Martinez-Nunez F., Soolanayakanahally R.Y., Stanton B., Guy R.D., Mansfield S.D., Douglas C.J., Cronk Q.C. Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). Mol. Ecol. 2015; 24(13):3243-3256. DOI 10.1111/mec.13126.

Hamzeh M., Dayanandan S. Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast *trnT-trnF* region and nuclear rDNA. Am. J. Bot. 2004;91(9):1398-1408. DOI 10.3732/ajb.91.9.1398.

Hollingsworth P.M., Graham S.W., Little D.P. Choosing and using a plant DNA barcode. PLoS One. 2011;6(5):e19254. DOI 10.1371/journal.pone.0019254.

Hsiao C., Chatterton N.J., Asay K.H., Jensen K.B. Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences. Genome. 1995;38(2):211-223.

Jansson S., Bhalerao R.P., Groover A.T. Genetics and Genomics of Populus. New York: Springer-Verlag, 2010. DOI 10.1007/978-1-4419-1541-2.

Jansson S., Douglas C.J. *Populus*: a model system for plant biology. Annu. Rev. Plant Biol. 2007;58:435-458. DOI 10.1146/annurev.arplant.58.032806.103956.

Jiang D., Feng J., Dong M., Wu G., Mao K., Liu J. Genetic origin and composition of a natural hybrid poplar *Populus* × *jrtyschensis* from two distantly related species. BMC Plant Biol. 2016;16:89. DOI 10.1186/s12870-016-0776-6.

Khadeeva N.V., Goryunova S.V., Kochumova A.A., Yakovleva Y.Y., Mel'nikova N.V., Zholobova O.O., Korotkov O.I., Kudryavtsev A.M. Genetic monitoring of populations of *Matthiola fragrans* (Bunge) using RAPD and AFLP analysis. Biol. Bull. 2011;38(4):325-331. DOI 10.1134/S1062359011040078.

Kostina M.V., Nasimovich J.A. On the taxonomy of *Populus* L. II. The importance of fruit characters for identification of cultivated and adventive species in the Moscow region. Byulleten' Moskovskogo Obshchestva Ispytateley Prirody = Bulletin of the Moscow Society of Naturalists. Biol. Ser. 2014; 119(5):74-79. (in Russian)

Kostina M.V., Puzyryov A.N., Nasimovich J.A., Parshevnikova M.S. Representatives of the sections *Aigeiros* Duby and *Tacamahaca* Spach (genus *Populus* L., Salicaceae) and their hybrids in cities of central and eastern European Russia. Skvortsovia. 2017;3(3):97-119.

Li D.Z., Gao L.M., Li H.T., Wang H., Ge X.J., Liu J.Q., Chen Z.D., Zhou S.L., Chen S.L., Yang J.B., Fu C.X., Zeng C.X., Yan H.F., Zhu Y.J., Sun Y.S., Chen S.Y., Zhao L., Wang K., Yang T., Duan G.W. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA. 2011;108(49):19641-19646. DOI 10.1073/pnas.1104551108.

McKown A.D., Klapste J., Guy R.D., Soolanayakanahally R.Y., La Mantia J., Porth I., Skyba O., Unda F., Douglas C.J., El-Kassaby Y.A., Hamelin R.C., Mansfield S.D., Cronk Q.C.B. Sexual homomorphism in dioecious trees: extensive tests fail to detect sexual dimorphism in *Populus*. Sci. Rep. 2017;7(1):1831. DOI 10.1038/s41598-017-01893-z.

Melnikova N.V., Borhert E.V., Martynov S.P., Okuneva I.B., Molkanova O.I., Upelniek V.P., Kudryavtsev A.M. Molecular genetic marker-based approaches to the verification of lilac *Syringa vulgaris* L. *in vitro* germplasm collections. Russ. J. Genet. 2009;45(1):85-90. DOI 10.1134/S1022795409010128.

Melnikova N.V., Borkhert E.V., Snezhkina A.V., Kudryavtseva A.V., Dmitriev A.A. Sex-specific response to stress in *Populus*. Front. Plant Sci. 2017;8:1827. DOI 10.3389/fpls.2017.01827.

Melnikova N.V., Konovalov F.A., Kudryavtsev A.M. Long terminal repeat retrotransposon *Jeli* provides multiple genetic markers for common wheat (*Triticum aestivum*). Plant Genet. Resour. 2011;9(2):163-165. DOI 10.1017/S1479262111000487.

Melnikova N.V., Kudryavtseva A.V., Zelenin A.V., Lakunina V.A., Yurkevich O.Y., Speranskaya A.S., Dmitriev A.A., Krinitsina A.A., Belenikin M.S., Uroshlev L.A., Snezhkina A.V., Sadritdinova A.F., Koroban N.V., Amosova A.V., Samatadze T.E., Guzenko E.V., Lemesh V.A., Savilova A.M., Rachinskaia O.A., Kishlyan N.V., Rozhmina T.A., Bolsheva N.L., Muravenko O.V. Retrotransposon-based molecular markers for analysis of genetic diversity within the genus *Linum*. BioMed Res. Int. 2014;2014:231589. DOI 10.1155/2014/231589.

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

Poczai P., Hyvonen J. Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. Mol. Biol. Rep. 2010;37(4):1897-1912. DOI 10.1007/s11033-009-9630-3.

Rae A.M., Street N.R., Rodríguez-Acosta M. *Populus* Trees. In: Kole C. (Ed.). Genome Mapping and Molecular Breeding in Plants. Vol. 7. Forest Trees. Berlin; Heidelberg: Springer-Verlag, 2007;1-28. DOI 10.1007/978-3-540-34541-1_1.

Roe A.D., MacQuarrie C.J., Gros-Louis M.C., Simpson J.D., Lamarche J., Beardmore T., Thompson S.L., Tanguay P., Isabel N. Fitness dynamics within a poplar hybrid zone. II. Impact of exotic sex on native poplars in an urban jungle. Ecol. Evol. 2014;4(10):1876-1889. DOI 10.1002/ece3.1028.

Tuskan G.A., Difazio S., Jansson S., Bohlmann J., Grigoriev I., Hellsten U., Putnam N., Ralph S., Rombauts S., Salamov A., Schein J., Sterck L., Aerts A., Bhalerao R.R., Bhalerao R.P., Blaudez D., Boerjan W., Brun A., Brunner A., Busov V., Campbell M., Carlson J., Chalot M., Chapman J., Chen G.L., Cooper D., Coutinho P.M., Couturier J., Covert S., Cronk Q., Cunningham R., Davis J., Degroeve S., Dejardin A., Depamphilis C., Detter J., Dirks B., Dubchak I., Duplessis S., Ehlting J., Ellis B., Gendler K., Goodstein D., Gribskov M., Grimwood J., Groover A., Gunter L., Hamberger B., Heinze B., Helariutta Y., Henrissat B., Holligan D., Holt R., Huang W., Islam-Faridi N., Jones S., Jones-Rhoades M., Jorgensen R., Joshi C., Kangasjarvi J., Karlsson J., Kelleher C., Kirkpatrick R., Kirst M., Kohler A., Kalluri U., Larimer F., Leebens-Mack J., Leple J.C., Locascio P., Lou Y., Lucas S., Martin F., Montanini B., Napoli C., Nelson D.R., Nelson C., Nieminen K., Nilsson O., Pereda V., Peter G., Philippe R., Pilate G., Poliakov A., Razumovskaya J., Richardson P., Rinaldi C., Ritland K., Rouze P., Ryaboy D., Schmutz J., Schrader J., Segerman B., Shin H., Siddiqui A., Sterky F., Terry A., Tsai C.J., Uberbacher E., Unneberg P., Vahala J., Wall K., Wessler S., Yang G., Yin T., Douglas C., Marra M., Sandberg G., Van de Peer Y., Rokhsar D. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science. 2006;313(5793):1596-1604. DOI 10.1126/science.1128691.

White T.J., Bruns T., Lee S.,Taylor J.W. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M.A., Gelfand D.H., Sninsky J.J., White T.J. (Eds.). PCR Protocols: A Guide to Methods and Applications. New York: Acad. Press, 1990;315-322. DOI 10.1016/B978-0-12-372180-8.50042-1.

Вавиловский журнал генетики и селекции. 2018;22(5):536-543
DOI 10.18699/VJ18.392

# GWAS of a soybean breeding collection from South East and South Kazakhstan for resistance to fungal diseases

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

[1] Institute of Plant Biology and Biotechnology, Almaty, Kazakhstan
[2] Kazakh National Agrarian University, Almaty, Kazakhstan
[3] Kazakh Research Institute of Agriculture and Plant Growing, Almalybak, Almaty Region, Kazakhstan
[4] Research Institute for Biological Safety Problems, Gvardeiskiy (Otar), Zhambyl Region, Kazakhstan

Soybean (*Glycine max* (L.) Merr) is an essential food, feed, and technical culture. In Kazakhstan the area under soybean is increasing every year, helping to solve the problem of protein deficiency in human nutrition and animal feeding. One of the main problems of soybean production is fungal diseases causing yields losses of up to 30 %. Modern genomic studies can be applied to facilitate efficient breeding research for improvement of soybean fungal disease tolerance. Therefore, the objective of this genome-wide association study (GWAS) was analysis of a soybean collection consisting of 182 accessions in relation to fungal diseases in the conditions of South East and South Kazakhstan. Field evaluation of the soybean collection suggested that *Fusarium* spp. and *Cercospora sojina* affected plants in the South region (RIBSP), and *Septoria glycines* – in the South East region (KRIAPP). The major objective of the study was identification of QTL associated with resistance to fusarium root rot (FUS), frogeye leaf spot (FLS), and brown spot (BS). GWAS using 4 442 SNP (single nucleotide polymorphism) markers of Illumina iSelect array allowed for identification of fifteen marker trait associations (MTA) resistant to the three diseases at two different stages of growth. Two QTL both for FUS (chromosomes 13 and 17) and BS (chromosomes 14 and 17) were genetically mapped, including one presumably novel QTL for BS (chromosome 17). Also, five presumably novel QTL for FLS were genetically mapped on chromosomes 2, 7, and 15. The results can be used for improvement of the local breeding projects based on marker-assisted selection approach.

Key words: soybean; fusarium root rot; frogeye leaf spot; brown spot; GWAS; SNP; QTL mapping.

## Полногеномный анализ ассоциаций с устойчивостью к грибным болезням в коллекции сои в условиях Юго-Восточного и Южного Казахстана

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

[1] Институт биологии и биотехнологии растений Комитета науки Министерства образования и науки Республики Казахстан, Алматы, Казахстан
[2] Казахский национальный аграрный университет, Алматы, Казахстан
[3] Казахский научно-исследовательский институт земледелия и растениеводства, пос. Алмалыбак, Алматинская область, Казахстан
[4] Научно-исследовательский институт проблем биологической безопасности Комитета науки Министерства образования и науки Республики Казахстан, пос. Гвардейский, Жамбылская область, Казахстан

Соя (*Glycine max* (L.) Merr) – важная пищевая, кормовая и техническая культура. В Казахстане площадь под соей увеличивается с каждым годом, что обусловлено ее важностью в решении проблемы дефицита белка в питании людей и кормлении животных. Одной из основных проблем производства сои являются грибные болезни, вызывающие потери урожая до 30 %. Для повышения эффективности селекции, направленной на улучшение устойчивости сои к болезням, могут быть использованы современные геномные технологии. Таким образом, целью настоящего исследования был полногеномный анализ ассоциаций (GWAS) в коллекции сои, состоящей из 182 образцов, на устойчивость к грибным болезням в условиях Юго-Восточного и Южного Казахстана. В результате полевой оценки коллекции сои обнаружены растения, пораженные *Fusarium* spp. и *Cercospora sojina* в Южном регионе (НИИПББ) и *Septoria glycines* – в Юго-Восточном регионе (КазНИИЗиР). Исследование было нацелено на идентификацию локусов количественных признаков (ЛКП), связанных с устойчивостью к основным заболеваниям, таким как фузариоз корневой гнили (FUS), церкоспороз (FLS) и септориоз (BS). GWAS с использованием 4442 SNP-маркеров (single nucleotide polymorphism) матрицы Illumina iSelect позволил идентифицировать 15 ассоциаций маркер–признак (MTA) на устойчивость к трем болезням на двух разных стадиях роста. Генетически картированы два ЛКП как для FUS (хромосомы 13 и 17), так и для BS (хромосомы 14 и 17), включая один предположительно новый ЛКП для BS, который был идентифицирован на хромосоме 17. Кроме того, пять предположительно новых ЛКП для FLS были идентифицированы на хромосомах сои 2, 7 и 15. Результаты исследования могут быть использованы для улучшения селекционных программ, в том числе маркер-опосредованной селекции.

Ключевые слова: соя; фузариоз корневой гнили; церкоспороз; септориоз; GWAS; SNP; ЛКП картирование.

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

Received 04.06.2018
Accepted for publication 24.06.2018


e-mail: yerlant@yahoo.com

Soybean (*G. max* (L.) Merrill.) is one of the most important legumes in the world due to its high nutritional value and protein content (Masuda, Goldsmith, 2009). In Kazakhstan, this crop is mainly cultivated in the South East region. According to the Agency for Statistics of the Republic of Kazakhstan, in 2017 soybean was grown in an area of 137.4 thousand hectares (http://www.fcc.kz/attachments/article/4325). For further development of the soybean industry, the Government of Kazakhstan has declared a new initiative to expand the soybean area to 400 thousand hectares by 2020 to ensure its yield at 1 million tons (Zatybekov et al., 2017).

The productivity of soybean largely depends on availability of well-adapted cultivars with approptiate flowering and maturity times to match various ecological environments of the country (Zhang et al., 2007). Our previous study based on evaluation of 120 soybean accessions in three different regions of Kazakhstan (Abugalieva et al., 2016) has confirmed the results of observations in other parts of the world (Contreras-Soto et al., 2018; Copley et al., 2018), which underline the importance of suitable flowering time for plant adaptation in a particular environmental niche.

Another important factor that severely limits the soybean productivity worldwide is susceptibility to harmful diseases (Yang X.B. et al., 2001; Vidic et al., 2013). For instance, 25 known diseases posed a constant threat to the productivity of soybean in the USA (Mueller et al., 2010). In China, out of eight most common diseases, six are caused by fungi. In Russia, reports suggest there are up to 32 soybean diseases (Zaostrovnykh, 2005; Kurilova, 2010; Polozhieva, Dubovitskaya, 2015). In Kazakhstan, more than ten fungal diseases of soybean have been identified (Mombekova et al., 2013; Didorenko et al., 2014), and with expansion of the area under the crop, it is an obvious necessity to study the genetic background associated with the tolerance to harmful pathogens.

The damage caused by various diseases is determined by environmental conditions, the biology and spread of the parasite, and the characteristics of breeding material (Faske et al., 2014). Different parts of the plant, including seeds, sprouts, roots, shoots, leaves, and beans can be severely affected by these diseases. In this respect, all soybean diseases can be separated into three large groups: 1) diseases of seeds, sprouts, and seedlings; 2) patches that affect various parts of the plant; 3) diseases that cause the plants to wilt (Faske et al., 2014). In general, the total yield loss from susceptibility to fungal diseases can reach up to 40 % (Zaostrovnykh, 2005). J.K. Pataky and S.M. Lim (1981) reported that soybean yield loss due to *S. glycines* was associated with reduction of seed weight. M.D. Dias et al. (2016) identified a highly significant correlations ($p < 0.01$) between yield and soybean *Colletotrichum truncatum* incidence on pods ($r = –0.85$). About 90 kg/ha of soybean grain were lost for each 1 % increment in the disease incidence.

Currently, a large number of genes controlling the resistance to various diseases and pests have been identified (Prabhu et al., 1999; Wang J. et al., 2010; Vidic et al., 2013). Several soybean mapping populations were developed for genetic localization of QTL and genes associated with the soybean diseases, such as rhizoctonia root rot (RRR, caused by *Rhizoctonia solani*) (Zhao et al., 2005), fusarium root rot (FUS, caused by *Fusarium* spp.) (Stacey, 2008), phytophthora root rot (PRR, caused by *Phytophthora sojae*) (Zhang et al., 2013), frogeye leaf spot (FLS, caused by *C. sojina*) (Mian et al., 1999) and sclerotinia stem rot (SCL, caused by *Sclerotinia sclerotiorum*) (Zhao et al., 2015). The majority of these studies were based on use of SSR (simple sequence repeat) microsatellite markers. However, with the development of SoySNP50K iSelect SNP (single nucleotide polymorphism) array (Song et al., 2013), most of the modern studies rely on the use of SNP markers, which are crucial for genome-wide association studies (GWAS) (Klein, 2007). GWAS is based on use of whole genome genotyping and a detailed phytopathological and morphological description of collections with a high level of genetic diversity (Klein, 2007). A survey of recent reports has shown successful use of GWAS for studying soybean resistance to fungal diseases (Bao et al., 2015; Iquira et al., 2015; Schneider et al., 2016; Qin et al., 2017).

The purpose of this study was to assess the tolerance of the soybean germplasm collection represented by 182 accessions from major soybean growing regions from all around the world to most harmful diseases spreading in the South and South East of Kazakhstan. The GWAS was applied for identification of marker-trait associations for resistance to FUS, FLS and BS.

## Materials and methods

The analyzed soybean collection consisted of 182 accessions, including 18 released cultivars and prospective breeding lines from Kazakhstan (Zatybekov et al., 2017). The accessions represented 12 countries from 5 geographic regions, including Western and Eastern Europe, North America, East Asia, and Kazakhstan. The collection was tested in the experimental plots of Research Institute of Biological Safety Problems (RIBSP, southern Kazakhstan) and Kazakh Research Institute of Agricultural Plant Production (KRIAPP, south-eastern Kazakhstan) in 2016–2017. Despite the environment similarities of the two localities, the conditions of soybean growth were different, as KRIAPP tested plants in irrigated, and RIBSP – in non-irrigated sites. Plants were grown in 1 meter long rows with a 30-cm distance between adjacent rows and a 5-cm gap between plants within rows. In total, the data for mean values of eight agronomic traits of the 182 soybean accessions harvested in two environments were subjected to further statistical analysis. The eight traits included the following data: days to seedling emergence (VE), days to flowering time (R2), days to development of pods (R4), days to full maturity of seeds (R8), plant height (PH), number of seeds per plant (NSP), thousand seeds weight (TSW), and yield per plant (YP).

Disease resistance analysis was carried out in relation to the three fungal diseases spread in the regions. In the South East the plants were analyzed for resistance to BS (caused by *S. glycines*), while in the South they were tested for resistance to FLS (*C. sojina*) and FUS (caused by a group of unidentified *Fusarium* pathogens indicated in this study as *Fusarium* spp.). The plant resistance to fungal diseases of the leaf surface was characterized based on a nine-point scale, where point 1 stood for highly resistant (no symptoms), 3 – for resistant (5–19 % foliage affected), 5 – for partially resistant (20–49 % of the foliage affected), 7 – for susceptible (50–79 % of the foliage affected), and 9 – for highly susceptible (up to 80 % of the foliage affected) (Hnetkovsky et al., 1996). The plant resistance to root rot was characterized based on a five-point scale,

where point 1 stood for healthy root without infection symptoms, 2 – for slight cortical necrosis or vascular discoloration, 3 – for moderate cortical necrosis or vascular discoloration, 4 – for extensive cortical or vascular tissue distroyed, and 5 – for withering and dying of a plant (Leath, Carroll, 1982). The disease severity after infection as a percentage of the affected area and the healthy part of plants was noted as well. Therefore, the resistance to all three diseases were marked with number 1 (for instance, FUS1), and the diseases severity were marked as number 2 (for instance, FUS2). Statistical analyses of obtained data were calculated by using Statistical Package for the Social Sciences (SPSS 16.0) (https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software) computer programs.

DNA samples were extracted and purified from single seeds of individual cultivars using commercial kits (Qiagene, CA, USA). The DNA concentration for each sample was adjusted to 50 ng/μl. All samples were genotyped using the soybean 5403 SNP Illumina iSelect array (Song et al., 2013) at the Traitgenetics GmbH (Gatersleben, Germany). The Illumina Infinium procedure was performed according to the manufacturer's protocol. SNP genotype analysis was carried out using the Illumina Genome Studio software (GS V2011.1). Population genetic analysis and principal coordinate analysis were performed using GenAlEx 6.5 (Peakall, Smouse, 2012).

The SNP dataset was filtered using a 10 % cutoff for missing data and markers with minor allele frequency >0.10 were considered for GWAS. Numbers of hypothetical groups ranging from $k = 1$ to 10 were assessed using 50,000 burn-in iterations followed by 100,000 recorded Markov-Chain iterations. To estimate the sampling variance of population structure inference, five independent runs were carried out for each $k$ by the STRUCTURE software (Pritchard et al., 2000). The output from STRUCTURE was analyzed for delta K value (ΔK) in STRUCTURE HARVESTER (Evanno et al., 2005). On the basis of the final $k$ values, Q-matrix for three identified clusters was developed. GWAS for resistance to the most harmful fungal diseases of soybean in South East and South Kazakhstan were studied using 4,442 SNP filtered against minor alleles. GWAS based on the MLM model, including options with Q and K matrices, was conducted using the TASSEL 5 software (Bradbury et al., 2007). Pairwise LD between markers was measured using linkage disequilibrium parameter ($r^2$) between alleles using R studio (Wimmer et al., 2012). The LD decay rate was estimated as the chromosomal distance at which the average pairwise correlation coefficient ($r^2$) dropped to a half of its maximum value (Wimmer et al., 2012). The GWAS for resistance to diseases and spread of the diseases during the plant growth was run separately, and as the first evaluation was marked as FUS1, FLS1, and BS1, the evaluations for the spread of the diseases during the plant growth were marked as FUS2, FLS2, and BS2.

## Results

### Diseases resistant

Field trial results at the experimental stations of the South East and South regions suggested a clear difference in the development of FUS, FLS, and BS on the leaf surfaces at the adult stage of the plant growth. While results at the RIBSP (South region) showed the occurrence of FUS and FLS, and luck of BS symptoms, the data at the KRIAPP (South East region) allowed the identification of only BS development (Fig. 1). At the RIBSP site, FUS showed stronger influence on plant growth as only 62.1 % of plants were resistant, whereas 77.4 % was resistant to FLS, and 100 % resistant to BS (see Fig. 1, *a*). At KRIAPP the collection showed 79.3 % resistance to BS with no signs of FUS and FLS throughout the plant growth period (see Fig. 1, *b*).

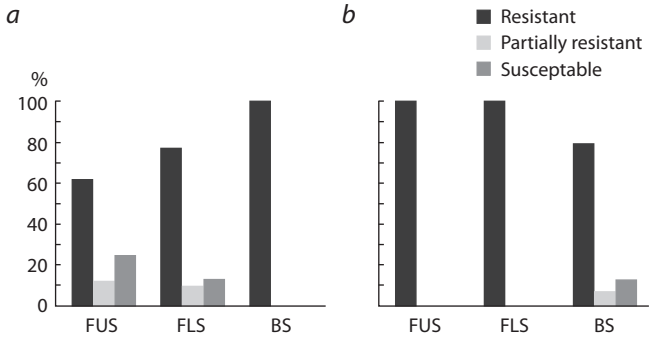



**Fig. 1.** Resistance of the soybean collection to three studied fungal diseases in the South and South East regions of Kazakhstan:

*a* – experimental plot of Research Institute of Biological Safety Problems; *b* – experimental plot of Kazakh Research Institute of Agricultural Plant Production.

Three-way ANOVA suggested that the origin of the accessions was significantly associated with reducing of YP in relationship to the development of FUS and FLS in South Kazakhstan (Table 1). Among tested accessions, there were clear examples of association between the high resistance and yield. For instance, cultivar (cv.) 'Santana' from France was highly resistant to FUS and had high YP (8.7 ± 0.26 g), while cv. 'Chernovickaya 7' from Ukraine was highly susceptible to FUS with the YP of only 1.8 ± 0.15. The collection included ten accessions from East Asia, which showed complete resistance to all three diseases.

### Phenotypic variation of the collection

Comparative assessment of five groups of samples in the studied soybean collection in two regions for 2016–2017 has not revealed sharp differences in the main agronomic traits. The varieties from North America and West Europe showed good potential in NSP, while the varieties from East Asia – the highest potential in TSW. Pearson correlation demonstrated that NSP negatively correlated with TSW ($p < 0.001$) and positively – with YP ($p < 0.001$). Overall highest average yield in the collection of 182 accessions for the two fields was recorded for the Supra variety from Canada (23.0 ± 3.86), followed by Cheremosh (18.7 ± 2.89) from Ukraine and Slaviya (19.3 ± 5.42) from Russia. Among the accessions from Kazakhstan, the Mysula variety showed the best YP (18.3 ± 3.45). It is interesting that in 2016, 14.8 % of the collection showed a higher yield result in comparison with standard variety Zhansaya, and in 2017 the number of higher yield accessions was even bigger (54.4 %).

Multivariate ANOVA suggested that FUS and FLS affected each studied trait both in RIBSP and KRIAPP, except FLS was not a factor for variation in TSW (see Table 1). On the other

**Table 1.** Multivariate ANOVA for the main agronomic traits by SPSS

| Factors | Traits | df | FUS | FLS | BS | Factors | Traits | df | FUS | FLS | BS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Resistance | NSP | 8 | *** | *** | ns | Origin | NSP | 44 | ** | * | ns |
| | TSW | 8 | * | ns | ns | | TSW | 44 | ns | ns | ns |
| | YP | 8 | *** | *** | ns | | YP | 44 | * | * | ns |
| | VER2 | 8 | ** | * | ns | | VER2 | 44 | ns | ns | ns |
| | VER8 | 8 | *** | ** | ** | | VER8 | 44 | ns | * | ns |
| Region | NSP | 17 | ns | ns | ns | Resistance × Region × Origin | NSP | 89 | ns | ns | ns |
| | TSW | 17 | ns | ns | ns | | TSW | 89 | ns | ns | ns |
| | YP | 17 | ns | ** | ns | | YP | 89 | ns | ns | ns |
| | VER2 | 17 | ns | ns | ns | | VER2 | 89 | ns | ns | ns |
| | VER8 | 17 | * | ns | ns | | VER8 | 89 | ns | ns | ns |

Note: df – degree of freedom. The F values are provided with significance level indicated by the asterisks. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns – not significant.

**Table 2.** Mean genetic diversity indexes in five soybean groups based on 4442 SNPs

| Population | East Europe | West Europe | East Asia | North America | Kazakhstan |
|---|---|---|---|---|---|
| N | 83 | 24 | 10 | 32 | 33 |
| Ne | 1.93 ± 0.006 | 1.72 ± 0.005 | 1.66 ± 0.006 | 1.67 ± 0.005 | 1.68 ± 0.005 |
| I | 0.76 ± 0.003 | 0.6 ± 0.003 | 0.55 ± 0.004 | 0.58 ± 0.003 | 0.58 ± 0.004 |
| h | 0.45 ± 0.002 | 0.40 ± 0.002 | 0.36 ± 0.003 | 0.38 ± 0.002 | 0.39 ± 0.002 |

Note: N – number of accessions; Ne – number of effective alleles; I – Shannon index; h – Nei's diversity index.

hand, BS significantly affected the duration of plant growth at the VER8 stage, the only one out of the eight studied traits (see Table 1).

### Genetic variation in the soybean collection based on SNP markers

Genotyping of the soybean collection using the Illumina iSelect SNP array revealed 5403 successful SNPs (74.03 % success) with 77.98 % variants being transitions and 22.01 % – transversions. The final data consisted of 4442 polymorphic SNPs spanned on 20 chromosomes with the average length of 47.4 Mb and the average number of SNPs per chromosome of 222.1. The number of markers per chromosome varied from 163 in Gm11 to 286 in Gm13 with the chromosome length ranging from 37.3 Mb in Gm16 to 62.2 Mb in Gm18. The average density of SNP map was one marker per every 213 Kb. The LD decay curve at the threshold $r^2 = 0.1$ was 20 Kbp (Supplemental Fig. 1)[1]. The PCoA allowed to separate 182 accessions based on their breeding origin and were split into five geographically distinct groups (see Supplemental Fig. 2). The smallest group was from East Asia (10 accessions), and the largest – from Eastern Europe (83 accessions, see Table 2). The PCoA analysis based on NeiP data showed that genotypes from two European groups were positioned separately from other three groups by the PCoA1 component (see Supplemental Fig. 2), while PCoA2 effectively separated the remaining three groups. The accessions from North America and Kazakhstan appeared to be the most close groups, while the accessions from East Asia were genetically more distant from the other four groups (see Supplemental Fig. 2).

### Association mapping

A total of 9 SNPs for 15 MTAs at the two stages of plant growth were identified to be associated with the resistance to three fungal diseases (Table 3). For each MTA separate QQ plots were generated to validate the significance of the associations. In addition, the results were statistically validated using a *t*-test for identification of a false positive MTA. All 15 identified MTAs were significant after *t*-test application. The results suggested that two MTAs were significant for resistance to FUS, five – for FLS, and two – for BS (see Table 3). Also, the physical position of each critical SNP marker was compared with positions of known QTLs (https://soybase.org/search/qtllist_by_symbol.php). In this study only two out of nine SNPs matched the positions of analogous QTL in a soybean genome (see Table 3).

The largest number of SNP markers in identified MTAs were located in chromosome 2, where mainly QTLs for resistance to *P. sojae* were genetically mapped (Fig. 2). The analysis of genome physical locations of associated SNP markers revealed that 3 SNPs were part of CDS (coding DNA sequence) and remained 6 SNPs were located in intergenic regions (Table 4).

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

[1] Supplementary Materials are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2018-22/appx9.pdf

**Table 3.** List of MTAs identified using the TASSEL software and matching with locations of corresponding QTL available in the Soybase database (https://soybase.org)

| Disease stages | SNP ID | Chr | Position | Allele | Allele freq | p-value | Add | $r^2$, % | Suggested QTL ID in this study | Known QTLs* |
|---|---|---|---|---|---|---|---|---|---|---|
| FUS1 | Gm13.20484995 | 13 | 20484995 | G/T | 43/134 | 2.38E–4 | 1.9 | 8.1 | q*Fus*. spp 13-1 | Fusarium lesion length 1-2, Phytoph 9-2 |
| | Gm17.8109237 | 17 | 8109237 | A/C | 134/41 | 7.54E–5 | –2.1 | 9.4 | q*Fus*. spp 17-1 | SCN 23-2 |
| FUS2 | Gm13.20484995 | 13 | 20484995 | G/T | 43/134 | 5.05E–4 | 1.9 | 7.3 | q*Fus*. spp 13-2 | Fusarium lesion length 1-2, Phytoph 9-2 |
| | Gm17.8109237 | 17 | 8109237 | A/C | 134/41 | 1.40E–5 | –2.6 | 11.4 | q*Fus*. spp 17-2 | SCN 23-2 |
| FLS1 | Gm02.9039246 | 2 | 9039246 | C/T | 72/104 | 1.52E–4 | 0.4 | 8.6 | q*Cer.s* 2-1-1 | Bean pyralid 1-2 |
| | Gm02.10140292 | 2 | 10140292 | A/G | 84/87 | 1.93E–4 | –2.6 | 8.5 | q*Cer.s* 2-1-2 | |
| | Gm07.36875730 | 7 | 36875730 | C/T | 48/130 | 8.97E–4 | 1.4 | 6.6 | q*Cer.s* 7-1 | SCN 40-4 |
| | Gm15.1007132 | 15 | 1007132 | A/C | 75/103 | 7.40E–4 | 13.5 | 6.6 | q*Cer.s* 15-1 | Fusarium lesion length 1-3 |
| FLS2 | Gm02.6596937 | 2 | 6596937 | G/T | 38/130 | 8.32E–4 | 1.1 | 6.9 | q*Cer.s* 2-2-1 | Phytoph 14-4, Bean pyralid 1-6 |
| | Gm02.9039246 | 2 | 9039246 | C/T | 72/104 | 8.78E–5 | 6.2 | 9.3 | q*Cer.s* 2-2-2 | Bean pyralid 1-2 |
| | Gm02.10140292 | 2 | 10140292 | A/G | 84/87 | 1.69E–4 | –4.4 | 8.6 | q*Cer.s* 2-2-3 | |
| | Gm07.36875730 | 7 | 36875730 | C/T | 48/130 | 7.69E–4 | 15.2 | 6.8 | q*Cer.s* 7-2 | SCN 40-4 |
| | Gm15.1007132 | 15 | 1007132 | A/C | 75/103 | 9.52E–4 | 13.5 | 6.3 | q*Cer.s* 15-2 | Fusarium lesion length 1-3 |
| BS1 | Gm14.4811528 | 14 | 4811528 | A/C | 87/87 | 7.23E–4 | 3.2 | 6.9 | q*Sep.g*14-1 | SDS 14-10, Sclero 8-2 |
| | Gm17.12684761 | 17 | 12684761 | C/T | 109/62 | 4.19E–4 | –3.5 | 7.7 | q*Sep.g*17-1 | |

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

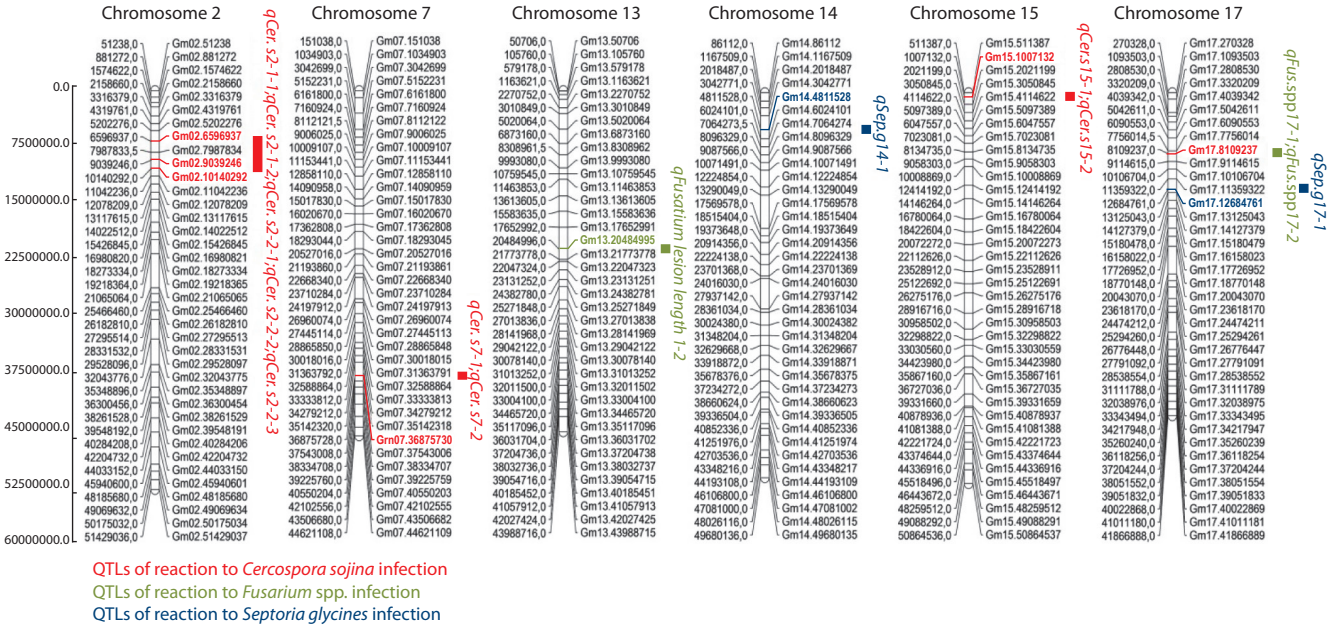



**Fig. 2.** Genetic map of identified SNP markers in identified MTAs for resistance to the three diseases analyzed in soybean population.

## Discussion

The analysis of three diseases from two regions of Kazakhstan revealed strong environmental influence on plant tolerance to studied pathogens, as FUS and FLS were the factor in the South, and BS – in South East parts of the country, respectively. ANOVA suggested that FUS and FLS affected nearly all studied traits, except for TSW, which was a factor in case of FUS, but not in case of FLS (see Table 1). The test also suggested that the origin of the plant material was essential for NSP and YP in both FUS and FLS studies. A different outcome

**Table 4.** Physical positions of identified SNPs in the soybean genome

| Pathogenes of three diseases | SNP ID | Positions in genome | Putative gene | Annotation of putative gene |
|---|---|---|---|---|
| *C. sojina* | Gm02.6596937 | Intron | Glyma02g08400 | Small heat-shock protein (hsp20) family |
| *C. sojina* | Gm02.9039246 | Intron | Glyma02g10920 | Unknown protein |
| *C. sojina* | Gm02.10140292 | CDS | Glyma02g11935 | 60S ribosomal protein L34 |
| *C. sojina* | Gm07.36875730 | Intron | Glyma07g31910 | Mitochondrial carnitine-acylcarnitine carrier protein |
| *Fusarium* spp. | Gm13.20484995 | CDS | Glyma13g16550 | Domain of unknown function (DUF3511) |
| *S. glycines* | Gm14.4811528 | Intron | Glyma14g06580 | Serine/threonine protein kinase |
| *C. sojina* | Gm15.1007132 | Intron | Glyma15g01550 | AUX/IAA family |
| *Fusarium* spp. | Gm17.8109237 | CDS | Glyma17g10780 | Thyroid receptor interacting protein related |
| *S. glycines* | Gm17.12684761 | Intron | Glyma17g15990 | Vesicle transport v-snare protein vti1-related |

was observed in BS analysis, as this disease affected plants only for the duration of the VER8 stage (see Table 1), which shows that the *S. glycines* did not affect plants in RIBSP and was barely important in KRIAPP. It is interesting that all ten studied accessions from East Asia and selected lines from Europe showed strong tolerance to all three diseases and could be directly involved in the breeding processes of soybean for resistance to studied fungi diseases.

On the other hand, the genetic study of the collection based on 4,442 polymorphic SNPs indicated a relatively close genetic relationship between the samples from Kazakhstan and North America, as PCoA test placed them together in left upper part of the graph (see Supplemental Fig. 2).

GWAS of the three diseases evaluated at the two stages of plant growth period allowed for identification of nine SNP markers associated with 15 MTAs (see Tables 3, 4). Two SNPs were identified in the GWAS of FUS on chromosomes 13 and 17 (see Fig. 2). The region on chromosome 13 matched with the well- known QTL (Fusarium lesion length 1–2) identified by M. Ellis et al. (2012). The authors had found that the region between the Satt160 and Satt149 markers was significantly associated with resistance to *Fusarium graminearum*. M. Kassem et al. (2006) reported that the Satt160 marker on chromosome 13 appeared to be a significant determinant of seed yield. It is interesting that this region was also associated with resistance to *P. sojae* (Wang H. et al., 2010). The region on the chromosome 17 has the same location with a QTL identified in GWAS for resistance to *Fusarium virguliforme* (Bao et al., 2015; Zhang et al., 2015). SNP marker Gm17.8109237 identified in this study has located approximately 4 Mb from SNP marker ss715611120_C_T identified by J. Zhang et al. (2015) and 6.7 Mb from SNP marker BARC-051665-11191 identified by Y. Bao et al. (2015).

The most significant amount of MTAs was found in GWAS for resistance to FLS (see Tables 3, 4). The SNP locations of identified five QTLs for FLS were mapped on chromosomes 2, 7, and 15, and did not match locations of the QTLs found in the previous study for resistance to this disease (Yang W. et al., 2001; Pham et al., 2015). A literature survey showed that one QTL for resistance to FLS matched the QTL previoulsy mapped on chromosome 13 (Pham et al., 2015), while another was positioned on chromosomes 16 (Yang W. et al., 2001). Therefore, the MTA found in this study presumably suggested that they are novel QTL for resistance to FLS.

In case of BS, the location of one out of two identified QTLs has matched the same region on chromosomes 14 with a QTL for resistance to sudden death syndrome (SDS) caused by *F. virguliforme* (Anderson et al., 2015). A physical position of associated SNP Gm14.4811528 for this QTL was in proximity of candidate gene Glyma14g06580 (Schmutz et al., 2010). The annotation of the gene is suggesting that it is a serine/threonine protein kinase, which is a common genetic factor often involved in controlling soybean diseases resistance (Cook et al., 2012). However, the identified position of the second MTA on chromosome 17 has possibly been reported for the first time. Therefore, in this study soybean QTL for resistances to FLS and BS were presumably novel ones. As the analyzed population in two regions has shown different reaction to tested diseases, these findings underline the importance of studying a genetically diverse collection of a particular soybean growing in a certain environmental niche. Identified MTAs may facilitate the discovery of new genes for resistance to diseases and a better understanding of genotype × environment interaction patterns. Also, the size and level of genetic variation in the studied genetic panels appear to be critical for the positive outcome of GWAS-based projects. It has been demonstrated that experiments with a sample size less than 384 accessions (Gurung et al., 2014) and large LD blocks (Zanke et al., 2014) might lead to identification of false positive associations. On the other hand, in the study by M.K. Turner et al. (2017), it was shown that smaller panels might allow for detection of false negative associations that would not have been detected in more extensive panels (Oyiga et al., 2017). Therefore, the results of this study using relatively small soybean collection size ($n$ = 182) may potentially relate to the above-mentioned findings by M.K. Turner and his coauthors.

With development of new genomic technologies, such as KASP (kompetitative allele-specific polymorphism) (Semagn et al., 2014), the designated SNP markers (see Tables 3, 4) for each of the identified MTAs for resistances to FUS, FLS and BS can be transformed into convenient types of DNA markers to enhance marker-associated selection projects in soybean.

Thus, the results of this study are a further contribution to the genetics and breeding of soybean associated with resistance to main fungal diseases.

## Acknowledgements

The authors would like to acknowledge the funding provided by the Ministry of Education and Sciences of the Republic of Kazakhstan under grant AP05131592.

## Conflict of interest

The authors declare they have no conflict of interest.

## References

Abugalieva S., Didorenko S., Anuarbek S., Volkova L., Gerasimova Y., Sidorik I., Turuspekov Y. Assessment of soybean flowering and seed maturation time in different latitude regions of Kazakhstan. PLoS One. 2016;11(12):e0166894. DOI 10.1371/journal.pone.0166894.

Anderson J., Akond M., Kassem M., Meksem K., Kantartzi S. Quantitative trait loci underlying resistance to sudden death syndrome (SDS) in MD96-5722 by 'Spencer' recombinant inbred line population of soybean. 3 Biotech. 2015;5(2):203-210. DOI 10.1007/s13205-014-0211-3.

Bao Y., Kurle J.E., Anderson G., Young N.D. Association mapping and genomic prediction for resistance to sudden death syndrome in early maturing soybean germplasm. Mol. Breeding. 2015;35:128. DOI 10.1007/s11032-015-0324-3.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;23:2633-2635. DOI 10.1093/bioinformatics/btm308.

Contreras-Soto R.I., Mora F., Lazzari F., Rott de Oliveira M.A., Scapium C.A., Chuster I. Genome-wide association mapping for flowering and maturity in tropical soybean: implications for breeding strategies. Breed. Sci. 2018;67:435-449. DOI 10.1270/jsbbs.17024.

Cook D.E., Lee T.G., Guo X., Melito S., Wang K., Bayless A.M., Wang J., Hughes T.J., Willis D.K., Clemente T.E., Diers B.W., Jiang J., Hudson M.E., Bent A.F. Copy number variation of multiple genes at *Rhg*1 mediates nematode resistance in soybean. Science. 2012;338:1206-1209. DOI 10.1126/science.1228746.

Copley T.R., Duceppe M.O., O'Donoughue L.S. Identification of novel loci associated with maturity and yield traits in early maturity soybean plant introduction lines. BMC Genomics. 2018;19:167. DOI 10.1186/s12864-018-4558-4.

Dias M.D., Pinheiro V.F., Café-Filho A.C. Impact of anthracnose on the yield of soybean subjected to chemical control in the north region of Brazil. Summa Phytopathol. 2016;42(1):18-23. DOI 10.1590/0100-5405/2114.

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

Ellis M., Wang H., Paul P., St. Martin S.K., McHale L., Dorrance A. Identification of soybean genotypes resistant to *Fusarium graminearum* and genetic mapping of resistance quantitative trait loci in the cultivar Conrad. Crop Sci. 2012;52(5):2224-2233.

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

Faske T., Kirkpatrick T., Zhou J., Tzanetakis I. Soybean diseases. Arkansas Soybean Production Handbook. 2014;11:1-18.

Gurung S., Mamidi S., Bonman J.M., Xiong M., Brown-Guedira G., Adhikari T.B. Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. PLoS One. 2014;9(9):e108179. DOI 10.1371/journal.pone.0108179.

Hnetkovsky N., Chang S.J.C., Doubler T.W., Gibson P.T., Lightfoot D.A. Genetic mapping of loci underlying field resistance to soybean sudden death syndrome (SDS). Crop Sci. 1996;36:393-400. DOI 10.2135/cropsci1996.0011183X003600020030x.

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

Kassem M.A.J., Shultz K., Meksem Y., Cho A.J., Wood M.J., Iqbal D., Lightfoot A. An updated 'Essex' by 'Forrest' linkage map and first composite interval map of QTL underlying six soybean traits. Theor. Appl. Genet. 2006;113:1015-1026. DOI 10.1007/s00122-006-0361-8.

Klein R.J. Power analysis for genome-wide association studies. BMC Genetics. 2007;8:58. DOI 10.1186/1471-2156-8-58.

Kurilova D.A. The harmfulness of soybean fusarium depending on the degree of damage to the plants. Maslichnye Kultury. Nauchno-tekhnicheskiy Byulleten Vserossiyskogo NII Maslichnyh Kultur = Oilseeds. Scientific and Technical Bulletin of the All-Russian Research Institute of Oilseeds. 2010;2(144-145):84-89. (in Russian)

Leath S., Carroll R.B. Screening for resistance to *Fusarium oxysporum* in Soybean. Plant Dis. 1982;66(12):1140-1143.

Masuda T., Goldsmith P.D. World soybean production: area harvested, yield, and long-tern projections. Int. Food Agribus. Man. Rev. 2009;12(4):143-161.

Mian M.A.R., Wang T., Phillips D.V., Alvernaz J., Boerma H.R. Molecular mapping of the *Rcs*3 gene for resistance to frogeye leaf spot in soybean. Crop Sci. 1999;39:1687-1691. DOI 10.1046/j.1439-0523.2001.00563.x.

Mian R., Bond J., Joobeur T., Mengistu A., Wiebold W., Snannon G., Wrather A. Identification of soybean genotypes resistant to *Cercospora sojina* by field screening and molecular markers. Plant Dis. 2009;93:408-411. DOI 10.1094/PDIS-93-4-0408.

Mombekova G.A., Shemshurova O.N., Seitbattalova A.I., Aitkhozhina N.A., Bekmakhanova N.E. Phytopathogens of sugar beet and soybean cultivated in soil and climatic conditions of Almaty region. NAN RK. 2013;4:8-11. (in Russian)

Mueller D., Robertson A., Sisson A., Tylka G. Soybean Diseases. Iowa State Univ. of Sci. and Technol., 2010.

Oyiga B.C., Sharma R.C., Baum M., Ogbonnaya F.C., Léon J., Ballvora A. Allelic variations and differential expressions detected at quantitative trait loci for salt stress tolerance in wheat. Plant Cell Environ. 2017;41(5):919-935. DOI 10.1111/pce.12898.

Pataky J.K., Lim S.M. Effects of septoria brown spot on the yield components of soybeans. Plant Dis. 1981;65:588-590.

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

Pham A.T., Harris D.K., Buck J., Hoskins A., Serrano J., Abdel-Haleem H. Fine mapping and characterization of candidate genes that control resistance to *Cercospora sojina* K. Hara in two soybean germplasm accessions. PLoS One. 2015;10(5):e0126753. DOI 10.1371/journal.pone.0126753.

Polozhieva Y.V., Dubovitskaya L.K. Soybean varieties evaluation on prevalence by complex of the root rots agents. Dalnevostochnyi Agrarnyi Vestnik = Far Eastern Agrarian Herald. 2015;3(35):35-38. (in Russian)

Prabhu R.R., Njiti V.N., Bell-Johnson B., Johnson J.E., Schmidt M.E., Klein J.H., Lightfoot D.A. Selecting soybean cultivars for dual resistance to soybean cyst nematode and sudden death syndrome using two DNA markers. Crop Sci. 1999;39(4):982-987.

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

Qin J., Song Q., Shi A., Li S., Zhang M., Zhang B. Genome-wide association mapping of resistance to *Phytophthora sojae* in a soybean [*Glycine max* (L.) Merr.] germplasm panel from maturity groups IV and V. PLoS One. 2017;12:e0184613. DOI 10.1371/journal.pone.0184613.

Schmutz J., Cannon S.B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D.L., Song Q., Thelen J.J., Cheng J., Xu D., Hellsten U., May G.D., Yu Y., Sakurai T., Umezawa T., Bhattacharyya M.K., Sandhu D., Valliyodan B., Lindquist E., Peto M., Grant D., Shu S., Goodstein D., Barry K., Futrell-Griggs M., Abernathy B., Du J., Tian Z., Zhu L., Gill N., Joshi T., Libault M., Sethuraman A., Zhang X.C., Shinozaki K., Nguyen H.T., Wing R.A., Cregan P., Specht J., Grimwood J., Rokhsar D., Stacey G., Shoemaker R.C., Jackson S.A. Genome sequence of the paleopolyploid soybean. Nature. 2010;463:178-183. DOI 10.1038/nature08670.

Schneider R., Rolling W., Song Q., Cregan R., Dorrance A.E., McHale L.K. Genome-wide association mapping of partial resistance to *Phytophthora sojae* in soybean plant introductions from the Republic of Korea. BMC Genomics. 2016;17(1):607. DOI 10.1186/s12864-016-2918-5.

Semagn K., Babu R., Hearne S., Olsen M. Single nucleotide polymorphism using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol. Breeding. 2014;33(1):1-14.

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

Stacey G. (Ed.) Genetics and Genomics of Soybean. Ser.: Plant Genetics and Genomics: Crops and Models (Vol. 2). Springer, 2008.

Turner M.K., Kolmer J.A., Pumphrey M.O., Bulli P., Chao S., Anderson J.A. Association mapping of leaf rust resistance loci in a spring wheat core collection. Theor. Appl. Genet. 2017;130:345-361. DOI 10.1007/s00122-016-2815-y.

Vidic M., Dordevic V., Petrovic K., Miladinovic J. Review of soybean resistance to pathogens. Ratar. Povrt. 2013;50(2):52-61. DOI 10.5937/ratpov50-4038.

Wang H., Waller L., Tripathy S., St. Martin S.K., Zhou L., Krampis K., Tucker D.M., Mao Y., Hoeschele I., Maroof S.M.A., Tyler B.M., Dorrance A.E. Analysis of genes underlying soybean quantitative trait loci conferring partial resistance to *Phytophthora sojae*. Plant Gen. 2010;3:23-40.

Wang J., Liu C.,Wang J., Qi Z., Li H., Hu G., Chen Q. An integrated QTL map of fungal diseases resistance in soybean (*Glycine max* L. Merr.): a method of meta-analysis for mining R genes. Agric. Sci. China. 2010;9(2):223-232. DOI 10.1016/S1671-2927(09)60087-0.

Wimmer V., Albrecht T., Auinger H., Schoen C. Synbreed: a framework for the analysis of genomic prediction data using R. Bioinformatics. 2012;28(15):2086-2087. DOI 10.1093/bioinformatics/bts335.

Yang W., Weaver D.B., Nielsen B.L., Qiu J. Molecular mapping of a new gene for resistance to frogeye leaf spot of soyabean in Peking. Plant Breed. 2001;120(1):73-78. DOI 10.1046/j.1439-0523.2001.00563.x.

Yang X.B., Feng F. Ranges and diversity of soybean fungal diseases in North America. Phytopathology. 2001;91(8):769-775.

Zanke C., Ling J., Plieske J., Kollers S., Ebmeyer E., Korzun V., Argillier O., Stiewe G., Hinze M., Beier S., Ganal M.W., Roder M.S. Genetic architecture of main effect QTL for heading date in European winter wheat. Front. Plant Sci. 2014;5:217. DOI 10.3389/fpls.2014.00217.

Zaostrovnykh V.I. Soybean diseases. Zaschita i Karantin Rastenii = Plant Protection and Quarantine. 2005;2:49-53. (in Russian)

Zatybekov A., Abugalieva S., Didorenko S., Gerasimova Y., Sidorik I., Anuarbek Sh., Turuspekov Y. GWAS of agronomic traits in soybean collection included in breeding pool in Kazakhstan. BMC Plant Biol. 2017;17(Suppl.4):63-70. DOI 10.1186/s12870-017-1125-0.

Zhang J., Singh A., Mueller D.S., Singh A.K. Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. Plant J. 2015;84:1124-1136. DOI 10.1111/tpj.13069.

Zhang J., Xia Ch., Wang X., Duan C., Sun S., Wu X., Zhu Zh. Genetic characterization and fine mapping of the novel Phytophtora resistance gene in a Chines soybean cultivar. Theor. Appl. Genet. 2013;126:1555-1561.

Zhang L., Kyei-Boahen S., Zhang J., Zhang M., Freeland T., Watson C., Liu X. Modifications of optimum adaptation zones for soybean maturity groups in the USA. Crop Management. 2007;6(1). DOI 10.1094/CM-2007-0927-01-RS.

Zhao G., Ablett G.R., Anderson T.R., Rajcan I., Schaafsma A.W. Inheritance and genetic mapping of resistance to rhizoctonia root and hypocotyl rot in soybean. Crop Sci. 2005;45:1441-1447. DOI 10.2135/cropsci2004.0560.

Zhao X., Han Y., Li Y., Liu D., Sun M., Zhao Y., Lu Ch., Li D., Yang Z., Huang L., Teng W., Qiu L., Zheng H., Li W. Loci and candidate gene identification for resistance to *Sclerotinia sclerotiorum* in soybean (*Glycine max* L. Merr.) via association and linkage maps. Plant J. 2015;82:245-255. DOI 10.1111/tpj.12810.

Вавиловский журнал генетики и селекции. 2018;22(5):544-552
DOI 10.18699/VJ18.393

# Alloplasmic recombinant lines (*H. vulgare*)-*T. aestivum* with 1RS.1BL translocation: initial genotypes for production of common wheat varieties

L.A. Pershina[1] ✉, L.I. Belova[1], N.V. Trubacheeva[1], T.S. Osadchaya[1], V.K. Shumny[1], I.A. Belan[2], L.P. Rosseeva[2], V.V. Nemchenko[3], S.N. Abakumov[4]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Federal State Budget Scientific Research Institution "Omsk Agrarian Scientific Center", Omsk, Russia
[3] Agrocomplex "Kurgansemena", Kurgan, Russia
[4] Federal State Unitary Enterprise "Ishimskoe", Tyumen Region, Russia

Alloplasmic lines are formed when the cytoplasm of one species is replaced by the cytoplasm of another as a result of repeated recurrent crosses of wide hybrids with the paternal genotype. Since the cytoplasm replacement results in new intergenomic interactions between a nucleus and cytoplasm leading to variability of plant characteristics, alloplasmic lines with restored fertility can be an additional source of biodiversity of cultivated plants. Earlier, recombinant alloplasmic lines (*H. vulgare*)-*T. aestivum* designated as L-17(1)–L-17(37) were formed from a plant with partially restored fertility of the $BC_3$ generation of barley-wheat hybrid *H. vulgare* (cv. Nepolegayushchii) × *T. aestivum* (cv. Saratovskaya 29). This male-sterile hybrid was consistently backcrossed with wheat varieties Mironovskaya 808 (twice) and Saratovskaya 29, and Mironovskaya 808 had a positive impact on the restoration of fertility. This paper presents the results of investigation into a group of recombinant alloplasmic lines (L-17$F_4$), as well as into doubled haploids (DH) lines – alloplasmic DH-17-lines obtained from anther culture of alloplasmic lines (L-17$F_2$). The most productive of these lines were used as initial breeding genotypes. Hybrid form Lutescens 311/00-22 developed from the crossing of the alloplasmic DH(1)-17 line (as maternal genotype) with euplasmic line Com37 (CIMMYT), the source of the 1RS.1BL wheat-rye translocation, proved to be successful for breeding. The presence of the 1RS.1BL translocation in the genome of the Lutescens 311/00-22 form and the L-311(1)–L-311(6) alloplasmic lines isolated from it did not lead to a decrease of fertility or sterility in the plants. This indicates that the chromosome of the 1BS wheat does not carry the gene(s) that determine the restoration of fertility in the studied (*H. vulgare*)-*T. aestivum* alloplasmic lines. Alloplasmic lines L-311(1)–L-311(6) showed their advantage in comparison with the standard varieties for resistance to leaf and stem rust, yield, and grain quality. The breeding tests performed at Omsk Agricultural Scientific Center, Agrocomplex "Kurgansemena", Federal State Unitary Enterprise "Ishimskoe" (Tyumen Region), using alloplasmic lines L-311(5), L-311(4) and L-311(6) resulted in varieties of spring common wheat Sigma, Uralosibirskaya 2 and Ishimskaya 11, respectively.

Key words: alloplasmic lines (*H. vulgare*)-*T. aestivum*; DH-lines; translocation 1RS.1BL; varieties of common wheat.

Received 15.03.2018
Accepted for publication 16.05.2018


✉ e-mail: pershina@bionet.nsc.ru

# Аллоплазматические рекомбинантные линии (*H. vulgare*)-*T. aestivum* с транслокацией 1RS.1BL: исходные генотипы для создания сортов мягкой пшеницы

Л.А. Першина[1] ✉, Л.И. Белова[1], Н.В. Трубачеева[1], Т.С. Осадчая[1], В.К. Шумный[1], И.А. Белан[2], Л.П. Россеева[2], В.В. Немченко[3], С.Н. Абакумов[4]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Федеральное государственное бюджетное научное учреждение «Омский аграрный научный центр», Омск, Россия
[3] Агрокомплекс «Кургансемена», Курган, Россия
[4] Федеральное государственное унитарное предприятие «Ишимское», Тюменская область, Россия

Аллоплазматические линии образуются при замещении цитоплазмы одного вида на цитоплазму другого в результате повторяющихся возвратных скрещиваний отдаленных гибридов с отцовским генотипом. Так как при замещении цитоплазмы между ядром и цитоплазмой возникают новые межгеномные взаимодействия, приводящие к изменчивости признаков растений, аллоплазматические линии с восстановленной фертильностью могут служить дополнительным источником биоразнообразия культурных растений. Ранее в наших работах были получены рекомбинантные аллоплазматические линии (*H. vulgare*)-*T. aestivum*, обозначенные как Л-17(1)–Л-17(37), сформированные от растения с частично восстановленной фертильностью $BC_3$ поколения ячменно-пшеничного гибрида *H. vulgare* (Неполегающий) × *T. aestivum* (Саратовская 29). Этот мужско-стерильный гибрид был последовательно беккроссирован сортами пшеницы Мироновская 808 (дважды) и Саратовская 29, где сорт Мироновская 808 оказал влияние на восстановление фертильности. В статье представлены результаты изучения группы рекомбинантных аллоплазматических линий Л-17$F_4$, а также линий гаплоидов с удвоенным числом хромосом – аллоплазматических ДГ-17-линий, полученных в результате культивирования пыльников линий Л-17$F_2$. Наиболее продуктивные из изученных линий включены в селекционный процесс. Успешной для селек-

ции оказалась гибридная форма Лютесценс 311/00-22, полученная от скрещивания аллоплазматической ДГ(1)-17-линии с эуплазматической линией Com37 (CIMMYT), источником пшенично-ржаной транслокации 1RS.1BL. Присутствие транслокации 1RS.1BL в геноме формы Лютесценс 311/00-22 и выделенных из нее аллоплазматических линий Л-311(1)–Л-311(6) не привело к снижению фертильности растений или их стерильности. Это указывает на то, что хромосома пшеницы 1BS не несет ген(ы), определяющие восстановление фертильности у изученных в настоящей работе аллоплазматических линий (*H. vulgare*)-*T. aestivum*. Линии Л-311(1)– Л-311(6) показали их преимущество по сравнению с сортами-стандартами по устойчивости к бурой ржавчине, стеблевой ржавчине, урожайности, качеству зерна. В результате селекционных испытаний в Омском аграрном научном центре, Агрокомплексе «Курган-семена», на предприятии «Ишимское» Тюменской области на основе аллоплазматических линий Л-311(5), Л-311(4) и Л-311(6) созданы сорта яровой мягкой пшеницы Сигма, Уралосибирская 2 и Ишимская 11 соответственно.

Ключевые слова: аллоплазматические линии (*H. vulgare*)-*T. aestivum*; ДГ-линии; транслокация 1RS.1BL; сорта мягкой пшеницы.

**HOW TO CITE THIS ARTICLE:**

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

Alloplasmic lines are formed as a result of repeated recurrent crosses of wide hybrids with the paternal genotype and combine the cytoplasm of the maternal species with the nuclear genome of the pollen parent (Tsunewaki, 1996). The new nuclear-cytoplasmic interactions formed due to cytoplasm replacement can cause epigenetic modifications of nuclear genes (Soltani et al., 2016), leading to changes at the level of transcription and metabolism (Crosatti et al., 2013; Soltani et al., 2016). These processes may lead to plant developmental disorders (Suzuki et al., 1995), changes in resistance to stress factors (Dhitaphichit et al., 1989; Keane, Jones, 1990; Buloychik et al., 2002; Talukder et al., 2015) and changes in the manifestation of agronomically important traits (Ekiz et al., 1998; Liu et al., 2002; Atienza et al., 2008; Klimushina et al., 2013). The most common sign of a nuclear-cytoplasmic conflict included in alloplasmic lines is cytoplasmic male sterility (CMS) (Tsunewaki, 1996; Bentolila et al., 2002).

From a practical perspective, alloplasmic lines are obtained from different species of cultivated plants and characterized by CMS that represents a system for obtaining hybrid seeds in heterotic breeding (Cisar, Cooper, 2002). Given that new intergenomic interactions arise during cytoplasm replacement, alloplasmic lines with restored fertility can be an additional source of biodiversity in cultivated plants (Liu et al., 2016). For wheat alloplasmic lines, examples of their practical use are not numerous. The VPM1 alloplasmic line carrying the cytoplasm of *Aegilops ventricosa* Tausch was the source of two commercial varieties of wheat Roason and Rendezvous (Jones et al., 1998) resulting from hybridization of the amphiploid (*Ae. ventricosa* × *T. persicum*) with common wheat variety Marne. This alloplasmic line and the varieties derived from it are resistant to a number of fungal pathogens due to the introgression of *Ae. ventricosa* genes *Pch-1* and *Sr38*/*Lr37*/*Yr17* (Delibes et al., 1988; Friebe et al., 1996). Common wheat variety ‘Xiaoshan 2134’ carrying the cytoplasm of *Ae. crassa*, is characterized by high grain quality, resistance to salinity and high yield (Liu et al., 2002). The authors ascribe these traits to the effect of nuclear-cytoplasmic heterosis. The lines of common wheat with cytoplasm *Triticum timopheevii* and *Secale cereale* characterized by tolerance to drought (Semenov et al., 2014) and high-quality gluten characteristics are considered promising for breeding (Klimushina et al., 2013; Semenov et al., 2016). With the use of rye cytoplasm, new forms of triticale have been obtained (Gordey et al., 2011).

Earlier, we reported the production of the recombinant alloplasmic lines of common wheat derived from back-cross progenies of barley-wheat hybrids *H. vulgare* × *T. aestivum* (Pershina et al., 1998) and restoration of their fertility (Pershina et al., 1999a). Alloplasmic lines (*H. vulgare*)-*T. aestivum* characterized by different fertility levels proved to be valuable models for studying the variability of both nuclear and organellar (mitochondrial and chloroplast) genomes in the process of nuclear-cytoplasmic co-adaptation (Bildanova et al., 2004; Aksyonova et al., 2005; Pershina et al., 2014). Some

of the recombinant alloplasmic lines with full fertility restoration were included in breeding of common wheat (Belan et al., 2017).

The aim of presented research was to generalize and analyse the results of obtaining and studying of recombinant and introgressive alloplasmic lines (*H. vulgare*)-*T. aestivum*, which have been successfully used in breeding and have become a source of the new varieties of spring common wheat.

## Materials and methods

**Production of recombinant and introgressive alloplasmic lines (*H. vulgare*)-*T. aestivum*.** The Nepolegayushchii variety of cultivated barley was crossed with spring common wheat Saratovskaya 29, and the pollinated flowers were treated with a solution of gibberellic acid. Using embryo cultivation, hybrid plants *H. vulgare* × *T. aestivum* were grown and were characterized by male sterility but female fertility (Pershina et al., 1998). The $F_1$ hybrids were consistently backcrossed with wheat Mironovskaya 808 (twice) and Saratovskaya 29. Among the plants from the $BC_3$ generation, one partially fertile plant was isolated (three seeds were set in one spike, and 34 in another) (Pershina et al., 1999a). Each fertile plants grown from these seeds became the source of recombinant alloplasmic lines designated L-17(1)–L17(37) (Fig. 1). To form each subsequent generation of alloplasmic lines, the seeds from the main spike of the most productive plant were selected. Alloplasmic DH-lines DH(1)-17$F_2$, DH(1)-17$F_2$, and DH(1)-17$F_2$ were derived by anther culture of alloplasmic line L-17$F_2$ (Pershina et al., 1999b).

In this paper, we present the results of a study of seven alloplasmic lines: L-17(3)$F_4$, L-17(4)$F_4$, L-17(9)$F_4$, L-17(12)$F_4$, L-17(18)$F_4$, L-17(21)$F_4$, and L-17(24)$F_4$; three alloplasmic DH lines: DH(1)-17$F_2$, DH(2)-17$F_2$, and DH(3)-17$F_2$; and six introgression alloplasmic lines: L-311(1)$F_{10}$–L-311(6)$F_{10}$. The introgression alloplasmic lines were isolated from the hybrid form L-311/00-22$F_5$ obtained as a result of the crossing of alloplasmic line DH(1)-17$F_2$ with euplasmic line Com37 (Belan et al., 2010). Using the results of chromosome C-banding, E.D. Badaeva identified the wheat-rye translocation 1RS.1BL in the alloplasmic lines of group L-311 (Pershina et al., 2013).

**Growing conditions and study of alloplasmic lines.** Alloplasmic lines L-17$F_4$ and DH-17$F_2$ were grown in the field of Institute of Cytology and Genetics near Novosibirsk in 1999 when there was no mass development of leaf rust or stem rust pathogens. Control wheat line Saratovskaya 29 (Sar29) was used to obtain barley-wheat hybrids and alloplasmic lines (*H. vulgare*)-*T. aestivum*. The alloplasmic lines of the L-311$F_{10}$ group were grown in the field of the Institute of Cytology and Genetics in 2017. Euplasmic line Om37 carrying translocation 1RS.1BL and isolated from variety Omskaya 37, and alloplasmic line L-17(3)$F_{12}$ were used as controls. Plants were grown on plots 50 cm wide in rows of 10, with a distance of 25 cm between rows. The germination rate was determined by the number of seedlings. During harvesting, the height of

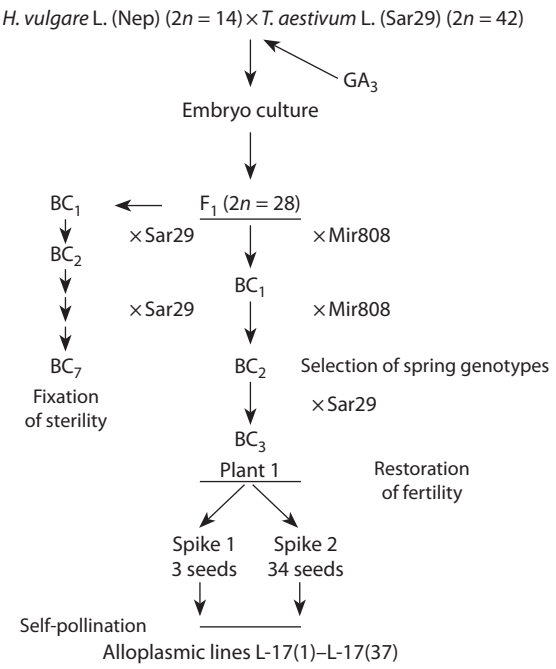



**Fig. 1.** Development of a set of recombinant (*H. vulgare*)-*T. aestivum* alloplasmic lines L-17 (Pershina et al., 1998, 1999a).

Designation: Nep – barley variety Nepolegayushchii; varieties of common wheat: Sar29 – Saratovskaya 29, Mir808 – Mironovskaya 808. $GA_3$ – gibberellic acid.

the plants, the number of productive spikes, the length of the main spike, the number of spikelets per main spike, the number of seeds per main spike, and the frequency of plants with full fertility (more than 35 seeds in the main spike) were estimated.

To confirm the presence of the translocation 1RS.1BL, the SCAR marker iag95 linked to the genes *Lr26* and *Sr31* localized in the short arm of the rye chromosome 1R (Mago et al., 2002), and genomic *in situ* hybridization (Mukai, Gill, 1991) were used in the lines included in this study. At least 20 plants were used per treatment. The differences between the average values of the lines were compared using *t*-tests. Data were analysed using Statistica v.7.0.61.0.

Alloplasmic hybrid form L-311/00-22 and alloplasmic lines L-311(1)–L-311(6) were grown in the field of Laboratory of Spring Common Wheat of Omsk Agrarian Scientific Center, following the methods described in (Belan et al., 2017). Since 2007, the L-311 alloplasmic lines have been tested according the full breeding scheme for BN-1, BN-2, and BN-3 (breeding nurseries of the first, second, and third year of study) and in a competitive variety trial nursery. Assessments of resistance to powdery mildew (*Blumeria graminis* f. sp. *tritici*), leaf rust (*Puccinia recondita* f. sp. *tritici*), and stem rust (*Puccinia graminis* f. sp. *tritici*) were made according to previously described methods (Belan et al., 2017). In addition, yield and grain quality parameters (1000-grain weight, seed protein content) were determined. Using analogous methods, alloplas-

**Table 1.** Characteristics of the alloplasmic lines of L-17$F_4$ and DH-17$F_2$ (Novosibirsk, field, 1999)

| Genotypes | Height of plants, cm | Number of productive spikes | Length of the main spike, cm | Number of spikelets per main spike | Number of seeds | | 1 000-grain weight | % of FF plants |
|---|---|---|---|---|---|---|---|---|
| | | | | | per main spike | per plant | | |
| Sar29 | 96.3 ± 2.0 | 4.3 ± 0.3 | 9.1 ± 0.1 | 13.8 ± 0.2 | 40.1 ± 1.1 | 152.3 ± 13.4 | 45.8 ± 0.7 | 80 |
| L-17(9)$F_4$ | 87.5 ± 2.1$^{(**)}$ | 3.8 ± 0.2 | 8.2 ± 0.2$^{(***)}$ | 12.9 ± 0.2$^{(**)}$ | 38.1 ± 1.9 | 122.8 ± 10.9 | 42.9 ± 0.9$^{(*)}$ | 60$^{(*)}$ |
| L-17(18)$F_4$ | 93.2 ± 1.9 | 2.9 ± 0.3$^{(**)}$ | 7.7 ± 0.1$^{(***)}$ | 12.1 ± 0.2$^{(***)}$ | 33.3 ± 1.7$^{(**)}$ | 88.6 ± 12.6$^{(**)}$ | 43.0 ± 0.8$^{(*)}$ | 65 |
| L-17(4)$F_4$ | 96.4 ± 1.3 | 5.4 ± 0.6 | 9.3 ± 0.2 | 13.6 ± 0.4 | 42.8 ± 2.1 | 194.6 ± 20.5 | 44.5 ± 0.8 | 85 |
| L-17(21)$F_4$ | 94.6 ± 1.6 | 5.2 ± 0.4 | 9.1 ± 0.1 | 14.2 ± 0.3 | 44.8 ± 1.5* | 192.1 ± 14.6* | 43.5 ± 0.6$^{(*)}$ | 95 |
| L-17(12)$F_4$ | 100 ± 1.7 | 5.6 ± 0.3** | 9.2 ± 0.1 | 14.5 ± 0.3 | 41.8 ± 1.0 | 197.9 ± 15.7* | 44.4 ± 0.7 | 95 |
| #L-17(3)$F_4$ | 85.6 ± 2.5$^{(**)}$ | 6.1 ± 0.3** | 9.4 ± 0.2 | 14.2 ± 0.3 | 42.6 ± 1.4 | 199.7 ± 15.4* | 44.9 ± 0.9 | 75 |
| #L-17(24)$F_4$ | 99.8 ± 1.7 | 6.6 ± 0.3*** | 9.3 ± 0.1 | 15.5 ± 0.2*** | 42.4 ± 0.9 | 227.9 ± 12.7*** | 45.7 ± 0.6 | 90 |
| #DH(1)-17$F_2$ | 96.1 ± 1.1 | 6.7 ± 0.3*** | 9.3 ± 0.2 | 15.2 ± 0.1*** | 45.4 ± 1.2** | 238.4 ± 12.3*** | 45.9 ± 0.5 | 85 |
| #DH(2)-17$F_2$ | 80.2 ± 1.3$^{(***)}$ | 5.5 ± 0.3*** | 9.3 ± 0.2 | 14.8 ± 0.2** | 45.2 ± 1.6* | 188.8 ± 10.9* | 46.6 ± 0.4 | 100 |
| #DH(3)-17$F_2$ | 90.9 ± 1.3$^{(*)}$ | 6.7 ± 0.4*** | 10.0 ± 0.1*** | 15.4 ± 0.2*** | 44.6 ± 1.4* | 239.7 ± 15.9*** | 44.9 ± 0.5 | 90 |

Note: FF – full fertility. The difference in comparison with the parent Saratovskaya 29 line is significantly less at $^{(*)}$ $p < 0.05$; $^{(**)}$ $p < 0.01$ and $^{(***)}$ $p < 0.001$; and significantly more at * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. # indicates the lines included in breeding.

mic line L-311(4) has been studied in the Agrocomplex "Kurgansemena" competitive variety trial nursery since 2013, and alloplasmic line L-311(6) – since 2014 in the Enterprise "Ishimskoe" (Tyumen Region) competitive variety trial nursery.

## Results and discussion

In this study recombinant alloplasmic lines (*H. vulgare*)-*T. aestivum* were inverstigated, in which barley variety Nepolegayushchii (Nep) was a source of the cytoplasm, and the recombinant nuclear genome was formed with the use of common wheat varieties Saratovskaya 29 (Sar29) and Mironovskaya 808 (Mir808) (see Fig. 1) (Pershina et al., 1999a). Use of the Mir808 variety in recurrent crosses of male sterile barley-wheat hybrid *H. vulgare* (Nep) × *T. aestivum* (Sar29) had a positive effect on fertility restoration: one of the plants of the $BC_3$ generation set seeds after self-pollination (see Fig. 1).

According to our data, the Mironovskaya yarovaya wheat variety recovered from Mir808 (Dorofeev et al., 1987), along with the Ulyanovka and Pyrotrix 28 varieties (Pershina et al., 1999a, 2014), which also restored full fertility of common wheat in the barley (*H. vulgare*) cytoplasm. When the backcrossing of barley-wheat hybrids *H. vulgare* × *T. aestivum* (Sar29) included only the Sar29 variety in the backcrossed progenies, fixation of complete sterility occurred (see Fig. 1) (Pershina et al., 1999a, 2014). In addition, depending on the genetic diversity of wheat varieties included in the backcrossing of barley-wheat hybrids, numerous alloplasmic lines (*H. vulgare*)-*T. aestivum* with different levels of fertility were obtained, some of which, when self-pollinated, segregated into plants with low fertility including those being completely sterile (Pershina et al., 1999a, 2014).

Alloplasmic lines L-17(1)–L-17(37) formed from a partially fertile $BC_3$ generation plant with a $2n = 42$ did not contain barley chromosomes and were fertile in $F_2$–$F_3$ generations (Bildanova et al., 2004). The germination of the recombinant L-17$F_4$ alloplasmic lines was similar to the control level (not less than 98 %), and all the plants were fertile (Table 1).

In the majority of alloplasmic lines, the frequency of plants with full fertility (seed set higher than 35 seeds in the main spike) was at the level of the Sar29 parent line, varying from 65 % in L-17(18)$F_4$ to 95 % in L-17(21)$F_4$ and L-17(12)$F_4$ (see Table 1). The full fertility was significantly lower than that of the control (60 %) in alloplasmic line L-17(9)$F_4$. The analysis of the productivity characteristics (the number of productive spikes, the length of the main spike, the number of spikelets and seeds per main spike, and the number of seeds per plant) differed between different alloplasmic lines. So, in L-17(9)$F_4$ and L-17(18)$F_4$, the values of some of these parameters were significantly lower than in the Sar29 control line, while in L-17(4)$F_4$ they reached the control level. Alloplasmic lines L-17(21)$F_4$, L-17(12)$F_4$, L-17(3)$F_4$ and L-17(24)$F_4$ exceeded line Sar29 in the number of seeds per plant. The increase in L-17(21)$F_4$ was due to an increase in the number of seeds per main spike; an increase of a number of productive spikes in alloplasmic lines L-17(3)$F_4$ and L-17(12)$F_4$; and an increase in the number of productive spikes per plant and the number of spikelets per main spike in alloplasmic line L-17(24)$F_4$. For 1000-grain weight, the values of this parameter were the same at the

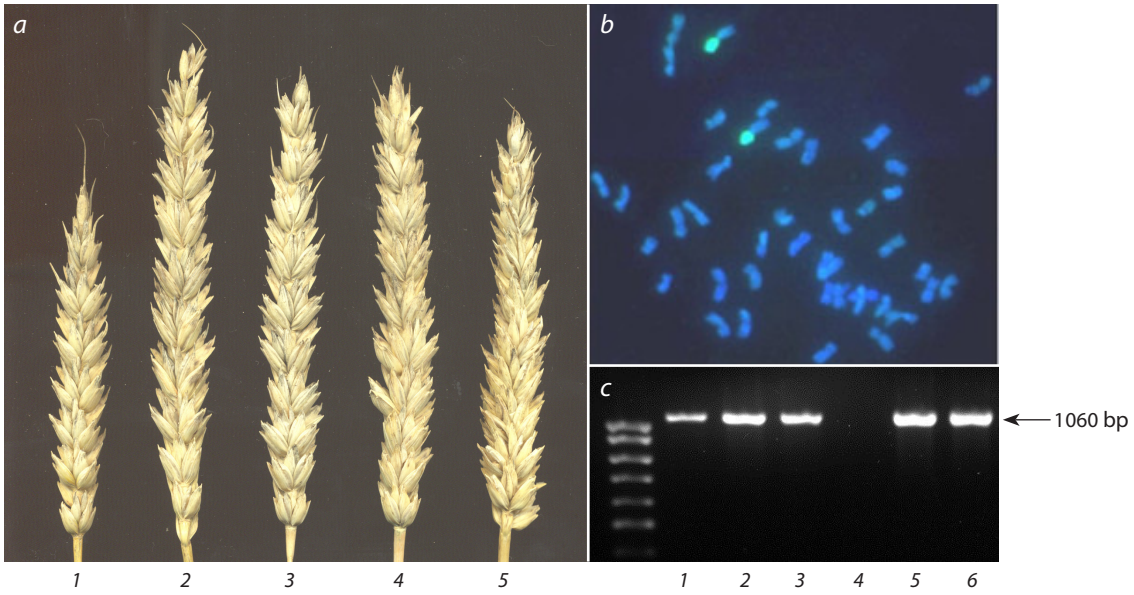



**Fig. 2.** The proof of the presence of wheat-rye translocation in the alloplasmatic lines of L-311(4)$F_{10}$, L-311(5)$F_{10}$, L-311(6)$F_{10}$ by the results of GISH and PCR analysis.

*a* – spikes: *1* – alloplasmic line L-17(3)$F_{12}$; *2–4* – alloplasmic lines L-311(4)$F_{10}$, L-311(5)$F_{10}$, L-311(6)$F_{10}$; *5* – euplasmic line Om37. *b* – GISH-analysis of alloplasmic lines L-311(4)$F_{10}$ (1RS chromosome is marked in green). *c* – PCR assay developed for iag95 marker linked with *Lr26* gene: *1* – rye *S. cereale* (control), *2* – Om37, *3* – L-311(4)$F_{10}$, *4* – L-17(3)$F_{12}$, *5* – L-311(5)$F_{10}$, *6* – L-311(6)$F_{10}$.

control level for the majority of the lines studied and was lower for lines F-17(9)$F_4$, L-17(18)$F_4$, and L-17(21)$F_4$ (see Table 1).

These data show that the new nuclear genome that formed as a result of recombination between two genomes of wheat varieties Saratovskya 29 and Mironovskaya 808 when interacting with the cytoplasm of barley had a positive effect not only on the restoration and maintenance of fertility of alloplasmic lines (*H. vulgare*)-*T. aestivum* but also on the traits that determine the productivity of plants. This is mostly due to the high ecological plasticity and high productivity of Saratovskaya 29 and Mironovskaya 808, which have been grown for a long time in many regions with different soil and climatic conditions and hybridized with each other and with other genotypes of wheat to produce many varieties (Dorofeev et al., 1987).

Another important result of our study is that fertility and the signs of high productivity were also evident in the DH lines derived from alloplasmic line L-17$F_2$. Earlier we had been able to obtain 20 green plants grown from anther culture of alloplasmic line L-17$F_2$ on a modified medium of P-II (Pershina et al., 1999b). Eleven of these plants with $2n = 42$ were fertile. These plants were used for development of a separate alloplasmic DH line. In the current paper we present the results of a study of the self-pollinated progeny of three alloplasmic DH lines: DH(1)-17$F_2$, DH(2)-17$F_2$, and DH(3)-17$F_2$ (see Table 1). In these lines, in comparison with line Sar29, the number of productive spikes, the number of spikelets and seeds per main spike, and the number of seeds per plant were significantly higher. Due to their high level of fertility and productivity, these alloplasmic DH lines, along with the original alloplasmic lines L-17(3)$F_4$ and L-17(24)$F_4$, were included in the breeding. In addition, among these alloplasmic lines, we found three more undersized lines such as L-17(3)$F_4$, DH(2)-17$F_2$, and DH(3)-17$F_2$.

To date, from a practical point of view, the hybrid combination obtained in Laboratory of Spring Common Wheat Breeding (Omsk Agrarian Scientific Center) by crossing alloplasmic line DH(1)-17$F_2$ (maternal genotype) with line Com37 (pollinator, provided for the study by A.I. Morgunov from the CIMMYT collection) has been the most effective. In the early breeding nurseries, hybrid alloplasmic form Lutescens 311/00-22 was reproduced as a population, from which, after selection based on agronomical valuable traits, six promising alloplasmic lines were identified: L-311(1), L-311(2), L-311(3), L-311(4), L-311(5), and L-311(6).

Using C-banding, E.D. Badaeva revealed in these alloplasmic lines a wheat-rye translocation identified as 1RS.1BL (Pershina et al., 2013), which was inherited in a number of self-pollinated generations. Figure 2 shows the spikes of alloplasmic lines L-311(4) $F_{10}$, L-311(5)$F_{10}$, L-311(6)$F_{10}$ and the presence of translocation of 1RS.1BL in these lines prooved by the results of GISH and PCR analysis. Since translocation 1RS.1BL was not detected the alloplasmic line of group L-17, it can be assumed that Com37 used as a pollinator in production of hybrid Lutescens 311/00-22, had this translocation. (Currently,

**Table 2.** Characteristics of alloplasmic lines L-311(1)$F_{10}$–L-311(6)$F_{10}$ in comparison with euplasmic line Om37-1RS.1BL and alloplasmic line L-17(3)$F_{12}$

| Genotypes | Number of productive spikes | Length of main spike, cm | Number of spikelets per main spike | Number of seeds | |
|---|---|---|---|---|---|
| | | | | per main spike | per plant |
| Om37 | 5.0 ± 0.5 | 10.5 ± 0.2 | 18.6 ± 0.3 | 41.5 ± 1.2 | 183.2 ± 18.1 |
| L-311(1)$F_{10}$ | 5.2 ± 0.3 | 11.5 ± 0.2** | 19.6 ± 0.3* | 47.9 ± 1.2** | 194.3 ± 18.5 |
| L-311(2)$F_{10}$ | 6.0 ± 0.4 | 11.8 ± 0.1*** | 19.9 ± 0.2*** | 49.5 ± 0.8*** | 231.8 ± 18.3 |
| L-311(3)$F_{10}$ | 6.1 ± 0.5 | 12.4 ± 0.1*** | 20.3 ± 0.3*** | 49.5 ± 1.2*** | 257.4 ± 16.2** |
| L-311(4)$F_{10}$ | 6.8 ± 0.4* | 12.4 ± 0.1*** | 20.9 ± 0.3*** | 56.8 ± 1.0*** | 263.5 ± 19.3** |
| L-311(5)$F_{10}$ | 6.5 ± 0.5* | 12.7 ± 0.3*** | 20.3 ± 0.4*** | 50.9 ± 2.1*** | 252.1 ± 22.2* |
| L-311(6)$F_{10}$ | 6.3 ± 0.4* | 12.8 ± 0.2*** | 21.2 ± 0.4*** | 53.8 ± 1.2*** | 260.3 ± 22.6* |
| L-17(3)$F_{12}$ | 4.7 ± 0.4 | 8.5 ± 0.2(***) | 14.6 ± 0.2(***) | 32.6 ± 1.2(***) | 153.7 ± 13.4 |

Note: The difference in comparison with euplasmic line Om37 is significantly greater at * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; and significantly less at (***) $p < 0.001$.

**Table 3.** Results of studying promising alloplasmic lines of group L-311 in a competitive variety trial nursery, Omsk, 2010

| Lines | Growing season, days | Productivity, t/ha | | 1 000-grain weight | | Protein content in the grain (N × 5.7), % | | Lesion of rust disease | |
|---|---|---|---|---|---|---|---|---|---|
| | | lines | ± of Om33 | lines | ± of Om33 | lines | ± of Om33 | leaf rust | stem rust |
| L-311(1) | 98 | 4.75 | 0.16 | 47.7 | 4.6 | 16.56 | 1.48 | 0 | 10 |
| L-311(2) | 97 | 5.00 | 0.41 | 44.9 | 1.8 | 18.24 | 3.16 | 1 | 20 |
| L-311(3) | 98 | 5.29 | 0.7 | 43.0 | -0.1 | 18.78 | 3.7 | 1 | 5 |
| L-311(4) | 98 | 5.05 | 0.46 | 44.4 | 1.3 | 17.78 | 2.7 | 5 | 5 |
| L-311(5) | 94 | 5.25 | 0.66 | 45.8 | 2.7 | 17.64 | 2.56 | 5 | 10 |
| L-311(6) | 93 | 5.00 | 0.4 | 45.5 | 2.4 | 17.53 | 2.45 | 5 | 10 |
| Om33-St | 93 | 4.59 | – | 43.1 | – | 15.08 | – | 80 | 50 |

Note: Om33-St is a variety of spring common wheat Omskaya 33 used as a reference.

the Com37 line and alloplasmic lines DH(1)-17$F_2$ have been lost.)

The fact that the substitution of chromosome 1BS for 1RS did not result in decrease of fertility or sterility in alloplasmic lines (*H. vulgare*)-*T. aestivum* indicates that the restoration and maintenance of fertility of these lines does not depend on the effect of *Rf*-genes controlling the restoration of wheat fertility in the alien cytoplasm and localized in chromosome 1BS (Tsunewaki, 2015).

In our previous work, other genotypes of alloplasmic lines (*H. vulgare*)-*T. aestivum* were obtained in which the introduction of wheat-rye translocation 1RS.1BL into the genome, including a combination with wheat-*Agropyrum* translocation 7DL-7Ai, did not have a negative effect on the fertility (Pershina et al., 2014).

The results of a study into the morphobiological traits of alloplasmic lines L-311(1)$F_{10}$–L-311(6)$F_{10}$ in comparison with alloplasmic line L-17(3)$F_{12}$ (has no the1RS.1BL translocation), and euplasmic line Om37 carrying this translocation, are presented in Table 2.

In all the six alloplasmic lines of group L-311$F_{10}$ carrying the1RS.1BL translocation the lengths of the main spike, the number of spikelets and the number of seeds per main spike were significantly higher than in control line L-17(3)$F_{12}$ and line Om37. In addition, alloplasmic lines L-311(3)$F_{10}$, L-311(4)$F_{10}$, L-311(5)$F_{10}$, and L-311(6)$F_{10}$ exceeded the control line in the number of seeds per plant, and the lines L-311(4)$F_{10}$, L-311(5)$F_{10}$, and L-311(6)$F_{10}$ – in the number of productive spikes.

Since rye chromosome arm 1RS carries a complex of genes (*Lr26*/*Sr31*/*Yr9*/*Pm8*) controlling resistance to fungal pathogens (Singh et al., 1990), the varieties with 1RS.1BL have become widespread throughout the world (Rabinovich, 1998; Schlegel, 2018). Mass cultivation of homogeneous varieties of wheat with this translocation has reduced the effectiveness of the gene resistance to fungal pathogens, including appearance of the highly aggressive *Ug99* race virulent to the *Sr31* gene (Pretorius et al., 2000). At the same time, the effect of 1RS.1BL translocation on the manifestation of agronomically important and adaptive

traits depends on the genetic background and plant growing regions (Hoffmann, 2008; Tahmasebi et al., 2015). The gene *Sr31* remains effective for protection against stem rust in different regions of Russia (Gultyaeva, 2017).

From the data obtained a conclusion can be made that the genetic background of the alloplasmic lines of group L-311 was favourable for realization of the positive effect of the 1RS.1BL translocation on both agronomically important and adaptive traits in the breeding trials regions including Omsk, Kurgan and Tyumen regions.

The study of alloplasmic lines L-311(1)–L-311(6) for consecutive self-pollinated generations in breeding nurseries and competitive variety trial nurseries was initially carried out in the fields of the Omsk Agrarian Scientific Center. It should be emphasized that leaf rust expansion had been noted there every year since 2001, and in 2007 the epiphytoty of this pathogen was recorded. Since 2009, the threat of stem rust has increased. In addition, powdery mildew is observed in this region every year.

Studying alloplasmic lines L-311(1)–L-311(6) demonstrated their moderate resistance to powdery mildew. In comparison with variety-standard Omskaya 33 susceptible to leaf and stem rust pathogens, the alloplasmic lines showed complex resistance to local populations of leaf and stem rust (Table 3).

Moreover, these alloplasmic lines exceeded the variety-standard for yield, 1000-grain weight, and seed protein content. Table 3 presents the results of studying alloplasmic lines L-311(1)–L-311(6) in 2010 in a competitive variety trial nursery. According to the tests results obtained from 2010 to 2012, middle-ripened alloplasmic line L-311(5) was identified as the most promising and was transferred in 2013 to the State Committee for Testing of New Varieties as variety Sigma. This variety was included in State Register of Selection Achievements of the Russian Federation in West Siberia in 2016.

The middle-ripened alloplasmic line L-311(4) was studied in the competitive variety trial nursery of Agrocomplex "Kurgansemena". Based on the three year results (2012–2015), variety Uralosibirskaya 2 was transferred to the State Committee for Testing of New Varieties in 2016. This variety was tested in the Urals and West Siberian region. In 2014, middle-early alloplasmic line L-311(6) was studied in the competitive variety trial nursery of Enterprise "Ishimskoe", Tyumen Region. Based to the four-year data this line was transferred to the State Committee for Testing of New Varieties as variety Ishimskaya 11 in West Siberia in 2017.

In addition, the efficiency of using alloplasmic lines from group L-311 for obtaining hybrids and the alloplasmic DH lines combining resistance genes to fungal pathogens localized in wheat-rye translocation 1RS.1BL with complexes of effective resistance genes from other sources has been shown (Pershina et al., 2013).

## Acknowledgements

The study is supported by budget funding for the state task (Project No. 3324-2018-0018), as well as the RFBR grant No. 17-04-01738. The authors are grateful to A.I. Morgunov for the provided material for crossing to obtain new breeding forms.

## Conflict of interest

The authors declare they have no conflict of interest.

## References

Aksyonova E., Sinyavskaya M., Danilenko N., Pershina L., Nakamura C., Davydenko O. Heteroplasmy and paternally oriented shift of the organellar DNA composition in barley-wheat hybrids during backcrosses with wheat parents. Genome. 2005;48(5):761-769.

Atienza S.G., Martin A., Peechioni N., Platani C., Cattivelli L. The nuclear-cytoplasmic interaction controls carotenoid content in wheat. Euphytica. 2008;159:325-331.

Belan I., Rosseeva L., Laikova L., Rosseev V., Pershina L., Trubacheeva N., Morgounov A., Zelenskiy Y. Utilization of new wheat genepool in breeding of spring bread wheat. Proc. of 8th Int. Wheat Conf. St. Petersburg, 1–4 June 2010. St. Petersburg, 2010;69-70.

Belan I.A., Rosseeva L.P., Meshkova L.V., Blokhina N.P., Pershina L.A., Trubacheeva N.V. Development of soft wheat varieties resistant to fungal diseases for West Siberia and the Ural. Vestnik Altayskogo Gosudarstvennogo Agrarnogo Universiteta = Bulletin of the Altai State Agricultural University. 2017;1(147):5-7. (in Russian)

Bentolila S., Alfonso A.A., Hanson M.R. A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male sterile plants. Proc. Natl. Acad. Sci. USA. 2002;99:10887-10892.

Bildanova L.L., Badaeva E.D., Pershina L.A., Salina E.A. Molecular study and C-banding of chromosomes in common wheat alloplasmic lines obtained from the backcross progeny of barley–wheat hybrids *Hordeum vulgare* L. (2*n* = 14) × *Triticum aestivum* L. (2*n* = 42) and differing in fertility. Russ. J. Genet. 2004;40(12):1383-1391.

Buloychik A.A., Voluevich E.A., Mikhno A.M. Genome and plasmon effects on the expression of defeated genes for resistance to brown rust in wheat. Tsitologiya i Genetika = Cytology and Genetics. 2002;36(2):11-19. (in Russian)

Cisar G., Cooper D.B. Hybrid wheat. In: Curtis B.C., Rajaram S. Gomez Macpherson H. (Eds). Bread Wheat: Improvement and Production. Rome: Food and Agriculture Organization, 2002;30:157-174.

Crosatti C., Quansah L., Maré C., Giusti L., Roncaglia E., Atienza S.G., Cattivelli L., Fait A. Cytoplasmic genome substitution in wheat affects the nuclear-cytoplasmic cross-talk leading to transcript and metabolite alterations. BMC Genomics. 2013;14:868-889. https://doi.org/10.1186/1471-2164-14-868.

Delibes A., Doussinault G., Mena M., López-Brana I., Garcia-Olmedo F. Eyespot resistance gene *Pch-1* from *Aegilops ventricosa* is associated with a different chromosome in wheat line H-93-70 than the resistance factor in "Roazon" wheat. Theor. Appl. Genet. 1988; 76(4):573-576. DOI 10.1007/BF00260911.

Dhitaphichit P., Jones P., Keane E.M. Nuclear and cytoplasmic gene control of resistance to loose smut (*Ustilago tritici* (Pers.) Rostr.) in wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 1989;78(6): 897-903.

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

Ekiz H., Safi Kiral A., Akçin A., Simsek L. Cytoplasmic effects on quality traits of bread wheat (*Triticum aestivum* L.). Euphytica. 1998;100(1):189-196.

Friebe B., Jiang J., Raupp W.J., McIntosh R.A., Gill B.S. Characterization on wheat-alien translocations conferring resistance to diseases and pests: current status. Euphytica. 1996;91(1):59-87.

Gordey I.A., Belko N.B., Lyusikov O.M. Sekalotritikum (×*Secalotriticum*): the genetic basis for the creation and formation of the genome. Minsk: Belaruskaya Navuka Publ., 2011. (in Russian)

Gultyaeva E.I. Rye translocations in cultivars of common wheat included in National Register of Breeding Achievements. Proc. of 3d Int. Conf. "Genetic Resources and Plant Breeding". Novosibirsk, 2017;16-17. (in Russian)

Hoffmann B. Alteration of drought tolerance of winter wheat caused by translocation of rye chromosome segment 1R. Cereal Res. Commun. 2008;36(2):269-278.

Jones P., Keane E.M., Osborne B.A. Effects of alien cytoplasmic variation on carbon assimilation and productivity in wheat. J. Exp. Bot. 1998;49(326):1519-1528.

Keane E.M., Jones P.W. Effects of alien cytoplasm substitution on the response of wheat cultivars to *Septoria nodorum*. Ann. Appl. Biol. 1990;117:299-317.

Klimushina M.V., Divashuk M.G., Karlov G.I., Mokhammed T.A.K., Semenov O.G. Analysis of allelic state of genes responsible for baking properties in allocytoplasmic wheat hybrids. Russ. J. Genet. 2013;49(5):530-538. DOI 10.1134/S1022795413050074.

Liu C.G., Wu Y.W., Hou H., Zhang C., Zhang Y., McIntosh R.A. Value and utilization of alloplasmic common wheats with *Aegilops crassa* cytoplasm. Plant Breed. 2002;121(5):407-410. DOI 10.1046/j.1439-0523.2002.755374.x.

Liu Y., Tang L., Xu Q., Ma D., Zhao M., Sun J., Chen W. Experimental and genomic evidence for the indica-type cytoplasmic effect in *Oryza sativa* L. ssp. *japonica*. J. Integr. Agric. 2016;15(10):2183-2191. DOI 10.1016/S2095-3119(15)61190-X.

Mago R., Spielmeyer W., Lawrence G. Identification and mapping of molecular markers linked to rust resistance genes located on chromosome 1RS of rye using wheat-rye translocation lines. Theor. Appl. Genet. 2002;104(8):1317-1324. DOI 10.1007/s00122-002-0879-3.

Mukai Y., Gill B.S. Detection of barley chromatin added to wheat by genomic *in situ* hybridization. Genome. 1991;34(3):448-452. DOI 10.1139/g91-067.

Pershina L.A., Belova L.I., Devyatkina E.P., Numerova O.M., Shumny V.K. Efficiency of haploid plant raise by anther culturing and remote crossing of cereals. Fiziologiya i Biokhimiya Kulturnykh Rasteniy = Physiology and Biochemistry of Cultivated Plants. 1999b;31(3):196-202. (in Russian)

Pershina L.A., Numerova O.M., Belova L.I., Devyatkina E.P., Shumny V.K. The effect of the genotypic diversity of *Hordeum vulgare* L. and *Triticum aestivum* L. on the crossability and production of partially fertile barley-wheat hybrids. Russ. J. Genet. 1998;34(10):1156-1163.

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. 1999a;35(2):176-183.

Pershina L.A., Osadchaya T.S., Badaeva E.D., Belan I.A., Rosseeva L.P. Androgenesis in anther cultures of cultivars and a promising form of spring common wheat of West Siberia differing in the presence or absence of wheat-alien translocations. Russ. J. Genet.: Appl. Res. 2013;3(4): 246-253. DOI 10.1134/S2079059713040096.

Pershina L.A., Trubacheeva N.V., Sinyavskaya M.G., Devyatkina E.P., Kravtsova L.A. Nuclear-cytoplasmic compatibility and the state of mitochondrial and chloroplast DNA regions in alloplasmic recombinant and introgressive lines (*H. vulgare* )-*T. aestivum*. Russ. J. Genet. 2014;50(10):1017-1024.

Pretorius Z.A., Singh R.P., Wagoire W.W., Payne T.S. Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis* f. sp. *tritici* in Uganda. Plant Dis. 2000;84(2):203. DOI 10.1094/PDIS.2000.84.2.203B.

Rabinovich S.V. Importance of wheat–rye translocations for breeding modern cultivar of *Triticum aestivum* L. Euphytica. 1998;100:323-340.

Schlegel R. Current List of Wheats with Rye and Alien Introgression. Version 05.16. 2018. Available at: http://www.rye-gene-map.de/rye-introgression

Semenov O.G., Divashuk M.G., Haitembu Gerhard Shangeshapwako, Mohammed Tawfeek Ahmed Kaid. DNA marker-assisted creation of allocytoplasmic wheat genotypes with high gluten quality. Vestnik RUDN = Bulletin of Peoples' Friendship University of Russia. 2016;1:7-14. (in Russian)

Semenov O.G., Mohammed Tawfeek Ahmed Kaid. Morphobiological characterization of genotypes of allocytoplasmic spring wheat forms according to their level of tolerance to drought stress. Vestnik RUDN = Bulletin of Peoples' Friendship University of Russia. 2014;2:5-14. (in Russian)

Singh N.K., Shepherd K.W., McIntosh R.A. Linkage mapping of genes for resistance to leaf, steam and stripe rust and ω-secalins on the short arm of rye chromosome 1R. Theor. Appl. Genet. 1990;80(5):609-616. DOI 10.1007/BF00224219.

Soltani A., Kumar A., Mergoum M., Pirseyedi S.M., Hegstad J.B., Mazaheri M., Kianian S.F. Novel nuclear-cytoplasmic interaction in wheat (*Triticum aestivum*) induces vigorous plants. Funct. Integr. Genomics. 2016;16(2):171-182. DOI 10.1007/s10142-016-0475-2.

Suzuki T., Nakamura C., Mori N., Kaneda C. Overexpression of mitochondrial genes in alloplasmic common wheat with a cytoplasm of wheatgrass (*Agropyron trichophorum*) showing depressed vigor and male sterility. Plant Mol. Biol. 1995;27(3):553-565.

Tahmasebi S., Heidari B., Pakniyat H., Dadkhodaie A. Consequences of 1BL/1RS translocation on agronomic and physiological traits in wheat. Cereal Res. Commun. 2015;43(4):554-566.

Talukder S.K., Vara Prasad P.V., Todd T., Babar M.A., Poland J., Bowden R., Fritz A. Effect of cytoplasmic diver-

sity on post anthesis heat tolerance in wheat. Euphytica. 2015;204:383-394. DOI 10.1007/s10681-014-1350-7.

Tsunewaki K. Plasmon analysis as the counterpart of genome analysis. In: Jauhar P.P. (Ed.). Methods of Genome Analysis in Plants. Boca Raton: CRC Press, 1996;271-299.

Tsunewaki K. Fine mapping of the first multi-fertility-restoring gene, $Rf^{multi}$, of wheat for three *Aegilops* plasmons, using 1BS-1RS recombinant lines. Theor. Appl. Genet. 2015;128:723-732. DOI 10.1007/s00122-015-2467-3.

Tsunewaki K. Fine mapping of the first multi-fertility-restoring gene, $Rf^{multi}$, of wheat for three *Aegilops* plasmons, using 1BS-1RS recombinant lines. Theor. Appl. Genet. 2015;128:723-732. DOI 10.1007/s00122-015-2467-3.

Вавиловский журнал генетики и селекции. 2018;22(5):553-559
DOI 10.18699/VJ18.394



# Resource potential of some species of the genus *Miscanthus* Anderss. under conditions of continental climate of West Siberian forest-steppe

O.V. Dorogina[1] ✉, O.Yu. Vasilyeva[1], N.S. Nuzhdina[1], L.V. Buglova[1], Yu.A. Gismatulina[2], E.V. Zhmud[1], G.A. Zueva[1], O.V. Komina[1], E.A. Tsybchenko[3]

[1] Central Siberian Botanical Garden, SB RAS, Novosibirsk, Russia
[2] Institute for Problems of Chemical and Energetic Technologies, SB RAS, Biysk, Russia
[3] Novosibirsk State Agrarian University, Novosibirsk, Russia

In the meantime, search for environmentally friendly renewable energy sources alternative to fossil fuel has been driven by energy security challenges including limited availability of fuel and energy price fluctuations. Therefore herbal perennial grasses with their rapid growth and prominent biomass yield increasingly make it a favorite choice as a valuable agricultural crop usable for cellulosic ethanol production. As an example, the genus *Miscanthus* Anderss. (silvergrass) comprises ca. 14–20 species including *M. sacchariflorus* (Maxim.) Hack., *M. sinensis* Anderss., *M. purpurascens* Anderss, and *M.* × *giganteus*, which appear to be an almost inexhaustible source of sustainable raw material, and several *Miscanthus* species were investigated as a potential biofuel energy crop with commercially viable way of its producing. Introduction and investigation of *Miscanthus* species were initiated in the Central Siberian Botanical Garden of the Siberian Branch of the Russian Academy of Sciences (CSBG SB RAS, Novosibirsk, Russia) based on the grass and ornamental plant collection in the late 1990s. The paper objective is studying the biological traits of three *Miscanthus* species introduced into the CSBG SB RAS, selection and genetic identification of cultivars and varieties as the most perspective agricultural crop. To evaluate the potential crop yield and selection prospects of *Miscanthus* species being competitive as a valuable biofuel energy crop, the authors have estimated seasonal rhythms of model species development in the continental climate conditions of West Siberia. The article characterizes different *Miscanthus* varieties obtained either by the *ex situ* or *in situ* methods; presents the biochemical analysis of plant material and molecular identification of three *Miscanthus* species introduced into the CSBG SB RAS. The seasonal development analysis of three selected varieties of *Miscanthus* (*M. sacchariflorus*, *M. sinensis*, and *M. purpurascens*) proved the hydrometeorological conditions to be advantageous for prominent biomass yield, e. g. contributory to use *Miscanthus* in West Siberia as an easy to grow cellulose-rich grass. Molecular markers applicable in DNA-identification and genetic passportization of *Miscanthus* varieties have been established, which are perspective as such an economically available plant material as alternative non-woody source of cellulose.

Key words: *Miscanthus*; bioenergy; phenology; biomorphology; reproductive biology; chemical composition; DNA sequencing; ITS locus; *trn*L-F intron.

# Ресурсный потенциал некоторых видов рода *Miscanthus* Anderss. в условиях континентального климата лесостепи Западной Сибири

О.В. Дорогина[1] ✉, О.Ю. Васильева[1], Н.С. Нуждина[1], Л.В. Буглова[1], Ю.А. Гисматулина[2], Е.В. Жмудь[1], Г.А. Зуева[1], О.В. Комина[1], Е.А. Цыбченко[3]

[1] Центральный сибирский ботанический сад Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт проблем химико-энергетических технологий Сибирского отделения Российской академии наук, Бийск, Россия
[3] Новосибирский государственный аграрный университет, Новосибирск, Россия

В настоящее время весьма актуальны задачи по поиску альтернативных древесине источников энергии, экологически безопасных и экономически доступных. В связи с этим особый интерес представляют виды травянистых растений с высокой скоростью роста и характеризующиеся высокими значениями нарастания надземной вегетативной массы, имеющие практическое применение в качестве источника биоэтанола. Примером может служить род *Miscanthus* Anderss. (веерник), включающий примерно 14–20 видов, в том числе *M. sacchariflorus* (Maxim.) Hack., *M. sinensis* Anderss. и *M. purpurascens* Anderss., а также *M.*×*giganteus*, которые являются практически неисчерпаемыми источниками возобновляемого сырья в области альтернативной энергетики. В Центральном сибирском ботаническом саду (ЦСБС) СО РАН (Новосибирск) на основе коллекции газонных и декоративных злаков в конце 1990-х гг. было начато формирование и изучение родового комплекса *Miscanthus* Anderss. Целью этого исследования стало изучение биологических особенностей видов *Miscanthus*: *M. sacchariflorus*, *M. sinensis* и *M. purpurascens*, интродуцированных в ЦСБС, отбор и генетическая идентификация перспективных форм в качестве технических сырьевых растений. Для оценки ресурсного потенциала и перспективы селекционной работы с родовым комплексом *Miscanthus* с целью хозяйственного использования в качестве технической (биоэнергетической) культуры в условиях лесостепи Западной Сибири были изучены сезонные ритмы развития модельных видов в условиях континентального климата в сравнении с муссонным и умеренно континентальным; охарактеризованы биоморфы, образующиеся *ex situ* и *in situ*; определен химический состав растительного сырья и проведена идентификация по молекулярно-генетическим мар-

Received 17.04.2018
Accepted for publication 06.06.2018


✉ e-mail: olga-dorogina@yandex.ru

керам трех видов *Miscanthus*, интродуцированных в ЦСБС СО РАН. Анализ сезонного развития трех отборных форм веерников (*M. sacchariflorus*, *M. purpurascens* и *M. sinensis*) показал, что гидротермические условия благоприятствовали получению вегетативной массы растительного сырья, т. е. использованию в качестве технической культуры в условиях лесостепи Западной Сибири. Выявлены формы для дальнейшей селекции и молекулярные признаки для разных видов мискантуса, которые можно использовать для идентификации и паспортизации форм и линий, перспективных для получения экономически доступного растительного сырья – альтернативных источников целлюлозы недревесного происхождения.

Ключевые слова: род *Miscanthus*; биоэнергетика; фенология; биоморфология; репродуктивная биология; химический состав; секвенирование; локус ITS, *trn*L-F.

**HOW TO CITE THIS ARTICLE:**

Dorogina O.V., Vasilyeva O.Yu., Nuzhdina N.S., Buglova L.V., Gismatulina Yu.A., Zhmud E.V., Zueva G.A., Komina O.V., Tsybchenko E.A. Resource potential of some species of the genus *Miscanthus* Anderss. under conditions of continental climate of West Siberian forest-steppe. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):553-559. DOI 10.18699/VJ18.394

According to domestic and foreign scholars, only 5 % of the vascular plants species of the world flora are studied regarding to their utility for humanity (Revin, 2000). Introduction in the culture of food, technical and ornamental plants of local and non-local flora allows reducing the load on resource species in their natural habitats. An integrated approach to solve these problems includes studying the biological and ecological characteristics of species *in situ* and *ex situ*, which has been discussed at the International Congresses of Botanical Gardens (Spain, 2004, Poland, 2007, Ireland, 2010, China, 2012) on repeated occasions.

A special group of objects is represented by species showing high polymorphism in nature and at the same time having a wide range of practical applications in various economic fields. The generic complex of *Miscanthus* Anderss. including *M. sacchariflorus* (Maxim.) Hack., *M. sinensis* Anderss., and *M. purpurascens* Anderss., originally known as ornamental grasses, is unique in this respect. However, over the past two decades, those species, as well as *M.* × *giganteus*, have been referred to the plant objects valued as practically inexhaustible sources of renewable raw materials for pulp production that in turn is a base product for developments in the field of alternative energy (Slynko et al., 2013).

The collection of lawn and ornamental grasses is a part of the bioresource scientific "Collection of Living Plants in the Open and Closed Ground" USU 440534 of the Central Siberian Botanical Garden of the Siberian Branch of the Russian Academy of Sciences (CSBG SB RAS). Forming the collection of living plants and studying the *Miscanthus* family complex started in the late 1990s mainly due to receiving the cloned material from other botanical gardens.

Since 2003, the collection was widened by *Miscanthus* samples from their natural habitats at the Russian Far East. The most significant replenishments occurred in 2013 (samples from the Gamov Peninsula, Primorsky Krai) and 2017 (Chuguev and Khasan districts of Primorsky Krai, the Kuril Islands – Shikotan, Kunashir, Iturup, and Kamchatsky Krai). Selection of promising plant raw materials as a valuable alternative biofuel material being the ultimate research objective, there is a need to study genetic polymorphism and DNA marking in these *Miscanthus* species.

*Miscanthus* is one of the most efficient solar energy accumulators on the planet (Dohleman, Long, 2009). According to physiologists' research, *Miscanthus* species have the highest potential productivity. Obtaining up to 40 tons of dry biomass per hectare is associated with $C_4$-type photosynthesis typical for all representatives of this genus. Unlike most of the traditionally cultivated $C_4$ plants, such as sugar cane and corn, *Miscanthus* are able to maintain a high intensity of photosynthesis even under rather low temperatures (Naidu et al., 2003; Anisimov et al., 2016), which explains the high productivity of these plants grown under more severe climatic conditions than in natural ones.

The prospects of *Miscanthus* growing in the Ob forest-steppe region is conditioned by the fact that $C_4$ photosynthesis is ineffective in low-light terms due to high energy requirements according to some scholars. In *Miskanthus* natural habitats (Vladivostok), the average annual number of sunshine hours is 2140, while in Novosibirsk – 2077, that is much higher than in Moscow (1731) and approximates to the parameters of Krasnodar (2110) and Yalta (2185).

The study objective is to investigate the biological features of *Miscanthus* species (*M. sacchariflorus*, *M. sinensis* and *M. purpurascens*) introduced into the CSBG SB RAS, along with selection and genetic identification of promising forms as technical raw plants.

## Materials and methods

Three species were chosen as model ones, their ecological and biological characteristics being studied since 2012 in the CSBG SB RAS. The introduced samples were collected at the Russian Far East. Taxonomic signs are presented by Voroshilov (1982) as follows:

*M. sacchariflorus*. A lemma has a short straight awn not protruding from a tylosis floccus. The floccus hairs are white, rarely reddish; a rhizome has thin creeping shoots. It grows at the territory of the Russian Federation in Primorye, the western and southern regions of the Amur River basin.

*M. sinensis*. Inflorescence rami are branched; both spikes are at some distance from the axes of spigs. Glumes are asperous with long hairs. Tylosis flocci of a lemma are rather sparse, squarrose after flowering, dirty white, but quite often

reddish as well. The lemma awn is considerably protruding from of a floccus, well visible, genuflexuous. A rhizome is shortened, thick. It occurs on Sakhalin and the Kuril Islands.

*M. purpurascens*. Inflorescence rami are unbranched (sometimes, except for a base itself); a sessile spikelet is located directly on the sprig axis. Glumes are only asperous. Flocci after flowering are often not so squarrose, denser and reddish. A rhizome is long, horizontal. The species inhabits Primorye. Records for Sakhalin and Kuril Islands are erroneous apparently.

*Miscanthus* growing rhythms and development were studied by techniques of phenological observations in the USSR botanical gardens (Methodology…, 1975) modified and supplemented in the course of the research by G.A. Zueva in compliance with objects' peculiarities. Hydrothermal characteristics of vegetation and rest periods were calculated from the data of Novosibirsk State Agrarian University and the weather station Ogurtsovo (Novosibirsk Region), and the climate monitoring data (http://www.pogodaiklimat.ru/) as well.

Biomorphological investigations were based on the data on *Miscanthus* life forms presented by A.B. and T.A. Bezdelevs (Bezdelev, Bezdeleva, 2006) for *M. sacchariflorus* and *M. sinensis*. Characteristics of *M. purpurascensis* life form are not given in this biomorphological report. For morphometric parameters, mean values ($X$), error ($s_x$), and a coefficient of variation ($V$, %) are presented.

The chemical composition of plant samples (leaves) collected in the CSBG SB RAS in the first decade of October was defined at the Laboratory of Bioconversion of the Institute for Problems of Chemical and Energetic Technologies SB RAS (Altaisky Krai).

*Miscanthus* plant specimens were grounded with scissors. Cellulose mass fraction determined by Kurschner method (in terms of absolutely dry raw materials – adrm), acid-insoluble lignin mass fraction (adrm), pentosan mass fraction (adrm), ash content (adrm), mass fraction of extractive substances – fat-waxy fraction (FF) (extractant – dichloromethane, adrm) were carried out according to standardized analysis of plant raw materials (Obolenskaya et al., 1991).

To extract the genomic DNA, CTAB method developed by Doyle and Doyle (1990) with dry plant material (leaves) was applied. The DNA concentration was determined by spectrophotometry using a BioSpectrometer Kinetic and μCuvette G1.0 microcuvettes (Eppendorf, Germany).

Primers ITS4 and ITS5 and the PCR protocol recommended by White et al. (1990) were used to amplify the internal transcribed spacer 1 and 2 regions (ITS1-2) of the nuclear ribosomal DNA. To amplify the *trn*L-F intergenic spacer of chloroplast DNA, universal primers "c" and "f" (Taberlet et al., 2007) were applied, as well as PCR protocol proposed by Amirahmadi et al. (2010). DNA loci amplification was performed on Thermal Cycler C1000 (Bio-Rad, USA). The length of the amplified fragments was analyzed by electrophoretic separation in an agarose gel (1.5 %) in Sub Cell Model 96 chambers under the current (4 V/cm). The amplified fragments were stained with SYBR-Green (Thermo Fischer Scientific, USA); visualizing and video recording of the separated PCR fragments were carried out with Gel-Doc XR + gel documentation and the Bio-Rad ImageLab Software systems.

The obtained individual DNA fragments of ITS1-2 and *trn*L-F intergenic spacers were purified from PCR components by AMPureXP sorption (Agencourt, USA) and sequenced directly using the BigDye Ready Reaction DNA Sequencing Kit v.3.1 kit (Thermo Scientific, USA). The Sanger reaction products were purified from unincorporated fluorescent dyes by centrifugation (900 g, 2 min) with a Sephadex G-50 Fine column (750 μl suspension) (GE Healthcare, USA), and scanned with 3130XL automatic genetic analyzers (Applied Biosystems, USA) at the Collaboration Centre "Genomics" (Novosibirsk). The obtained sequenograms were applied to mark species of the genus *Miscanthus*: *M. sacchariflorus*, *M. sinensis* and *M. purpurascens*. The comparison was made in nucleotide bases pairs (nbp).

MEGA 7.0.25 software (Kumar et al., 2016) was used to align the nucleotide sequences by the Clustal-W method, to calculate p-distances and to determine GC-nucleotides content; a scheme of genetic relationships of the analyzed samples was constructed by the Maximum Parsimony method.

## Results

### *Miscanthus* morphology and phenorhythm peculiarities under the conditions of West Siberian forest-steppe

To study the phenorhythmics in *Miscanthus* three species, the hydrothermal character of vegetation periods in various regions was analyzed (Table 1): (1) natural habitat zone (Vladivostok, mid-latitude monsoon climate); (2) region with a long-term successful experience of *Miscanthus* raw materials commercial production (Penza, mid-latitude moderate-continental climate); (3) experimental sites in the risk crop farming zone (Novosibirsk, mid-latitude continental climate).

Table 1 shows that for their spring regrowth and development *Miscanthus* gain from the thermal regime in conditions of Penza and even Novosibirsk (May), however, in late September–October this advantage over natural habitats is leveled. It is the heat accumulation in the autumn months that contributes to a full-fledged generative development and seed production in Primorsky Krai.

When comparing the temperature and solid precipitation in Novosibirsk, Penza and Vladivostok in winter, it was revealed that unlike the warmer growing conditions in Vladivostok in September–October, the winter temperature regime in natural habitats is more severe in December–January as opposed to Penza (Table 2). Compared to Novosibirsk, Vladivostok in winter has less snow precipitation, but this does not damage wintering reproduction buds *in situ*. Thus, 100 % winter hardiness of the genus *Miscanthus* representatives in Novosibirsk is explained by the historically developed and genetically fixed adaptive capabilities of the Far Eastern plants.

Studying the seasonal development of plants introduced into the culture makes it possible to draw more reliable conclusions on prospects, or, on the contrary, the inadequate stability of species and varieties, only if the long-term research period includes either dry vegetative season or the one with heavy winter precipitation, or extreme overwinter. For six-year long observations of *M. sacchariflorus*, *M. sinensis*, *M. purpurascens* in the forest-steppe conditions of West Siberia, it has been noted the following:

**Table 1.** Hydrothermal character of the vegetative season in geographical regions of *Miscanthus* cultivation, representing three types of climate

| Indicator | April | May | June | July | August | September | October | November |
|---|---|---|---|---|---|---|---|---|
| Novosibirsk, type of climate: continental | | | | | | | | |
| Mean maximum, °C | 7.8 | 18.4 | 22.9 | 25.0 | 22.4 | 15.6 | 7.2 | –3.7 |
| Average temperature, °C | 2.3 | 11.7 | 16.8 | 19.2 | 16.5 | 10.1 | 2.9 | –7.1 |
| Mean minimum, °C | –2.0 | 6.0 | 11.4 | 14.1 | 11.5 | 5.9 | –0.3 | –10.0 |
| Precipitation norm, mm | 28 | 34 | 50 | 72 | 49 | 42 | 46 | 38 |
| Penza, type of climate: moderate continental | | | | | | | | |
| Mean maximum, °C | 12.1 | 20.7 | 24.5 | 26.4 | 24.6 | 18.1 | 9.7 | 0.6 |
| Average temperature, °C | 6.5 | 13.9 | 18.0 | 19.9 | 17.9 | 12.1 | 5.4 | –2.3 |
| Mean minimum, °C | 1.5 | 7.5 | 12.0 | 13.8 | 12.0 | 7.2 | 1.9 | –4.7 |
| Precipitation norm, mm | 33 | 41 | 63 | 59 | 50 | 52 | 46 | 45 |
| Vladivostok, type of climate: monsoon climate | | | | | | | | |
| Mean maximum, °C | 9.4 | 14.7 | 18.1 | 21.9 | 24.1 | 20.2 | 13.4 | 3.3 |
| Average temperature, °C | 4.9 | 9.8 | 14.0 | 18.3 | 20.8 | 16.4 | 9.4 | –0.3 |
| Mean minimum, °C | 1.9 | 6.7 | 11.4 | 16.2 | 18.5 | 13.5 | 6.4 | –3.1 |
| Precipitation norm, mm | 62 | 95 | 89 | 164 | 145 | 92 | 59 | 41 |

**Table 2.** Hydrothermal character of the winter period in the geographical regions of *Miscanthus* cultivation in three climate types

| Indicator | December | January | February | March |
|---|---|---|---|---|
| Novosibirsk, continental climate | | | | |
| Mean maximum, °C | –10.1 | –12.2 | –9.8 | –2.1 |
| Average temperature, °C | –13.9 | –16.1 | –14.3 | –7.0 |
| Mean minimum, °C | –17.5 | –19.6 | –17.9 | –11.1 |
| Precipitation norm, mm | 31 | 26 | 15 | 17 |
| Penza, moderate continental climate | | | | |
| Mean maximum, °C | –4.4 | –5,5 | –5.1 | 1.0 |
| Average temperature, °C | –7.4 | –8.7 | –9.1 | –3.4 |
| Mean minimum, °C | –10.3 | –11.9 | –12.5 | –7.2 |
| Precipitation norm, mm | 41 | 38 | 31 | 35 |
| Vladivostok, monsoon climate | | | | |
| Mean maximum, °C | –6.0 | –8.2 | –4.1 | 2.1 |
| Average temperature, °C | –9.9 | –12.2 | –8.5 | –1.9 |
| Mean minimum, °C | –12.6 | –15.3 | –11.3 | –4.7 |
| Precipitation norm, mm | 26 | 10 | 25 | 35 |

– in 2012, the vegetation period was abnormally dry and hot, there was 38 % of the precipitation norm in June, and only 6 % in July; an average monthly air temperature exceeded the average annual one by 4.7 °C in June, and by 3.1 °C in July;
– in 2013, the vegetation period had excess moisture, the precipitation exceeded an average long-term data being as much as 115 % in July, 284 % in August (171 mm), and 135 % in September;
– the early wintering in 2013 (November) was relatively mild, especially considering the almost twofold excess of snow precipitation;
– in 2016 the early wintering was extreme, when in November during the entire second decade, the minimum air temperatures were below –20 °C, and even –25 °C for six days, with a slight excess of mean annual solid precipitation;
– the wintering in 2012–2013 was the harshest one (first of all, for phanerophytes, not for hemicryptophytes, which group the involved plants refers to), when 24 days in December were characterized by a minimum air temperature below –25 °C, of which 10 days were the coldest (from –30.0 to –41.5 °C).

However, the severe temperature conditions did not significantly affect *M. sacchariflorus*, *M. sinensis* and *M. purpurascens* winter hardiness.

It was found that during all the years of observation, the three plant species of *Miscanthus* had the same set of phenophases of vegetative and generative development (regrowth, tubing, earing and flowering). The exception was 2013, when the earing and flowering phases did not occur. It's explained not so much by the excessive precipitation leading to a turbulent vegetative development of plants similar to the Far Eastern monsoon climate, but lower air temperatures fixed in Novosibirsk at the same time. Probably, the heat shortage slowed the intrarenal differentiation of generative organs to such an extent that the plants couldn't reach even the earing phase before the first autumn frosts.

*M. sacchariflorus* stably differs from *M. sinensis* and *M. purpurascens* by earlier (5–10 days) regrowth; the advance persists even after entering the other phenophases. It should be noted that in its natural habitats of Primorsky Krai this species occupies the northernmost areas (e. g. Sikhote-Alin).

**Table 3.** Morphometric parameters of vegetative organs of *Miscanthus* in West Siberian forest-steppe (*ex situ*)

| Length of vegetative shoot, cm | | Number of leaves, pcs. | | Leaf width*, cm | | Leaf length*, cm | | Straw diameter, cm | |
|---|---|---|---|---|---|---|---|---|---|
| $X \pm s_x$ | *V*, % | $X \pm s_x$ | *V*, % | $X \pm s_x$ | *V*, % | $X \pm s_x$ | *V*, % | $X \pm s_x$ | *V*, % |
| *M. purpurascens* | | | | | | | | | |
| 163.1 ± 4.3 | 13.1 | 10.5 ± 0.1 | 5.6 | 1.5 ± 0.1 | 13.4 | 57.2 ± 1.1 | 9.6 | 0.47 ± 0.02 | 21.80 |
| *M. sinensis* | | | | | | | | | |
| 213.0 ± 7.2 | 13.2 | 11.3 ± 0.2 | 6.4 | 2.1 ± 0.1 | 16.3 | 66.4 ± 1.8 | 10.3 | 0.51 ± 0.02 | 16.24 |
| *M. sacchariflorus* | | | | | | | | | |
| 220.7 ± 0.3 | 12.5 | 11.4 ± 0.2 | 7.1 | 2.1 ± 0.1 | 16.9 | 67.5 ± 2.2 | 10.9 | 0.51 ± 0.03 | 16.33 |

* The fourth top leaf.

**Table 4.** Chemical composition of *Miscanthus* samples ($X \pm s_x$, %) introduced in the CSBG SB RAS, Novosibirsk

| Samples | Cellulose by Kurschner* | Lignin* | Pentosans* | Ash* | FF |
|---|---|---|---|---|---|
| *M. purpurascens* | 48.4 ± 1.0 | 23.1 ± 0.5 | 21.7 ± 0.5 | 4.36 ± 0.05 | 2.2 ± 0.5 |
| *M. sinensis* | 49.1 ± 1.0 | 23.3 ± 0.5 | 20.7 ± 0.5 | 3.00 ± 0.05 | 2.6 ± 0.5 |
| *M. sacchariflorus* | 53.3 ± 1.0 | 28.1 ± 0.5 | 21.3 ± 0.5 | 5.66 ± 0.05 | 2.4 ± 0.5 |

Note: * In terms of absolutely dry raw material; FF – fat-waxy fraction.

The earliest entry into generative phases was marked for all the three species in 2014, which was characterized by early spring warming and thereafter by earlier soil thawing.

The most references (Greef et al., 1997; Nishiwaki et al., 2011; Gifford et al., 2015) devoted to various aspects of research and utilization of *Miscanthus* species, forms, hybrids and varieties point out that the concerned gene pool collection requires serious systematic verification.

To a large extent, the taxonomic accessory of vascular plants is traditionally determined by the quantitative and qualitative characteristics of generative organs and aboveground vegetative ones. However, in the *Miscanthus* species, the structure of the underground organs should also be taken into account. So, while studying *M. sacchariflorus* biomorphology in the CSBG SB RAS, it was revealed that the species life form in the Ob forest-steppe conditions matches its life form description made by A.B. and T.A. Bezdelevs (Bezdelev, Bezdeleva, 2006) for natural habitats in Primorye: a perennial summer-green herbaceous thin-long-rhizome sympodially growing polycarpic with an elongated erect shoot.

These authors described two biomorphs for *M. sinensis*. However, in our opinion, their description of the life form "a perennial summer-green herbaceous thick-long-rhizome loose-tussock sympodially growing polycarpic with a semi-rosellate erect shoot" is more relevant to *M. purpurascens*, while "a perennial summer-green herbaceous short-rhizome loose-tussock sympodially growing polycarpic with a semi-rosellate erect shoot" is referred to *M. sinensis*, for which particularization is character under local conditions.

The dynamics of shoot formation and morphometric parameters in culture were compared with shoots development of *M.* × *giganteus* under conditions of mid-latitude moderate continental climate (Penza). It is believed that one of the parental forms of *Miscanthus giganteus* is *M. sinensis* (Table 3).

Table 3 shows that the length of generative shoots in *Miscanthus* three species varies within 163–220 cm. In Penza the range of generative shoots of *Miscanthus giganteus* was 160–207 cm (Gushchina et al., 2016). The species studied in Novosibirsk (CSBG SB RAS) and Penza are similar in this parameter, as well as in the leaves width.

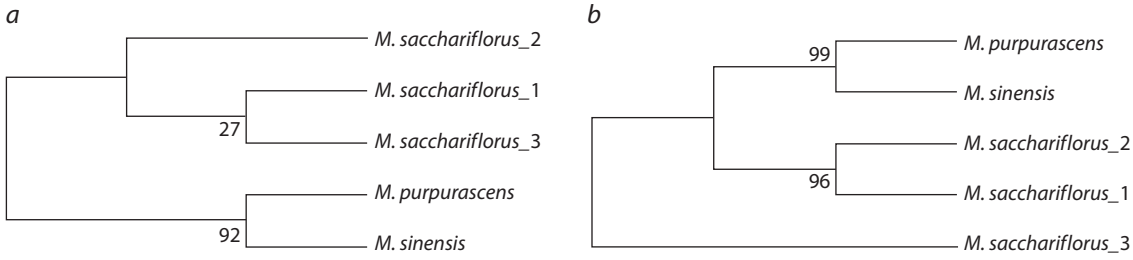
It should be noted that the stems and leaves characteristics are of particular importance when using *Miscanthus* as a technical crop. As these studies were carried on within the framework of SB RAS Integration Project "Fundamental basics of ethylene production from *Miscanthus*" (http://www.sbras.ru/files/files/pril_pso-230_15-08-17.pdf), a special attention was paid to obtaining raw materials and further analysis of its chemical composition for the content of cellulose, as well as lignin, pentosans, ash and fat-waxy fraction, which results are tabulated in Table 4.

As it turns out, *M. sinensis* and *M. purpurascens* are close not only by habitat regions, morphological features, but in chemical composition as well. *M. sacchariflorus* reveals higher values of both cellulose and lignin. Lignin high content can be associated with the greater ash content (6 %). In commercial production, it is caused by the higher concentration of soil acid-insoluble salts, which may result from the excessive mineral fertilizer application to obtain the maximum yield. However, to process such biomass into cellulose is much more difficult. It is noteworthy that under the CSBG SB RAS conditions, all the three *Miscanthus* species were grown on aligned agrarian background.

### *Miscanthus* species genetic marking

To determine the genetic polymorphism and DNA marking in *Miscanthus* species, the variability of the length and nucleotide composition of ITS1-2 nDNA locus sequences and *trn*L-F chlDNA intergenic spacer were studied in three species: *M. sacchariflorus*, *M. sinensis* and *M. purpurascens*.

Based on the analysis of the ITS1-2 internal transcribed spacer of nuclear ribosomal DNA of *Miscanthus*, the variability of this region length was revealed, comprising




Genetic relatedness between *Miscanthus* representatives based on the ITS (*a*) and the *trn*L-F (*b*) data sets using Maximum Parsimony method.

The bootstrap values are given in the nodes.

440 bp in *M. purpurascens*, 486 bp – *M. sinensis* and 656–671 bp – *M. sacchariflorus*, correspondingly. The length of *trn*L-F xlDNA intergenic spacer chlDNA in *Miscanthus* varied in a smaller range: from 893 bp (*M. sinensis*) to 962 bp (*M. sacchariflorus*_2). The resulting *trn*L-F sequences are marked by a low content of G and C nucleotides (an average of 31.4 % for a locus) while for the ITS locus an average GC-nucleotide counts 61 %. ITS sequences are characterized by the presence of polyguanine fragments; *trn*L-F locus sequence – of a large number of poly-A and poly-T-blocks.

According to the sequencing data on nuclear (ITS) and chloroplast DNA (*trn*L-F) loci, the schemes were created (see the Figure), which reflect genetic polymorphism and related connections between representatives of the genus *Miscanthus*.

Both schemes reliably reveal the close relationship between *M. sinensis* and *M. purpurascens* (the bootstrap support value is over 90 %). Based on the analysis of different marker loci, two different models of clustering *M. sacchariflorus* samples were identified. According to sequencing ITS region, all three samples form a common clade (see the Figure, *a*). At the same time, a sample of *M. sacchariflorus*_3 and two samples of *M. sacchariflorus*_1 and *M. sacchariflorus*_2 form two separate clades according to the *trn*L-F locus analysis (see the Figure, *b*).

## Discussion

The hydrothermal conditions of all vegetative periods (2012–2017), when the seasonal development of three selected forms of *Miscanthus* was studied (one of each species – *M. sacchariflorus*, *M. purpurascens* and *M. sinensis*), were favorable to the vegetative mass of plant raw materials production, i. e. their usage as a technical culture. Assessing the ornamental qualities of the objects revealed the only vegetative period of 2013, when not any sample entered the generative phase due to excess moistening. Neither earing nor flowering (when impressive thyrses develop being ornamental even over the winter snow) was observed.

Morphometric parameters of vegetative (technologically significant) Miscanthus organs in West Siberian forest-steppe were not inferior to the ones under the conditions of the moderate continental climate, varying at a low or medium level, thus making it possible to predict the yield (raw stock) rather precisely.

A general morphological peculiarity of *M. sacchariflorus* (which occupies northernmost habitats in monsoon and moderate continental climates) is the structure of underground organs, namely the thin-long-rhizome biomorph. Phenorhythm types and life forms of *M. sacchariflorus* and *M. sinensis* correspond to those under the natural conditions of Primorsky Krai. This testifies their rather high adaptive capabilities, and the prospects for selection of the most technologically (including biochemically) productive forms taking into account the intraspecific polymorphism.

Based on the analysis of the genetic relationships of the genus *Miscanthus* representatives and the sequencing data of two ITS1-2 and *trn*L-F loci, it should be concluded that obtained results are quite consistent in toto. The different nature of the phylogenetic trees branching can be conditioned by various types of inheritance (biparental and uniparental) character for nuclear and chloroplast markers, respectively, chosen to construct trees.

The revealed molecular features of *Miscanthus* different species can be used to identify and certify the *Miscanthus* forms and lines promising as available plant material, suitable for perspective environmentally safe alternative biofuel production.

## Acknowledgements

The investigations were carried out according to the complex program of fundamental scientific research of SB RAS II.1. The projects block "Studying the morphogenetic potential of the genus *Miscanthus* Anderss. and the possibilities of its use in introduction and selection under conditions of West Siberia", the state registration AAAA-A17-117120640059-4. The materials were taken from the bioresource scientific "Collection of living plants in open and closed ground" USU 440534.

## Conflict of interest

The authors declare no conflict of interest.

## References

Amirahmadi S., Kazempour Osaloo S., Maassoumi A.A. Loss of chloroplast *trn*$L_{UAA}$ intron in two species of *Hedysarum* (Fabaceae): evolutionary implications. Iran. J. Biotechnol. 2010;8(3):150-155.

Anisimov A.A., Hohlov N.F., Tarakanov I.G. Photoperiodic regulation of development in various Miscanthus species (*Miscanthus* spp.). Izvestiya Timiryazevskoy Selskokhozyaystvennoy Akademii = Izvestiya of the Timiryazev Agricultural Academy. 2016;(6):56-72. (in Russian)

Bezdelev A.B., Bezdeleva T.A. Life Forms of Seed Plants of the Russian Far East. Vladivostok: Dalnauka Publ., 2006. (in Russian)

Dohleman F.G., Long S.P. More productive than maize in the midwest: how does Miscanthus do it? Plant. Physiol. 2009;150(4):2104-2115. DOI 10.1104/pp.109.139162.

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

Gifford J.M., Chae W.B., Juvik J.A., Swaminathan K., Moose S.P. Mapping the genome of *Miscanthus sinensis* for QTL associated with biomass productivity. GCB Bioenergy. 2015;7(4):797-810. DOI 10.1111/gcbb.12201.

Greef J.M., Deuter M., Jung C., Schondelmaier J. Genetic diversity of European *Miscanthus* species revealed by AFLP fingerprinting. Genet. Resour. Crop Evol. 1997;44(2):185-195. DOI 10.1023/A:1008693214629.

Guschina V.A., Agapkin N.D., Borisova E.N. Adaptation of giant miscanthus first year of life to the conditions of the middle Volga region. Proceedings of the 3rd International Scientific and Practical Conference "Problems and Monitoring of Natural Ecosystems". 2016;14-18. (in Russian)

Kumar S., Stecher G., Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2016;33(7):1870-1874. DOI 10.1093/molbev/msw054.

Methodology of Phenological Observations in botanical gardens of the USSR: Board of botan. gardens of the USSR. Moscow, 1975;18-27. (in Russian)

Naidu S.L., Moose S.P., Al-Shoaibi A.K., Raines C.A., Long S.P. Cold tolerance of $C_4$ photosynthesis in *Miscanthus* × *giganteus*: adaptation in amounts and sequence of $C_4$ photosynthetic enzymes. Plant Physiol. 2003;132:1688-1697. DOI 10.1104/pp.103.021790.

Nishiwaki A., Mizuguti A., Kuwabara S., Toma Y., Ishigaki G., Miyashita T., Yamada T., Matuura H., Yamaguchi S., Rayburn A.L., Akashi R., Stewart J.R. Discovery of natural *Miscanthus* (Poaceae) triploid plants in sympatric populations of *Miscanthus sacchariflorus* and *Miscanthus sinensis* in southern Japan. Am. J. Bot. 2011; 98(1):154-159. DOI 10.3732/ajb.1000258.

Obolenskaya A.V., Elnitskaya Z.P., Leonovich A.A. Laboratory Course on the Chemistry of Wood and Cellulose. Moscow, 1991. (in Russian)

Revin P. Speech at the XVIth International Botanical Congress. Bulletin of the Board of botanic gardens of Russia and the branch of the Botanic Gardens Conservation International. 2000;11:38-47. (in Russian)

Slynko N.M., Goryachkovskaya T.N., Shekhovtsov S.V., Bannikova S.V., Burmakina N.V., Starostin K.V., Rozanov A.S., Nechiporenko N.N., Veprev S.G., Shumny V.K., Kolchanov N.A., Peltek S.E. The biotechnological potential of the new crops, Miscanthus cv. Soranovsky. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2013;17(4/1):765-771. (in Russian)

Taberlet P., Coissac E., Pompanon F., Gielly L., Miquel C., Valentini A., Vermat T., Corthier G., Brochmann C., Willerslev E. Power and limitation of the chloroplast *trn*L (UAA) intron for plant DNA-barcoding. Nucleic Acids Res. 2007;35(3):e14. DOI 10.1093/nar/gkl938.

Voroshilov V.N. Field guide to Soviet Far East plants. Moscow: Nauka, 1982. (in Russian)

White T.J., Bruns T., Lee S., Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR Protocols: A Guide to Methods and Applications. 1990;18(1):315-322. DOI 10.1016/b978-0-12-372180-8.50042-1.

Вавиловский журнал генетики и селекции. 2018;22(5):560-567
DOI 10.18699/VJ18.395

# Тенденция преодоления устойчивости к бурой ржавчине интрогрессивных линий мягкой пшеницы с генетическим материалом *Aegilops speltoides* Tausch

Л.Я. Плотникова[1] ✉, Л.В. Мешкова[2], Е.И. Гультяева[3], О.П. Митрофанова[4], И.Ф. Лапочкина[5]

[1] Омский государственный аграрный университет имени П.А. Столыпина, Омск, Россия
[2] Омский аграрный научный центр, Омск, Россия
[3] Всероссийский научно-исследовательский институт защиты растений, Санкт-Петербург, Пушкин, Россия
[4] Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений имени Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
[5] Московский научно-исследовательский институт сельского хозяйства (Немчиновка), Московская область, Россия

Развитие бурой ржавчины, вызываемой грибом *Puccinia triticina* Erikss., приводит к существенным потерям урожая мягкой пшеницы *Triticum aestivum* L. Вид *Aegilops speltoides* Tausch (2*n* = 14, геном SS) считается перспективным источником генов для защиты пшеницы от болезней. Целью исследований был мониторинг устойчивости образцов *Ae. speltoides*, интрогрессивных линий и сортов мягкой пшеницы с генетическим материалом этого вида к западносибирской популяции возбудителя бурой ржавчины. Оценка устойчивости растений в условиях естественного инфекционного фона показала, что образцы *Ae. speltoides* были иммунны к бурой ржавчине, однако наблюдается тенденция повышения восприимчивости интрогрессивных линий и сортов пшеницы. Отмечено снижение защитного действия генов *Lr28*, *Lr36* и *Lr35*, но *Lr47* сохраняет свою эффективность в западносибирском регионе, что подтверждено оценкой вирулентности омской популяции *P. triticina* к перечисленным генам. В период исследований (2003–2017 гг.) была преодолена резистентность линий Од 26/89, 156/90, аналогов сорта Новосибирская 67 АНК-39 (B, C) и линии Л-500 из коллекции «Арсенал». Линии Од (35/1, 35/89, 210/90, 278/89), АНК-39 (A, D, E), Л-501, сорта Челяба 75 и Mit оставались высокоустойчивыми в течение всего периода проведения испытаний. По данным молекулярного анализа и фитопатологического тестирования, в изученных образцах (за исключением линий сорта Thatcher и Pavon) не выявлены известные гены *Lr28*, *Lr36* и *Lr47* от *Ae. speltoides*. На основании данных, полученных с помощью ДНК-маркеров, сделано предположение о том, что сорт Челяба 75, семь устойчивых линий серий Од и АНК-39 содержат транслокации с геном *LrSp*. В остальных образцах вероятно наличие дополнительных неидентифицированных генов устойчивости *Ae. speltoides*. Тенденцию снижения устойчивости к бурой ржавчине интрогрессивных линий и сортов с генами *Ae. speltoides* необходимо учитывать при селекции мягкой пшеницы, предназначенной для выращивания в Западной Сибири.

Ключевые слова: *Triticum aestivum*; *Aegilops speltoides*; интрогрессивные линии; *Puccinia triticina*; вирулентность; *Lr*-гены; молекулярные маркеры.

## A tendency towards leaf rust resistance decrease in common wheat introgression lines with genetic material from *Aegilops speltoides* Tausch

L.Ya. Plotnikova[1] ✉, L.V. Meshkova[2], E.I. Gultyaeva[3], O.P. Mitrofanova[4], I.F. Lapochkina[5]

[1] Omsk State Agrarian University named after P.A. Stolypin, Omsk, Russia
[2] Omsk Agrarian Scientific Center, Omsk, Russia
[3] All-Russian Institute of Plant Protection, St. Petersburg, Pushkin, Russia
[4] Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
[5] Moscow Agricultural Research Institute (Nemchinovka), Moscow region, Russia

Leaf rust, caused by the fungus *Puccinia triticina* Erikss., inflicts serious crop loss of common wheat *Triticum aestivum* L. The species *Aegilops speltoides* Tausch (2*n* = 14, SS) is considered a promising issue of genes to protect bread wheat from diseases. The objective of this study was the monitoring of resistance to leaf rust of *Ae. speltoides* accessions and introgressive lines and cultivars with genetic material of this species to the Western Siberian population of fungus. The estimation of specimens in the field conditions on natural infectious background showed that the *Ae. speltoides* accessions were immune to leaf rust, however, a tendency towards increasing susceptibility of the introgressive lines and varieties was detected. The protective effect of the known genes *Lr28*, *Lr36* and *Lr35* decreased, but *Lr47* remained efficient in West Siberia, as confirmed by the results of testing of the Omsk population of *P. triticina* for virulence to the mentioned genes. During the study (2003-2017) the resistance has been overcome of lines Od 26/89, 156/90, analogs of cv. Novosibirskaya 67 – ANK-39 (B, C), and L-500 from the "Arsenal" collection. High resistance to leaf rust was preserved of lines Od (35/1, 35/89, 210/90, 278/89); ANK-39 (A, D, E); L-501 and cvs. Chelyaba 75 and Mit. Analysis of DNA markers and phytopathological tests showed that the studied variet-

Received 19.01.2018
Accepted for publication 09.05.2018


✉ e-mail: lya.plotnikova@omgau.org

ies and lines lacked the known genes *Lr28*, *Lr36*, and *Lr47* from *Ae. speltoides* (except for the lines of cvs. Thatcher and Pavon). On the base of analysis of DNA markers, it was assumed that Chelyaba 75 and seven resistant lines from the Od- and ANK-39-series have translocations bearing the *LrSp* gene. Presumably, the rest of samples possess additional not yet identified genes of *Ae. speltoides*. The trend of overcoming of resistance to leaf rust of introgressive lines and varieties with *Ae. speltoides* genes should be taken into consideration in common wheat breeding in Western Siberia.

Key words: *Triticum aestivum*; *Aegilops speltoides*; introgressive lines; *Puccinia triticina*; virulence; *Lr*-genes; molecular markers.

**КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:**

Плотникова Л.Я., Мешкова Л.В., Гультяева Е.И., Митрофанова О.П., Лапочкина И.Ф. Тенденция преодоления устойчивости к бурой ржавчине интрогрессивных линий мягкой пшеницы с генетическим материалом *Aegilops speltoides* Tausch. Вавиловский журнал генетики и селекции. 2018;22(5):560-567. DOI 10.18699/VJ18.395

**HOW TO CITE THIS ARTICLE:**

Plotnikova L.Ya., Meshkova L.V., Gultyaeva E.I., Mitrofanova O.P., Lapochkina I.F. A tendency towards leaf rust resistance decrease in common wheat introgression lines with genetic material from *Aegilops speltoides* Tausch. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):560-567. DOI 10.18699/VJ18.395 (in Russian)

Бурая ржавчина, вызываемая грибом *Puccinia triticina* Erikss., распространена во всех регионах возделывания мягкой пшеницы *Triticum aestivum* L. и приводит к ежегодным – 5–10 %, а в годы эпифитотий – 50–70 % потерям урожая (Eversmeyer, Kramer, 2000). Для стабилизации фитопатологической обстановки в агроценозах необходимо повышать разнообразие сортов по генам устойчивости, в связи с чем в селекционных учреждениях планомерно проводят работу по переносу генетического материала других видов в геном пшеницы. В Каталоге символов генов пшеницы описано более 80 генов устойчивости к бурой ржавчине (гены *Lr*) (McIntosh et al., 2013).

Введение чужеродных генов в сорта – длительный процесс, поскольку интрогрессии, особенно множественные, резко ухудшают хозяйственно ценные признаки растений (Friebe et al., 1996; Salina et al., 2015). Жесткие требования к урожайности и качеству зерна пшеницы вынуждают селекционеров использовать ограниченный набор чужеродных генов *Lr* (Мартынов и др., 2015). Молекулярно-генетический скрининг сортов из Государственного реестра селекционных достижений, допущенных к использованию (2005–2013 гг.), показал, что большая их часть, преимущественно возделываемых в Поволжье, Южно-Уральском и Западно-Сибирском регионах, защищена двумя генами, *Lr19* и *Lr9*, от *Agropyron elongatum* (Host.) Beuv. и *Aegilops umbellulata* Zhuk. соответственно.

Для расширения генетической базы селекции пшеницы перспективными источниками генов устойчивости к болезням считаются виды рода *Aegilops* L. секции Sytopsis, прежде всего вид *Aegilops speltoides* Tausch ($2n = 2x = 14$, геном SS). Результаты фитопатологических оценок, в том числе многолетних, образцов этого вида, представленных в коллекциях различных стран мира, показали, что большинство из них проявляет иммунитет или высокую устойчивость к местным популяциям возбудителя бурой ржавчины (Gill et al., 1985; Shah et al., 2000; Михайлова, 2006; Anikster et al., 2005; Тырышкин и др., 2012; Holubec et al., 2014). В настоящее время в геном мягкой пшеницы от *Ae. speltoides* перенесены шесть генов устойчивости, из них *Lr28*, *Lr36*, *Lr47*, *Lr51*, *Lr66* проявляются на стадии проростков, а *Lr35* – у взрослых растений (McIntosh et al., 2013).

В России имеется богатый опыт предселекционной работы с образцами *Ae. speltoides*. Во Всероссийском институте генетических ресурсов растений имени Н.И. Вавилова (ВИР, Санкт-Петербург) были получены интрогрессивные линии с групповой резистентностью к грибным болезням и гаметоцидным геном *Gc* (Одинцова и др., 1991). В Московском научно-исследовательском институте сельского хозяйства (Немчиновка) создана коллекция интрогрессивных линий пшеницы Арсенал с генами диких видов, включая *Ae. speltoides* (Лапочкина, 2005). В Национальном центре зерна им. П.П. Лукьяненко получены устойчивые к ржавчинным болезням линии на основе геномно-замещенной формы пшеницы Авродес (тетра-Аврора × *Ae. speltoides*) (Давоян и др., 2012). Первый российский сорт мягкой пшеницы с генетическим материалом *Ae. speltoides* – Челяба 75 (с геном *LrSp*) – включен в Государственный реестр селекционных достижений, допущенных к использованию, в 2012 г. (Тюнин и др., 2017).

Для перспективной селекции важна информация о степени устойчивости к болезни источников генов *Ae. speltoides* в основных регионах возделывания пшеницы, а также о генетическом контроле признака. В связи с этим целью исследований был мониторинг устойчивости к бурой ржавчине образцов *Ae. speltoides*, интрогрессивных линий и сортов мягкой пшеницы, полученных с участием этого вида, к западносибирской популяции бурой ржавчины.

## Материалы и методы

**Растительный материал.** В исследования были включены образцы вида *Ae. speltoides* (кк-2, 12, 43, 452, 453) и 20 сортов и линий мягкой пшеницы с его генами из коллекции генетических ресурсов растений Всероссийского института генетических ресурсов растений им. Н.И. Вавилова (ВИР). Известные гены от *Ae. speltoides* несли линии

яровой мягкой пшеницы сорта Thatcher Tc*Lr28* (RL-6079, транслокация T4AS.4AL-7S#2S), Tc*Lr35* (RL-6083, T2BS-2SS.2SL), Tc*Lr36* (ER84018, 6BS), сорт Pavon (*Lr47*) (к-44748, 7AS-7S#1S-7AS.7AL). Интрогрессивные линии серии Од (26/89, 35/1, 35/89, 69/89, 156/90, 210/90, 278/89) созданы И.Г. Одинцовой с сотрудниками на основе амфидиплоида *T. dicoccum* × *Ae. speltoides*, часть из них несла гаметоцидный ген *Gc*, приводящий к элиминации пыльцы с рецессивными аллелями устойчивости (линии-кукушки), в других он отсутствовал (Одинцова и др., 1991). Аналоги сорта пшеницы Новосибирская 67 АНК-39(A–E) (кк-65529, 65528, 65524, 65525, 65526 соответственно) получены С.Ф. Ковалем на базе разных линий Од (Коваль и др., 2001). Сорт яровой пшеницы Челяба 75 создан с участием линии Од (линия-кукушка) и несет ген *LrSp* (Тюнин и др., 2017). Линии Л-500 и Л-501 (кк-62903, 62904) получены И.Ф. Лапочкиной (2005). Сорт озимой пшеницы Mit (к-58063, США) имеет в родословной *Ae. speltoides* (Genetic Resources Information System…, http://www.wheatpedigree.net/). В качестве стандартов восприимчивости использовали сорта яровой мягкой пшеницы Памяти Азиева (среднеранний), Новосибирская 67 (среднеспелый), Серебристая (среднепоздний).

**Фитопатологические оценки.** Развитие бурой ржавчины на посевах оценивали в лесостепной зоне юга Западной Сибири (г. Омск) на естественном инфекционном фоне в 2003–2017 гг. Массовое развитие бурой ржавчины в зоне наблюдается в конце июля–августе, когда у растений формируется и созревает зерно. Тип реакции растений на заражение в полевых и лабораторных условиях определяли по пятибалльной шкале: 0 – иммунитет, без симптомов; 0; – некрозы без пустул; 1 – устойчивый, очень мелкие пустулы с некрозом; 2 – умеренно устойчивый, пустулы мелкие или средние, с некрозом; 3 – умеренно восприимчивый, пустулы средней величины без некроза; 4 – восприимчивый, пустулы крупные, часто сливающиеся (Mains, Jackson, 1926). Реакцию 0–2 балла считали устойчивой, 3–4 балла – восприимчивой. Степень поражения растений (в %) оценивали по сравнительной шкале (Peterson et al., 1948).

**Изучение популяции *P. triticina*.** Спорообразцы для анализа популяции *P. triticina* собирали в 2013–2016 гг. на сортах яровой мягкой пшеницы в разных районах Омской области. Изучение популяции по вирулентности проводили путем определения реакций проростков линий, содержащих гены *Lr28*, *Lr36*, *Lr47*, на заражение изолятами гриба. Анализ аллельного состава генов вирулентности выполняли по методу Л.А. Михайловой (2006). Частоту вирулентных клонов в популяции (в %) определяли по результатам анализа 300–400 изолятов.

**Молекулярно-генетические исследования.** Для постулирования генов *Lr28*, *Lr35*, *Lr47* и *Lr66* у образцов пшеницы использовали ДНК-маркеры SCS421; Sr39≠22r; PS10; S13-R16 соответственно. Дополнительно в анализ включили маркеры генов, широко распространенных в российских сортах пшеницы: *Lr9* (маркер SCS5), *Lr19* (SCS265), *Lr20* (STS628), *Lr24* (Sr24≠12, Sr24≠50), *Lr26* (SCM9), *Lr39* (=*Lr41*) (GDM39), *Lr34* (L34DIN9/L34Plus), *Lr37* (Ventriup/LN2). ДНК выделяли из трех листьев 5–7-дневных проростков пшеницы по методике Д.Б. Дорохова и Э. Клоке (1997). Амплификацию ДНК проводили по ранее описанным протоколам (Marais et al., 2010; Гультяева и др., 2017; http://maswheat.ucdavis.edu). Для постулирования генов устойчивости в образцах пшеницы дополнительно провели фитопатологический тест с использованием тест-клонов патогена, маркированных вирулентностью к *Lr9*, *Lr19* и *Lr26* (Михайлова, 2006).

## Результаты

В период 2003–2017 гг. отмечено ежегодное интенсивное развитие бурой ржавчины на посевах (поражение восприимчивых сортов 70–100 %), за исключением сезонов с длительной засухой (2008, 2010 и 2012 гг.).

В связи с усилением развития бурой ржавчины и преодолением эффективности большинства известных генов устойчивости (Плотникова и др., 2015) в 2014–2017 гг. в исследования был включен набор образцов *Ae. speltoides*. Линии с генетическим материалом *Ae. speltoides* были получены в 1980–2000 гг. и проявляли высокую устойчивость к бурой ржавчине в районах их создания (Одинцова и др., 1991; Лапочкина и др., 1996; Коваль и др., 2001; McIntosh et al., 2013). Все они проявляли иммунитет на стадии проростков (данные не приводятся) и взрослых растений в полевых условиях в Омской области (таблица).

Линия Tc*Lr28* с генетическим материалом *Ae. speltoides* длительное время сохраняла иммунитет, но в 2011 г. было отмечено появление единичных пустул, а в 2013 и 2017 гг. степень поражения повысилась до 30 %. Устойчивость линии Tc*Lr36* в 2003–2009 гг. варьировала, а с 2011 г. симптомы болезни отмечались ежегодно (5–30 %). Среди известных генов *Ae. speltoides* самым эффективным остается *Lr47*, но с 2013 г. на растениях сорта Pavon, донора этого гена, регулярно отмечается появление пустул патогена (степень поражения 1–10 %). Ген *Lr35* определяет возрастную устойчивость к ржавчине, его действие усиливается с фазы колошения и слабо зависит от условий среды (Плотникова, Штубей, 2009). В 2003–2014 гг. даже при интенсивных вспышках болезни поражение Tc*Lr35* не превышало 5–30 %, но в 2016–2017 гг. оно усилилось до 40–50 %, что свидетельствует о снижении эффективности гена.

В начале 90-х гг. XX в. все линии И.Г. Одинцовой были иммунны к популяции бурой ржавчины в Омской области. В 1998 г. была преодолена устойчивость Од 69/89 (не опубликовано), в 2014 г. – Од 156/90, в 2017 г. – Од 26/89 (20 %), остальные четыре линии сохраняют иммунитет до настоящего времени. Серия линий АНК-39(A–E) была получена на основе линий Од. В полевых условиях они длительное время проявляли иммунитет, но в 2014 г. была преодолена резистентность АНК-39C, а в 2017 г. – АНК-39B. Остальные линии серии АНК-39 (A, D, E) сохраняют высокую устойчивость (поражение не более 5–10 %) (см. таблицу).

В 2014 г. набор интрогрессивных линий был дополнен сортами Челяба 75, Mit и линиями Л-500 и Л-501 из коллекции «Арсенал». По данным фитопатологических оценок 2014–2017 гг., сорт Челяба 75 и линия Л-501 сохраняют высокую устойчивость к бурой ржавчине, поражение сорта Mit варьировало от умеренного (30 % в 2016 г.) до слабого (1–10 % в 2014, 2017 гг.), а линия Л-500 оказалась

Results of the field estimation of leaf rust severity in *Ae. speltoides* accessions, introgression lines, and common wheat varieties with *Ae. speltoides* genes (score / %), 2003–2017

| Variety or accession* | Years | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2003 | 2004 | 2005 | 2006 | 2007 | 2009 | 2011 | 2013 | 2014 | 2015 | 2016 | 2017 |
| Pamyati Azieva | 4/70 | 4/70 | 4/90 | 4/70 | 4/90 | 4/80 | 4/80 | 4/100 | 4/80 | 4/20 | 4/80 | 4/100 |
| Saratovskaya 29 | 4/80 | 4/70 | 4/100 | 4/70 | 4/100 | 4/100 | 4/80 | 4/80 | 4/90 | 4/50 | 4/100 | 4/100 |
| Novosibirskaya 67 | 4/70 | 4/80 | 4/100 | 4/80 | 4/100 | 4/100 | 4/90 | 4/90 | 4/100 | 4/70 | 4/100 | 4/100 |
| Serebristaya | – | – | – | – | – | – | – | 4/100 | 4/90 | 4/80 | 4/100 | 4/100 |
| k-2 *Ae. speltoides* | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 0/0 |
| k-12 *Ae. speltoides* | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 0/0 |
| k-43 *Ae. speltoides* | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 0/0 |
| k-452 *Ae. speltoides* | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 0/0 |
| k-453 *Ae. speltoides* | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 0/0 |
| Tc*Lr28* | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/5 | 4/20 | 4/10 | 4/5 | 3/20 | 4/20 |
| Tc*Lr35* | 4/10 | 4/5 | 4/30 | 4/20 | 4/30 | 4/20 | 4/20 | 4/10 | 4/30 | 4/10 | 4/40 | 4/50 |
| Tc*Lr36* | 3/5 | 0 | 0 | 4/10 | 4/10 | 0/0 | 4/10 | 4/30 | 4/10 | 4/5 | 4/20 | 4/20 |
| Pavon *Lr47* | – | – | – | – | – | – | 0/0 | 410 | 4/5 | 4/1 | 4/1 | 4/5 |
| Od 35/1 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 |
| Od 35/89 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 |
| Od 210/90 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 |
| Od 278/89 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 |
| Od 26/89 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/1 | 4/5 | 4/10 | 4/20 |
| Od 156/90 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/5 | 4/60 | 4/80 | 4/70 | 4/60 |
| Od 69/89 | 4/90 | 4/70 | 4/100 | 4/70 | 4/100 | 4/80 | 4/80 | 4/80 | 4/80 | 4/80 | 4/80 | 4/ 80 |
| ANK-39A | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/5 | 0/0 | 4/10 |
| ANK-39V | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/5 | 4/10 | 4/40 |
| ANK-39S | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/1 | 4/30 | 4/70 | 4/40 | 4/100 | 4/100 |
| ANK-39D | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 0/0 | 4/10 |
| ANK-39E | 2/10 | 0 | 2/5 | 2/5 | 2/10 | 0/0 | 0/0 | 0/0 | 2/5 | 0/0 | 0/0 | 4/5 |
| L-500 | – | – | – | – | – | – | – | – | 4/40 | 4/10 | 4/70 | 4/100 |
| L-501 | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 0/0 |
| Chelyaba 75 | – | – | – | – | – | – | – | – | 0/0 | 0/0 | 0/0 | 4/5 |
| Mit | – | – | – | – | – | – | – | – | 4/1 | 4/5 | 4/30 | 4/10 |

*k, accessions from the VIR collection; Tc, near isogenic lines of var. Thatcher; Od, I.G. Odintsova's lines; ANK-39, immune analogs of var. Novosibirskaya 67; L, lines from the Arsenal collection.

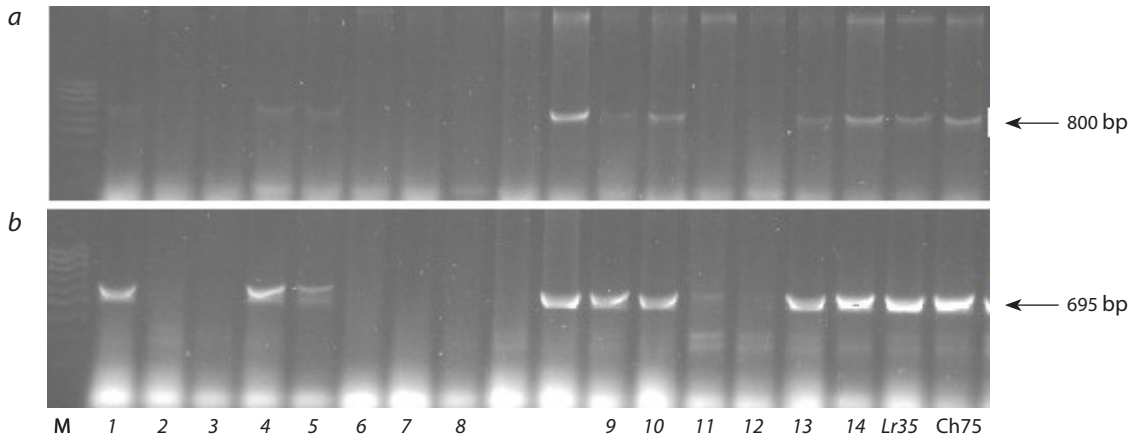
восприимчива к болезни. В целом, согласно результатам оценок различных интрогрессивных линий и сортов с генами *Ae. speltoides*, отмечено постепенное повышение их восприимчивости к болезни, особенно усилившееся в 2013–2017 гг.

Для подтверждения полученных результатов был изучен состав популяции *P. triticina* по вирулентности к генам *Ae. speltoides Lr28*, *Lr36*, *Lr47*. Установлено, что значительная доля изолятов *P. triticina* несла аллели вирулентности *pp28*, *pp36*, *pp47*. При этом в период 2013–2016 гг. доля клонов (в среднем по области) с аллелями *pp28* возросла в 2.5 раза (с 15.6 до 38.4 %), а с *pp36* и *pp47* – снизилась в 2.1 и 3.3 раза (с 30.1 до 19.7 % и с 15.1 до 4.5 % соответственно).

В результате молекулярного анализа у изучаемых образцов пшеницы не выявлено генов *Lr9*, *Lr19*, *Lr24*, *Lr37*, *Lr39*, *Lr20*, *Lr34 Lr28* и *Lr47*, что согласуется с данными фитопатологического тестирования. У сорта Челяба 75 (с геном *LrSp*), линий Од (210/90, 278/89, 35/1, 35/89) и АНК-39 (A, E, D), сохранивших устойчивость к бурой ржавчине, показано наличие ампликона размером 800 п. н. для маркера Sr39≠22r и 695 п. н. для маркера S13-R16 (рисунок).

## Обсуждение

Мягкая пшеница возделывается в форме монокультуры на обширной площади в Поволжье, на Южном Урале, в Западной Сибири, Алтайском крае, Северном Казахстане.




Electrophoretic resolution of PCR products of wheat accessions with markers (*a*) Sr39≠22r and (*b*) S13-R16.

Lanes: M, molecular weight ladder 100 bp (Fermentas); *1*, ANK-39E; *2*, Mit; *3*, ANK-39V; *4*, ANK-39A; *5*, ANK-39D; *6*, ANK-39S; *7*, L-500; *8*, L-501; *9*, Od 210/90; *10*, Od 278/89; *11*, Od 26/89; *12*, Od 156/90; *13*, Od 35/1; *14*, Od 35/89; *Lr35*, line Tc*Lr35*; Ch75, var. Chelyaba 75.

В 70–90-х гг. XX в. вспышки бурой ржавчины в этих регионах происходили каждые четыре-шесть лет из десяти (Санин, 1997). В последние десятилетия регулярное развитие бурой ржавчины и усиление ее вредоносности отмечены в Поволжье и на Южном Урале (Маркелова и др., 2014; Тюнин и др., 2017). Проведенные нами исследования показали, что на юге Западной Сибири в 2003–2017 гг. сильное развитие ржавчины (поражение сортов до 70–100 %) происходило в 12 из 15 сезонов, т. е. частота вспышек болезни существенно возросла.

Западносибирская популяция *P. triticina* весьма агрессивна, ее клоны несут 13–15 генов вирулентности как к генам устойчивости, введенным в коммерческие сорта (*Lr9*), так и к не использованным в селекции (Плотникова и др., 2015). Это вынуждает включать в селекционный процесс новые гены устойчивости, а также проводить мониторинг их эффективности в регионе.

С учетом усиления развития бурой ржавчины пшеницы в Западной Сибири была проведена оценка устойчивости пяти образцов *Ae. speltoides* и установлено, что они сохранили иммунитет в 2013–2017 гг. Ранее в работах других авторов было показано, что образцы *Ae. speltoides* иммунны к популяциям *P. triticina* Северо-Западного и Северо-Кавказского регионов (Тырышкин и др., 2012; Гультяева и др., 2014; Волкова и др., 2016), что свидетельствует о стабильной устойчивости вида *Ae. speltoides* к болезни.

Возрастная устойчивость считается неспецифической, а гены *Lr13* и *Lr34* послужили основой для создания сортов пшеницы с длительной устойчивостью (durable resistance) к бурой ржавчине (Parlevliet, 2002). В наших экспериментах ген *Lr35* обеспечивал устойчивость взрослых растений до 2015 г., но в 2016–2017 гг. поражение линии Tc*Lr35* достигло 40–50 %, что демонстрирует снижение его эффективности в Омской области.

Анализируя итоги 15-летнего изучения интрогрессивных линий и сортов с генами *Ae. speltoides*, следует отметить тенденцию снижения их резистентности к болезни. Суммарная доля среднеустойчивых и восприимчивых образцов (поражение от 20 до 100 %) в 2017 гг. возросла до 42 %, что позволяет предположить накопление в популяции *P. triticina* генов, вирулентных к генам *Ae. speltoides*.

Наши данные об эффективности генов устойчивости *Ae. speltoides* к бурой ржавчине несколько отличаются от результатов, полученных ранее в других регионах. Ген *Lr28* обеспечивал высокую устойчивость растений в Ленинградской и Челябинской областях (Гультяева и др., 2014; Тюнин и др., 2017). В Новосибирской области до 2016 г. были эффективны гены *Lr28* и *Lr35* (Сочалова, Пискарев, 2017). Такие различия могут быть объяснены комплексом причин: различиями популяций *P. triticina*, наборами возделываемых сортов (включая озимые), особенностями климата, значительными расстояниями между регионами и т. д. В Новосибирской области появление новых рас патогенов отмечалось позже, чем в Омской области. Так, накопление вирулентных к гену *Lr9* клонов в Омской области произошло в 2007 г., а в Новосибирской – в 2008 г. Аналогично первая за долгие годы вспышка стеблевой ржавчины в Омской области зарегистрирована в 2008 г., а в Новосибирской – в 2010–2011 гг. (Мешкова и др., 2008; Сочалова, 2016).

На основании результатов молекулярного анализа и фитопатологических тестов нами установлено, что в исследуемых образцах отсутствуют известные гены *Lr28*, *Lr36*, *Lr47* от *Ae. speltoides*. Это позволяет предположить, что устойчивость изученных образцов обеспечивается неизвестными генетическими локусами. Однако до настоящего времени хромосомная локализация этих генов не изучена. В линиях Л-500 и Л-501, созданных путем прямой гибридизации мягкой пшеницы с *Ae. speltoides*, показано наличие трех неизвестных рецессивных *Lr*-генов (Лапочкина и др., 1996).

Сведения о генетическом контроле резистентности к бурой ржавчине линий И.Г. Одинцовой отсутствуют. У всех линий серии АНК-39 (А–Е), созданных на основе линий Од, показан моногенный доминантный контроль устойчивости к болезни (Плотникова, Кузьмина, 2017). Исходя из родословной устойчивого сорта Mit, в его создании принимал участие *Ae. speltoides*. По данным разных источников, в этом сорте присутствует также ком-

бинация генов *Lr1*, *Lr10*, *Lr34* и, возможно, *Lr13* (Genetic Resources Information System…, http://www.wheatpedigree.net/). Однако первые два гена не эффективны в Западной Сибири (Плотникова и др., 2015), присутствие *Lr34* не подтверждено нами с помощью молекулярного анализа, а *Lr13* определяет количественную устойчивость растений (4 балла/20 %) (Плотникова, Штубей, 2009). Не исключено, что резистентность сорта Mit обеспечивается генами, унаследованными от *Ae. speltoides*. В будущем необходимо изучить хромосомную локализацию интрогрессированных из генома *Ae. speltoides* локусов у образцов, использованных в настоящем исследовании.

В нашей работе для сортов Челяба 75, четырех линий Од и трех линий АНК-39 (A, D, E) получены сходные результаты молекулярного маркирования, что позволяет предположить наличие у них одинаковых чужеродных локусов резистентности. Ранее было установлено, что маркеры для идентификации генов *Lr35* и *Lr66* не являются строго специфичными (Гультяева и др., 2014). Маркер Sr39≠22r амплифицирует фрагменты не только у образцов *Ae. speltoides*, линии Tc*Lr35* и сорта Челяба 75, но и у образцов других видов – *T. timopheevii*, *Ae. tauschii*. При валидации маркера S13-R16 для гена *Lr66*, локализованного в хромосоме 3A, было показано, что он выявляется только у линии Tc*Lr35* и сорта Челяба 2. С помощью набора молекулярно-цитологических методов (C-бэндинга и FISH) ранее было установлено, что сорт Челяба 75 имеет транслокацию *2DS.2SL* с геном *LrSp* (Адонина и др., 2016). Следует отметить, что у некоторых линий серий Од, АНК-39 и Л-501 не выявлены фрагменты, специфичные для маркеров Sr39≠22r и S13-R16, несмотря на то, что растения долгое время сохраняли устойчивость. Это подтверждает предположение о наличии в материале дополнительных неидентифицированных генов устойчивости. О возможности присутствия неизвестных генов устойчивости *Ae. speltoides* в генотипах интрогрессивных линий сообщали и другие авторы (Давоян и др., 2012; Маркелова и др., 2014; Миков и др., 2016).

В связи с усилением поражаемости ржавчиной интрогрессивных образцов встает вопрос об источнике генов вирулентности. Как правило, основной причиной преодоления генов устойчивости считают возникновение и размножение вирулентных клонов на широко возделываемых сортах с соответствующим геном. Однако первый сорт пшеницы с геном *Ae. speltoides* – Челяба 75 (*LrSp*) – был рекомендован для внедрения в производство на Южном Урале только в 2012 г., а в других российских сортах гены этого вида не установлены (Тюнин и др., 2017). Ранее в Западной Сибири были выявлены также клоны, вирулентные к ювенильным генам устойчивости *T. timopheevii*, хотя материал этого вида в сортах отсутствует (Плотникова и др., 2015). Одним из объяснений может быть занос инфекции из соседних регионов. В Казахстане естественные фитоценозы занимают значительные территории, в них достаточно широко представлены виды рода *Aegilops*, включая *Ae. crassa* (геном DMS) (Есимбекова и др., 2015). Не исключено, что в естественных фитоценозах произрастает и промежуточный хозяин *P. triticina – Isopyrum fumarioides*. Можно предположить, что гены вирулентности к видам рода *Aegilops* возникают в клонах гриба, существующих на диких злаках, а затем постепенно переносятся в Западно-Сибирский регион.

В последние годы появились сообщения об активном использовании генов *Ae. speltoides* в селекционных программах различных учреждений России, созданы первые устойчивые сорта с его генами (Челяба 75 и Уральская кукушка (Давоян и др., 2012; Маркелова и др., 2014; Сибикеев и др., 2015; Адонина и др., 2016; Тюнин и др., 2017). Следует отметить, что линии И.Г. Одинцовой и коллекции «Арсенал» использовались в качестве источников устойчивости в регионах, расположенных на пути перемещения циклонических масс (Поволжье, Южный Урал, Западная Сибирь). Известно, что при возникновении благоприятных условий редкие вирулентные фенотипы могут стремительно размножиться и преодолеть устойчивость сортов, возделываемых на большой площади. Это произошло с сортами Аврора и Кавказ (ген *Lr26*) в Северо-Кавказском регионе в 1973 г., а также с сортами, защищенными геном *Lr9*, в Омской области в 2007 г. (Воронкова и др., 1975; Мешкова и др., 2008). Поскольку ржавчинные грибы способны к дальним воздушным миграциям, то использование генетически однородных по устойчивости сортов в соседних регионах весьма опасно.

Таким образом, многолетний мониторинг развития бурой ржавчины на растениях с генами *Ae. speltoides* в Западной Сибири показал, что устойчивость значительной части (42 %) интрогрессивных линий и сортов мягкой пшеницы была частично или полностью преодолена патогеном. На основании полевых наблюдений установлено снижение защитного действия известных генов *Lr28*, *Lr36* и *Lr35*, хотя ген *Lr47* сохраняет высокую эффективность в регионе. Высокую устойчивость к бурой ржавчине продемонстрировали часть линий Од 35/1, 35/89, 210/90, 278/89), линии АНК-39 (A, D, E) и сорт Челяба 75, линия Л-501 и сорт Mit. Анализ материала с помощью ДНК-маркеров и фитопатологических тестов не доказал присутствие в нем известных генов *Lr28*, *Lr36*, *Lr47*. В линиях серий Од и АНК-39, сохранивших устойчивость к болезни, предполагается наличие транслокации с геном *LrSp*. В остальных линиях и сорте Mit возможны дополнительные неизвестные гены от *Ae. speltoides*. Тенденцию преодоления генов устойчивости к бурой ржавчине, интрогрессированных от *Ae. speltoides*, необходимо учитывать при селекции сортов мягкой пшеницы для Западной Сибири и соседних регионов.

## Conflict of interest

The authors declare no conflict of interest.

## References

Adonina I.G., Leonova I.N., Badaeva E.D., Salina E.A. Genotyping of hexaploid wheat varieties from different Russian regions. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2016;20(1):44-50. DOI 10.18699/VJ16.107. (in Russian)

Anikster Y., Manisterski J., Long D.L., Leonard K.J. Resistance to leaf rust, stripe rust, and stem rust in *Aegilops* ssp. in Israel. Plant Dis. 2005;89:303-308. DOI 10.1094/PD-89-0303.

Davoyan E.R., Davoyan R.O., Bebyakina I.V., Davoyan O.R., Zubanova Yu.S., Zinchenko A.N., Kravchenko A.M. Identification of leaf-rust resistance genes in species of *Aegilops* L., synthetic forms,

and introgression lines of common wheat. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2012;16(1):116-122. (in Russian)

Dorohov D.B., Cloke E. Rapid and economic technology of RAPD-analysis of plant genomes. Molekulyarnaya Genetika = Molecular Genetics. 1997;3(4):443-450. (in Russian)

Esimbekova M.A., Bulatova K.M., Kushanova R.Zh., Mukin K.B. Biodeverssity of wild cereals of the *Aegilops* L. genius in Kazakhstan in the context of wheat breeding. Izvestiya Timiryazevskoy Selskokhozyaystvennoy Akademii = Izvestiya of Timiryazev Agricultural Academy. 2015;6:5-18. (in Russian)

Eversmeyer M.G., Kramer C.L. Epidemiology of wheat leaf and stem rust in the central great plains of the USA. Annu. Rev. Phytopathol. 2000;38:491-513.

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

Genetic Resources Information System for Wheat and Triticale. Available online at http://www.wheatpedigree.net/

Gill B.S., Sharma H.C., Raupp W.J., Browder L.E., Hatchett J.H., Harvey T.L., Moseman J.G., Waines J.G. Evaluation of Aegilops species for resistance to wheat powdery mildew, wheat leaf rust, Hessian fly, and greenbug. Plant Dis. 1985;69(4):314-316.

Gultyaeva E.I., Aristova M.K., Shaidayuk E.L., Mironenko N.V., Kazartsev I.A., Akhmetova A., Kosman E. Genetic differentiation of *Puccinia triticina* Erikss. in Russia. Russ. J. Genet. 2017;53(9):998-1005. DOI 10.1134/S1022795417070031.

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

Holubec V., Hanzalova A., Dumalasova V., Bartos P. Aegilops conservation and collection evaluation in the Czech Republic. J. Syst. Evol. 2014;52(6):783-789. DOI 10.1111/jse.12117.

Koval S.F., Koval V.S., Shamanin V.P. Isogenic Lines of Wheat. Omsk: Omskblankizdat Publ., 2001. (in Russian)

Lapochkina I.F. Alien genetic variability and its importance in plant breeding. The identified plant gene pool and breeding. St. Petersburg: Vavilov Institute of General Genetics Publ., 2005;684-739. (in Russian)

Lapochkina I.F., Solomatin D.A., Grishina E.E., Vishnyakova Kh.S., Pukhalskiy V.A. Common wheat lines with genetic material from *Aegilops speltoides* Tausch. Russ. J. Genetics. 1996; 32(12):1438-1442.

Mains E.B., Jackson H.S. Physiological specialization in the leaf rust wheat *Puccinia triticina* Erikss. Phytopathology. 1926;16(1): 89-95.

Marais G.F., Bekker T.A., Eksteen B., McCallum T., Marais A.S. Attempts to remove gametocidal genes co-transferred to common wheat with rust resistance from *Aegilops speltoides.* Euphytica. 2010;171:71-85. DOI 10.1007/s10681-009-9996-2.

Markelova T.S., Ivanova O.V., Baukenova E.A., Naryshkina E.A., Salmova M.F. Creation of donors of multiple resistance to fungal diseases of spring common wheat by gene introgression from wild species and wheat relatives. Agrarnyy Vestnik Yugo-Vostoka = Agricultural Bulletin of South-East. 2014;1-2:25-27. (in Russian)

Martynov S.P., Dobrotvorskaya T.V., Mitrofanova O.P. Genealogical analysis of the use of aegilops (*Aegilops* L.) genetic material in wheat (*Triticum aestivum* L.) Russ. J. Genet. 2015;51(9):855-862. DOI 10.1134/s1022795415090070.

McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers J., Morris C., Appels R., Xia X.C. Catalogue of gene symbols for wheat. 12th Int. Wheat Genet. Symp. 8-13 September 2013, Yokohama, Japan. Available online at http://www.shigen.nig.ac.jp/wheat/

Meshkova L.V., Rosseeva L.P., Shreider E.R., Sidorov A.V. Virulence of pathotypes of the wheat leaf rust agent to Th*Lr9* in regions of Siberia and the Urals. 2nd All-Russian Conf. "Modern Problems of Plant Immunity to Harmful Organisms". St. Petersburg, 29 September–2 October 2008. St. Petersburg, 2008;70-73. (in Russian)

Mikhailova L.A. Genetics of the Relationships between the Leaf Rust Agent and Wheat. St. Petersburg: VIR Publ., 2006. (in Russian)

Mikov D.S., Davoyan E.R., Zubanova Yu.S., Davoyan R.O., Boldakov D.M. Molecular marker-assisted identification of leaf rust resistance genes *Lr28*, *Lr35*, *Lr36*, *Lr47*, *Lr51* in accessions of Aegilops speltoides, the synthetic form Avrodes, and its derivatives. Proceedings of the 9th All-Russian Conference of Young Scientists. 2016;88-89. (in Russian)

Odintsova I.G., Agafonova N.A., Boguslavsky R.L. Introgressive lines of common wheat with resistantance to leaf rust transferred from *Aegilops speltoides*. Source material and problems of wheat and triticale breeding. Trudy po Prikladnoy Botanike, Genetike i Selektsii = Proceedings on Applied Botany, Genetics, and Breeding. 1991;142:106-110. (in Russian)

Parlevliet J.E. Durability of resistance against fungal, bacterial and viral pathogens; present situation. Euphytica. 2002;124:147-156.

Peterson R.F., Campbell A.B., Hannah A.E. A diagrammatic scale for estimating rust intensity of leaves and stem of cereals. Can. J. Res. 1948;Sect. C. 26:496-500.

Plotnikova L.Ya., Kuzmina S.P. Genetic and physiological bases of resistance to leaf rust in introgressive lines of Novosibirskaya 67 variety with genes from *Aegilops speltoides* Tausch. Proceedings of the 2nd All-Russian (national) scientific conference "The Role of Agrarian Science in the Sustainable Development of Rural Areas", Novosibirsk, December 25, 2017. Novosibirsk, 2017;100-102. (in Russian)

Plotnikova L.Ya., Pozherukova V.E., Meshkova L.V., Mitrofanova O.P., Degtyarev A.I., Aidosova A.T. The resistance of Timofeevi wheat to *Puccinia triticina* in West Siberia. Mikologiya i Fitopatologiya = Mycology and Phytopathology. 2015;49(2):116-125. (in Russian)

Plotnikova L.Ya., Stubey T.Yu. Appearance of resistance to Puccinia triticina determined by the genes *Lr13*, *Lr22b*, and *Lr35* in adult plants. Mikologiya i Fitopatologiya = Mycology and Phytopathology. 2009;43(3):258-271. (in Russian)

Salina E.A., Adonina I.G., Badaeva E.D., Stasyuk A.I., Leonova I.N., Shishkina A.A., Divashuk M.G., Stankova E.V., Khuat T.M., Syukov V.V., Karlov G.I. A *Thinopyrum intermedium* chromosome in bread wheat cultivars as a source of genes conferring resistance to fungal diseases. Euphytica. 2015;204:91-101. DOI 10.1007/s10681-014-1344-5.

Sanin S.S. Phytosanitary monitoring: current state and ways of improvement. Problems of Optimization of the Phytosanitary Condition of Plant Growing. Proceedings of the All-Russian Congress on Plant Protection, St. Petersburg, December 1995. St. Petersburg, 1997;166-176. (in Russian)

Shah T.M., Ahmed J., Asghar M., Iqbal N., Farooq S. Evaluation of annual wild grass species for leaf rust resistance. Pak. J. Biol. Sci. 2000;3(3):469-472.

Sibikeev S.N., Voronina S.A., Badaeva E.D., Druzhin A.E. Study of resistance to leaf and stem rust in *Triticum aestivum – Aegilops speltoides* lines. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2015;19(2):165-170. (in Russian)

Sochalova L.P. Sources of wheat resistance genes to leaf and stem pathogens in the Novosibirsk region. Zernovoe Khozjaistvo Rossii = Grain Economy of Russia. 2016;2:45-49. (in Russian)

Sochalova L.P., Piskarev V.V. Resistance of spring common wheat varieties to agents of infectious diseases in the changing climate of West Siberia. Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Techniques of Agroindustrial Complex. 2017;31(2):21-25. (in Russian)

Tyryshkin L.G., Syukov V.V., Zakharov V.G., Zuev E.V., Gashimov M.E., Kolesova M.A., Chikida N.N., Yershova M.A., Belousova M.H. Sources of effective resistance of common wheat and its relatives to fungal diseases: search, creation, and use in breeding. Trudy po Prikladnoy Botanike, Genetike i Selektzii = Proceedings on Ap-

plied Botany, Genetics, and Breeding. 2012; 170:187-201. (in Russian)

Tyunin V.A., Shreyder E.R., Gultyaeva E.I., Shaydayuk E.L. Characteristics of virulence of *Puccinia triticina* populations and the potential of the *Lr24*, *Lr25*, *LrSp* genes for spring common wheat breeding in the Southern Ural. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2017;21(5):523-529. DOI 10.18699/VJ17.269. (in Russian)

Volkova G.V., Kremneva O.Yu., Shumilov Yu.V., Gladkova E.V., Vaganova O.F., Mitrofanova O.P., Lysenko N.S., Chikida N.N., Khakimova A.G., Zuev E.V. Immunologic estimation of accessions of wheat, its rare species, and Aegilops from the collection of N.I. Vavilov All-Russian Institute of Plant Genetic Resources and selection of sources with group resistance. Vestnik Zaschity Rastenii = Bulletin of Plant Protection. 2016;89(3):38-39. (in Russian)

Voronkova A.A., Dubonosov T.S., Panarin I.V. Causes of leaf rust outbreaks in the Krasnodar region. Rust of Cereals. Moscow: Kolos Publ., 1975;80-88. (in Russian)

Вавиловский журнал генетики и селекции. 2018;22(5):568-573
DOI 10.18699/VJ18.396

# *Drosophila melanogaster* inhabiting northern regions of European Russia are infected with *Wolbachia* which adversely affects their life span

N.V. Roshina[1, 2], A.V. Symonenko[2], A.V. Krementsova[2, 3], E.A. Tsybul'ko[2], V.E. Alatortsev[2], E.G. Pasyukova[2] ✉, D.V. Mukha[1] ✉

[1] Vavilov Institute of General Genetics, RAS, Moscow, Russia
[2] Institute of Molecular Genetics, RAS, Moscow, Russia
[3] Emanuel Institute of Biochemical Physics, RAS, Moscow, Russia

*Wolbachia* is a genus of bacteria causing intracellular infection in the natural populations of *Drosophila melanogaster* on all continents. In *D. melanogaster*, *Wolbachia* affects various life history traits, behaviour, sensitivity to stress and viral infection. The phenotypic effects of *Wolbachia* might evolve to promote its further spreading, increasing the interest in exploring the spread of *Wolbachia*, in particular, at the boundaries of the *D. melanogaster* habitat, in association with the effects on vital traits of host species. In this paper, we present data on the level of *Wolbachia* infection in two *D. melanogaster* populations from the northern regions of European Russia: Alexandrov (56.41° N, 38.72° E) and Valday (58.02° N, 33.24° E). The flies were collected in private apple gardens located in two small hamlets without supermarkets or fruit markets, from 2010 to 2015. The both populations demonstrated the same level of infection: in average, 69.7 % of the inbred lines (ILs) obtained from single females of the Alexandrov population and 68.4 % of ILs obtained from single females of the Valday population. The infection rate varied from year to year showing a tendency to reduction, its overall level being within the range previously observed in other habitats. Life spans were compared in sub-lines of the same IL, one infected with *Wolbachia* and the other treated with tetracycline healing this infection. In four out of five ILs, the lifespan of both males and females was severely affected by *Wolbachia*; in different ILs, the mean life spans reduced from 1.8 to 5.4 times and from 1.4 to 2.4 times, respectively. Our results confirm that, despite *D. melanogaster* widespread distribution, the *Wolbachia* effect on their life span has been mostly negative.

Key words: *Drosophila melanogaster*; *Wolbachia* sp.; PCR, natural populations; symbiont; life span.

**HOW TO CITE THIS ARTICLE:**
Roshina N.V., Symonenko A.V., Krementsova A.V., Tsybul'ko E.A., Alatortsev V.E., Pasyukova E.G., Mukha D.V. *Drosophila melanogaster* inhabiting northern regions of European Russia are infected with *Wolbachia* which adversely affects their life span. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):568-573. DOI 10.18699/VJ18.396

Received 06.04.2017
Accepted for publication 10.05.2018


✉ e-mail: egpas@rambler.ru; dmitryVmukha@gmail.com

# *Drosophila melanogaster*, обитающие на севере европейской части России, заражены *Wolbachia*, негативно влияющей на их продолжительность жизни

Н.В. Рощина[1, 2], А.В. Симоненко[2], А.В. Кременцова[2, 3], Е.А. Цыбулько[2], В.Е. Алаторцев[2], Е.Г. Пасюкова[2] ✉, Д.В. Муха[1] ✉

[1] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[2] Институт молекулярной генетики Российской академии наук, Москва, Россия
[3] Институт биохимической физики им. Н.М. Эмануэля Российской академии наук, Москва, Россия

*Wolbachia* – род бактерий, вызывающих внутриклеточную инфекцию и широко распространенных в природных популяциях *Drosophila melanogaster* на всех континентах. Вольбахии оказывают влияние на ряд признаков *D. melanogaster*, определяющих ее приспособленность, а также на поведение мух, их чувствительность к стрессу и вирусной инфекции. Влияние вольбахии на фенотип вида-хозяина могло способствовать расселению бактерии, что увеличивает интерес к дальнейшему изучению ее географического распространения, в частности на границах среды обитания *D. melanogaster*, в сочетании с воздействием на жизненно важные признаки вида-хозяина. В настоящей работе приведены данные об уровне зараженности вольбахией двух популяций *D. melanogaster* из северных регионов Европейской России. Мух собирали с 2010 по 2015 г. в частных яблоневых садах, расположенных в двух небольших деревнях без супермаркетов или фруктовых рынков, недалеко от городов Александров (56.41° с. ш., 38.72° в. д.) и Валдай (58.02° с. ш., 33.24° в. д.). Обе популяции были заражены вольбахией на сходном уровне. В среднем было заражено 69.7 % инбредных линий (ИЛ), полученных от индивидуальных самок, выловленных близ Александрова, и 68.4 % ИЛ, полученных от индивидуальных самок валдайской популяции. Уровень заражения менялся из года в год, демонстрируя тенденцию к снижению, при этом в среднем находился в пределах, ранее наблюдавшихся в других местообитаниях. Сравнивалась продолжительность жизни в сублиниях одних и тех же ИЛ: одна сублиния оставалась зараженной вольбахией, а другая была вылечена тетрациклином. В четырех из пяти протестированных ИЛ продолжительность жизни как самцов, так и самок оказалась сильно сниженной в присутствии вольбахии; в разных ИЛ вольбахия сокращала продолжительность жизни в 1.8–5.4 раза у самцов и в 1.4–2.4 раза у самок. Полученные результаты подтверждают, что, несмотря на широкое распространение, вольбахия может негативно влиять на продолжительность жизни.

Ключевые слова: *Drosophila melanogaster*; *Wolbachia* sp.; ПЦР; природные популяции; симбионт; продолжительность жизни.

*Wolbachia* is a genus of bacteria causing intracellular infection and represents the most widespread maternally transmitted facultative arthropod endosymbionts (Werren et al., 1995; Werren, Windsor, 2000; Hingelbroeker et al., 2008). From 40 % (Zug, Hammerstain, 2012) to 76 % (Jeyaprakash, Hoy, 2000) of arthropod species are estimated to have been infected. *Wolbachia* is widespread in the natural populations of *D. melanogaster* on all continents (Hoffmann et al., 1994, 1998; Solignac et al., 1994; Riegler et al., 2005; Ilinsky, Zakharov, 2007; Verspoor, Haddrill, 2011; Richardson et al., 2012; Ventura et al., 2012). However, the *Wolbachia* infection has never been characterized for the northern European populations of *D. melanogaster*.

*Wolbachia* is able to affect the life history traits of their hosts, mainly their reproduction, by causing cytoplasmic incompatibility, parthenogenesis induction, male-killing, and feminization, or by altering fecundity, fertility, and progeny survival (reviewed in Maistrenko et al., 2015). In *D. melanogaster*, apart from various life history traits (Olsen et al., 2001; Fry et al., 2004; Markov et al., 2009; Sharon et al., 2010; Serga et al., 2014), *Wolbachia* also affects sensitivity to stress (Brownlie et al., 2009; Versace et al., 2014) and viral infection (Teixeira et al., 2008; Martinez et al., 2015; Lindsey et al., 2018), sleep and aggressive behaviour (Rohrscheib et al., 2015; Bi et al., 2018). Among other traits, lifespan was reported to be influenced by *Wolbachia* (Min, Benzer, 1997; Brummel et al., 2004; Alexandrov et al., 2007; Carrington et al., 2009). It is believed that some effects of *Wolbachia* on the traits of their hosts evolved to promote spreading of bacteria (Werren et al., 2008). This increases the interest in exploring *Wolbachia* spreading, in particular, near the boundaries of the *D. melanogaster* habitats, in association with effects on vital traits of the host species.

In this paper, we describe the effect *Wolbachia* had in *D. melanogaster* inhabiting two localities in the northern regions of European Russia, which can be regarded as the outskirts of this species' natural habitat. The level of infection declined in years, but remained relatively high in the both populations. Our study has demonstrated that the life spans of both the males and females were negatively affected by the presence of *Wolbachia* in four of the five cases tested. This is a confirmation that *Wolbachia* has an important impact on the vital traits of their host organisms.

## Materials and methods

**Fly strains.** Flies were collected in private apple gardens located in two small hamlets without supermarkets or fruit markets, at a distance of several kilometers from Alexandrov, Russia (56.41° N, 38.72° E) and Valday, Russia (58.02° N, 33.24° E). The flies were collected manually from the surface of apple heaps (Alexandrov) or with baited traps (Valday). Fly collections were carried out from late August to early October of 2010, 2011, 2012, 2014 (Alexandrov) and in 2014, 2015 (Valday).

From 40 to 60 isofemale lines per population per year were started from females caught in nature. In their progeny, the male phenotype was checked to avoid contamination by *D. simulans*. Each line was maintained by brother × sister inbreeding during 20–22 generations. The inbred lines (IL) which survived inbreeding were further checked for the presence of *Wolbachia* and used for the lifespan measurements. Each IL was considered as genetically homogeneous (Falconer, Mackay, 1996) and represented one haploid genome of the population.

**Tests for *Wolbachia*.** DNA was extracted from 20–50 flies of the same genotype following the standard phenol-chloroform protocol. All ILs were checked for the presence of *Wolbachia* via quantitative PCR (MiniOpticon real-time PCR detection system, Bio-Rad) with primers to the *Wolbachia* 16S rRNA gene, 5′-CATACCTATTCGAAGGGATAG-3′ and 5′-AGCTTCGAGTGAAACCAATTC-3′ (Werren, Windsor, 2000). The lines showing positive results were treated with tetracycline (0.25 mg per 1 mL of fly food (Holden et al., 1993), with modifications) for three generations followed by three generations of recovery, before they were used in life span assays.

**Lifespan assays.** The lifespan was measured according to (Roshina et al., 2014). Five virgin flies of the same genotype and sex, all collected on the same day from cultures with moderate density, were placed in replicate vials. Flies were transferred weekly to vials with fresh food containing approximately 5 mL of standard medium without live yeast on the surface. Dead flies were recorded daily. Experiments comparing fly life spans were conducted simultaneously. Sample sizes were 50 flies/sex/line. Lifespan was estimated for each fly as the number of days alive from the day of eclosion to the day of death. Mean and median lifespan and survival curves were primarily used to characterize lifespan.

**Statistical analyses.** Fisher's exact test was used to compare the proportions of infected ILs. Standard descriptive statistic analysis of lifespan (Wilmoth, Horiuchi, 1999; Carey, 2003) was performed to determine mean lifespan and accompanying variances, standard deviations and standard errors; median, minimum and maximum lifespans; lifespans of the lower and upper quartiles, 10 and 90 percentiles. Survival curves were estimated using the Kaplan–Meier procedure. The nonparametric, distribution-free Mann–Whitney test and Kolmogorov–Smirnov test were used to evaluate the statistical significance of the difference between the survival curves.

## Results

The presence of *Wolbachia* was checked in inbred lines (ILs), each obtained from a single female collected in the private apple gardens located near Alexandrov, Russia (56.41° N, 38.72° E) and Valday, Russia (58.02° N, 33.24° E). From 20 to 75 lines per population per year were studied (Table 1). The level of infection was high in Alexandrov in 2010 and 2011, moderate in 2012 and became significantly lower in 2014 ($p < 0.01$, compared to 2010 or to 2011, Table 1). In Valday 2014, the level of infection was significantly higher than in Alexandrov 2014 ($p < 0.05$, see Table 1) and didn't change in 2015, although showed a tendency to reduction. When averaged over years, the level of infection was similar in the both populations: 69.7 % in Alexandrov and 68.4 % in Valday.

To assess the effects of *Wolbachia* on the lifespan, five infected ILs from Alexandrov 2010 were divided into two sub-lines each and, while one sub-line was maintained in the same way as earlier, the other was treated with tetracycline healing the *Wolbachia* infection. After treatment and subse-

**Table 1.** The level of *Wolbachia* infection in the natural populations of *D. melanogaster* inhabiting northers regions of European Russia

| Population | Year | Number of ILs | Number of infected ILs | % of infected ILs |
|---|---|---|---|---|
| Alexandrov | 2010 | 20 | 16 | 80.0 |
| | 2011 | 45 | 37 | 82.2 |
| | 2012 | 28 | 19 | 67.9 |
| | 2014 | 75 | 38 | 50.7 |
| Valday | 2014 | 60 | 42 | 70.0 |
| | 2015 | 48 | 32 | 66.7 |

**Table 2.** Parameters of male and female life spans in the Alexandrov ILs ($N = 50$)

| Line | Sex | *Wolbachia* | Mean | Median | Lower quartile | Upper quartile | *p*-values for comparisons with the infected sub-line | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Mann–Whitney test | Kolmogorov–Smirnov test |
| 1 | ♂ | + | 7 ± 0.5 | 7 | 4 | 27 | | |
| | | – | 38 ± 1.9 | 39 | 10 | 47 | **0.0000001** | **< 0.001** |
| | ♀ | + | 17 ± 0.9 | 19 | 11 | 23 | | |
| | | – | 41 ± 1.5 | 42 | 39 | 49 | **0.0000001** | **< 0.001** |
| 11 | ♂ | + | 25 ± 1.0 | 24 | 19 | 30 | | |
| | | – | 57 ± 1.6 | 57 | 52 | 64 | **0.0000001** | **< 0.001** |
| | ♀ | + | 20 ± 1.1 | 19 | 15 | 26 | | |
| | | – | 53 ± 2.1 | 51 | 45 | 68 | **0.0000001** | **< 0.001** |
| 19 | ♂ | + | 27 ± 1.3 | 31 | 18 | 35 | | |
| | | – | 50 ± 1.1 | 49 | 44 | 54 | **0.0000001** | **< 0.001** |
| | ♀ | + | 27 ± 1.2 | 27 | 21 | 33 | | |
| | | – | 46 ± 1.3 | 47 | 43 | 52 | **0.0000001** | **< 0.001** |
| 30 | ♂ | + | 31 ± 1.1 | 35 | 27 | 38 | | |
| | | – | 56 ± 2.3 | 55 | 45 | 68 | **0.0000001** | **< 0.001** |
| | ♀ | + | 31 ± 1.6 | 32 | 22 | 43 | | |
| | | – | 44 ± 1.4 | 44 | 31 | 56 | **0.0000001** | **< 0.001** |
| 33 | ♂ | + | 40 ± 1.7 | 39 | 34 | 47 | | |
| | | – | 44 ± 1.7 | 41 | 38 | 48 | 0.067205 | > 0.10 |
| | ♀ | + | 45 ± 1.7 | 48 | 40 | 54 | | |
| | | – | 49 ± 1.9 | 51 | 44 | 59 | 0.111280 | > 0.10 |

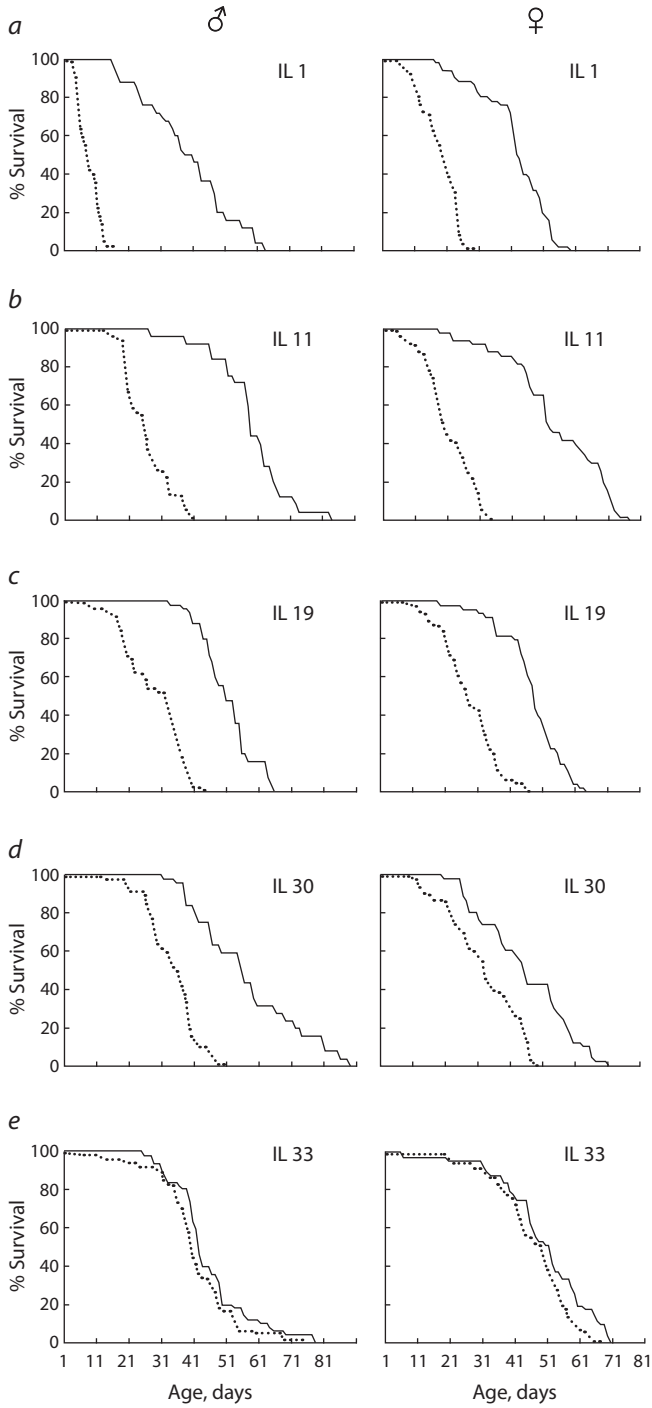
Note: Significant *p*-values are in bold case.

quent recovery, male and female life spans were measured in the both sub-lines of each of five ILs.

The mean life spans of infected males and females were relatively low (Table 2, Figure), the lowest being only seven days. In four ILs, the life spans of cured males and females were significantly longer than those of infected males and females (see Table 2, Figure): depending on the line, the mean life spans were increased from 1.8 to 5.4 times in males and from 1.4 to 2.4 times in females. The positive effect was thus smaller in females than in males. In one of the ILs, the life spans of infected males and females were not different from the life spans of cured males and females. In this line, the life spans of infected males and females were the highest, compared to the other four ILs.

## Discussion

Though the level of *Wolbachia* infection was analyzed in many natural populations of *D. melanogaster* all over the world, the number of populations studied is limited and they are distributed unevenly across the Earth's territory. The level of infection can be as high as 100 %, which was registered in Australian (Hoffmann et al., 1998) and Brazilian (Ventura et al., 2012) populations, and as low as 5 % (in Australian populations, Hoffmann et al., 1998) and 1 % (in Sub-Saharan African populations, Verspoor, Haddrill, 2011). In the large-scale work of Ilinsky and Zakharov (2007) populations from Ukraine, Belarus, Moldova, Caucasus, Central Asia, Ural, Udmurtia, Altai, West and East Siberia, and the Far East were studied in different years, from 1974 to 2005, and it was shown that the




Life spans of *Wolbachia*-infected males and females (dotted line) compared to the life spans of tetracycline-cured males (solid line).

*Wolbachia* infection rate varied from 70 % in Middle Asia to 40 % in Eastern Europe (Ukraine, Belarus, Moldova). Populations of the southern regions of European Russia (Caucasian) were analyzed by Bykov and co-authors (2014) who reported that the levels of *Wolbachia* infection were moderate in this region and higher than in the populations of Eastern Europe described in (Ilinsky, Zakharov, 2007). However, the published data are related only to the populations inhabiting the southern latitudes of the northern hemisphere and the northern latitudes of the southern hemisphere. Nothing has been previously known about the presence of *Wolbachia* in the northern populations of the northern hemisphere, including European Russia. Our study has demonstrated that *Wolbachia* presents in *D. melanogaster* inhabiting latitudes as high as 58° N. Infection frequencies appear to be moderate and vary over years.

While analyzing the frequency distribution of *Wolbachia* infection in various geographic regions, it should be taken into account that the studies were conducted in different years and seasons, and, overall, were nonsystematic. Monitoring of infection frequencies in Australian natural populations indicated that they could be both stable and fluctuate over time (Hoffmann et al., 1998). In Caucasian population, minor variations in the infection level were observed (Bykov et al., 2014). In our study, we observed a decline in infection frequencies over years. Laboratory cage studies showed that low larval density contributed to the stability of infection frequencies (Hoffmann et al., 1998). A plausible explanation would be that either the presence of *Wolbachia* is beneficial for the host or contamination occurs easier at high population density. However, there is no sufficient evidence to state that natural fluctuations in the level of infection are not random. Whether the observed fluctuations are explained by the fitness benefits, which *Wolbachia* provides for the host so that to persist in populations or, alternatively, by its unfavorable effects on fitness, remains to be elucidated.

Life span is a vital trait affected by *Wolbachia*. Interestingly, opposite effects were described by different authors: Min and Benzer (1997) and Alexandrov and co-authors (2007) demonstrated that, in infected flies, the life span decreased, while Brummel and co-authors (2004) reported the increased life span in infected flies. In the flies from Alexandrov population, healing from *Wolbachia* increased the life span and substantially slowed down aging in four out of five cases. Why both male and female life spans did not improve due to tetracycline treatment in one of the lines remains an enigma. One possible explanation is that the life span of infected flies in this line was the highest, compared to other ILs. Generally, it remains largely unclear why the *Wolbachia* infection in some cases leads to life extension, while in other cases shortens flies' life span. Previously, it was suggested that the effect on the life span depends on the strain of *Wolbachia* and ambient temperature (Brummel et al., 2004; Rohrscheib et al., 2016). At high temperatures of 25 to 29 °C, the deleterious effects of virulent *Wolbachia* strains on the host lifespan manifest themselves much more clearly, whereas at lower temperatures of 16 to 19 °C, the lifespan impact may be negligent or even slightly positive (Rohrscheib et al., 2016). During the late summer and early autumn, the outdoor temperature in Alexandrov varies from approximately 15 to 20 °C (https://ru.climate-data.org/location/929946/); accordingly, detrimental lifespan effects of *Wolbachia* may be compensated and pathogenic virulent strains could spread in the natural population of flies. At the laboratory conditions of 25 °C, the progeny of flies caught in late summer and early autumn would suffer evident lifespan shortening from the infection, as we have reported in this paper. In addition and in part as opposed to what has been said above, the divergence in longevity between different *D. melanogaster* lines was demonstrated to be mostly associated with the host genetic background rather than the type of *Wolbachia* infection (Carrington et al., 2009). Eventually, we

have to recognize that in spite of some available interesting data, it is not yet exactly clear how the life spans measured in the laboratory are related to the life spans under natural conditions. Further research is needed to describe the processes underlying the actual population dynamics of *Wolbachia* to the full extent.

What molecular mechanisms can determine the effect of *Wolbachia* endosymbiont on the *D. melanogaster* life span? In recent years, it has been demonstrated that *Wolbachia* can affect expression of the several genes related to aging. For example, it has been suggested that *Wolbachia* is able to upregulate insulin signaling downstream of InR (Insulin Receptor), thus negatively affecting the life span of flies (Ikeya et al., 2009). Specific interactions between insulin signaling and *Wolbachia* in lifespan regulation were confirmed in (Grönke et al., 2010). Effects of mutations in *Indy* also depended on the presence of *Wolbachia* (Toivenen et al., 2007). Further studies are needed to shed light on the molecular basis of *Wolbachia* effects on lifespan.

## Acknowledgements
This study was funded by the RFBR grant No. 15-54-46009 CT_a, the Presidium of the Russian Academy of Sciences Program No. 41 "Biodiveristy of natural systems and biological resources of Russia" (0112-2018-0022) and Russian State Project (0112-2018-0010).

## Conflict of interest
The authors declare they have no conflict of interest.

## References
Alexandrov I.D., Alexandrova M.V., Goryacheva I.I., Rochina N.V., Shaikevich E.V., Zakharov I.A. Removing endosymbiotic *Wolbachia* specifically decreases lifespan of females and competitiveness in a laboratory strain of *Drosophila melanogaster*. Russ. J. Genet. 2007;43(10):1147-1152. DOI 10.1134/S1022795407100080.

Bi J., Sehgal A., Williams J.A., Wang Y.F. Wolbachia affects sleep behavior in *Drosophila melanogaster*. J. Insect Physiol. 2018;27:81-88. DOI 10.1016/j.jinsphys.2018.02.011.

Brownlie J.C., Cass B.N., Riegler M., Witsenburg J.J., Iturbe-Ormaetxe I., McGraw E.A., O'Neill S.L. Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia pipientis*, during periods of nutritional stress. PLoS Pathog. 2009;5(4): e1000368. DOI 10.1371/journal.ppat.1000368.

Brummel T., Ching A., Seroude L., Simon A.F., Benzer S. *Drosophila* lifespan enhancement by exogenous bacteria. Proc. Natl. Acad. Sci. USA. 2004;101:12974-12979. DOI 10.1073/pnas.0405207101.

Bykov R.A., Ilinskii Yu.Yu., Voloshina M.A., Zakharov I.K. Prevalence and genotypic diversity of the symbiotic bacterium *Wolbachia* in the *Drosophila melanogaster* population of Nalchik. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2014;18(2):315-319. (in Russian)

Carey J.R. Longevity: The Biology and Demography of Life Span. Princeton, NT: Princeton Univ. Press, 2003.

Carrington L.B., Leslie J., Weeks A.R., Hoffmann A.A. The *popcorn* Wolbachia infection of *Drosophila melanogaster*: can selection alter Wolbachia longevity effects? Evolution. 2009;63(10):2648-2657. DOI 10.1111/j.1558-5646.2009.00745.x.

Falconer D.S., Mackay T.F.C. Introduction to Quantitative Genetics. 4th edn. Longman: Harlow, 1996.

Fry A.J., Palmer M.R., Rand D.M. Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. Heredity. 2004;93:379-389. DOI 10.1038/sj.hdy.6800514.

Grönke S., Clarke D.F., Broughton S., Andrews T.D., Partridge L. Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. PLoS Genet. 2010;6(2):e1000857. DOI 10.1371/journal.pgen.1000857.

Hoffmann A.A., Clancy D.J., Merton E. Cytoplasmic incompatibility in Australian populations of *Drosophila melanogaster*. Genetics. 1994; 136:993-999.

Hoffmann A.A., Hercus M., Dagher H. Population dynamics of the Wolbachia infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. Genetics. 1998;148:221-231.

Holden P.R., Jones P., Brookfield J.F. Evidence for a *Wolbachia* symbiont in *Drosophila melanogaster*. Genet. Res. 1993;62(1):23-29. DOI 10.1017/S0016672300031529.

Ikeya T., Broughton S., Alic N., Grandison R., Partridge L. The endosymbiont *Wolbachia* increases insulin/IGF-like signalling in *Drosophila*. Proc. R. Soc. B. 2009;276(1674):3799-3807. DOI 10.1098/rspb.2009.0778.

Ilinsky Y.Y., Zakharov I.K. The endosymbiont *Wolbachia* in Eurasian populations of *Drosophila melanogaster*. Russ. J. Genet. 2007; 43(7):748-756. DOI 10.1134/S102279540707006X.

Jeyaprakash A., Hoy M.A. Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76 % of sixty-three arthropod species. Insect. Mol. Biol. 2000;9(4):393-405. DOI 10.1046/j.1365-2583.2000.00203.x.

Lindsey A.R.I., Bhattacharya T., Newton I.L.G., Hardy R.W. Conflict in the intracellular lives of endosymbionts and viruses: A mechanistic look at *Wolbachia*-mediated pathogen-blocking. Viruses. 2018; 10(4):141. DOI 10.3390/v10040141.

Maistrenko O.M., Serga S.V., Vaiserman A.M., Kozeretska I.A. Effect of *Wolbachia* Infection on Aging and Longevity-Associated Genes in *Drosophila*. In: Vaiserman A.M., Moskalev A.A., Pasyukova E.G. (Eds.) Life Extension: Lessons from Drosophila. Vol. 3. Healthy Ageing and Longevity. Springer Int. Publ., 2015.

Markov A.V., Lazebny O.E., Goryacheva I.I., Antipin M.I., Kulikov A.M. Symbiotic bacteria affect mating choice in *Drosophila melanogaster*. Anim. Behav. 2009;77(5):1011-1017. DOI 10.1016/j.anbehav.2009.01.011.

Martinez J., Ok S., Smith S., Snoeck K., Day J.P., Jiggins F.M. Should symbionts be nice or selfish? Antiviral effects of Wolbachia are costly but reproductive parasitism is not. PLoS Pathog. 2015;11(7): e1005021. DOI 10.1371/journal.ppat.1005021.

Min K.T., Benzer S. *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent, causing degeneration and early death. Proc. Natl. Acad. Sci. USA. 1997;94(20):10792-10796. DOI 10.1073/pnas.94.20.10792.

Olsen K., Reynolds K.T., Hoffmann A.A. A field cage test of the effects of the endosymbiont *Wolbachia* on *Drosophila melanogaster*. Heredity. 2001;86(6):731-737. DOI 10.1046/j.1365-2540.2001.t01-1-00892.

Richardson M.F., Weinert L.A., Welch J.J., Linheiro R.S., Magwire M.M., Jiggins F.M., Bergman C.M. Population genomics of the Wolbachia endosymbiont in *Drosophila melanogaster*. PLoS Genet. 2012;8(12):e1003129. DOI 10.1371/journal.pgen.1003129.

Riegler M., Sidhu M., Miller W.J., O'Neill S.L. Evidence for a global *Wolbachia* replacement in *Drosophila melanogaster*. Curr. Biol. 2005;15(15):1428-1433. DOI 10.1016/j.cub.2005.06.069.

Rohrscheib C.E., Bondy E., Josh P., Riegler M., Eyles D., van Swinderen B., Weible II M.W., Brownlie J.C. *Wolbachia* influences the production of octopamine and affects *Drosophila* male aggression. Appl. Environ. Microbiol. 2015;81(14):4573-4580. DOI 10.1128/AEM.00573-15.

Rohrscheib C.E., Frentiu F.D., Horn E., Ritchie F.K., van Swinderen B., Weible M.W., Brownlie J.C. Intensity of mutualism breakdown is determined by temperature not amplification of *Wolbachia* genes. PLOS Pathog. 2016;12(9):e1005888. DOI 10.1371/journal.ppat.1005888.

Roshina N.V., Symonenko A.V., Krementsova A.V., Trostnikov M.V., Pasyukova E.G. Embryonic expression of shuttle craft, a Drosophila gene involved in neuron development, is associated with adult life-

span. Aging (Albany NY). 2014;6(12):1076-1093. DOI 10.18632/aging.100712.

Serga S., Maistrenko O., Rozhok A., Mousseau T., Kozeretska I. Fecundity as one of possible factors contributing to the dominance of the *w*Mel genotype of *Wolbachia* in natural populations of *Drosophila melanogaster*. Symbiosis. 2014;63(1):11-17. DOI 10.1007/s13199-014-0283-1.

Sharon G., Segal D., Ringo J.M., Hefetz A., Zilber-Rosenberg I., Rosenberg E. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA. 2010; 107(46):20051-20056. DOI 10.1073/pnas.1009906107.

Solignac M., Vautrin D., Rousset F. Widespread occurrence of the proteobacteria *Wolbachia* and partial cytoplasmic incompatibility in *Drosophila melanogaster*. C. R. Acad. Sci. III. 1994;317:461-470.

Teixeira L., Ferreira Á., Ashburner M. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. PLoS Biol. 2008;6(12):e1000002. DOI 10.1371/journal.pbio.1000002.

Toivenen J.M., Walker G.A., Martinez-Diaz P., Bjedov I., Driege Y., Jacobs H.T., Gems D., Partridge L. No influence of *Indy* on lifespan in *Drosophila* after correction for genetic and cytoplasmic background effects. PLoS Genet. 2007;3(16):e95. DOI 10.1371/journal.pgen.0030095.

Ventura I.M., Martins A.B., Lyra M.L., Andrade C.A., Carvalho K.A., Klaczko L.B. Spiroplasma in *Drosophila melanogaster* populations: prevalence, male-killing, molecular identification, and no association with *Wolbachia*. Microb. Ecol. 2012;64(3):794-801. DOI 10.1007/s00248-012-0054-6.

Versace E., Nolte V., Pandey R.V., Tobler R., Schlötterer C. Experimental evolution reveals habitat-specific fitness dynamics among *Wolbachia* clades in *Drosophila melanogaster*. Mol. Ecol. 2014;23(4):802-814. DOI 10.1111/mec.12643.

Verspoor R.L., Haddrill P.R. Genetic diversity, population structure and Wolbachia infection status in a worldwide sample of Drosophila melanogaster and D. simulans populations. PLoS One. 2011;6:e26318. DOI 10.1371/journal.pone.0026318.

Werren J.H., Baldo L., Clark M.E. *Wolbachia*: master manipulators of invertebrate biology. Nat. Rev. Microbiol. 2008;6:741-751. DOI 10.1038/nrmicro1969.

Werren J.H., Guo L., Windsor D.W. Distribution of *Wolbachia* among neotropical arthropods. Proc. R. Soc. B. 1995;262:197-204. DOI 10.1098/rspb.1995.0196.

Werren J.H., Windsor D.M. *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? Proc. R. Soc. B. 2000;267:1277-1285. DOI 10.1098/rspb.2000.1139.

Wilmoth J.R., Horiuchi S. Rectangularization revisited: variability of age at death within human populations. Demography. 1999;36:475-495. DOI 10.2307/2648085.

Zug R., Hammerstein P. Still a host of hosts for *Wolbachia*: Analysis of recent data suggests that 40 % of terrestrial arthropod species are infected. PLoS One. 2012;7(6):e38544. DOI 10.1371/journal.pone.0038544.

Вавиловский журнал генетики и селекции. 2018;22(5):574-585
DOI 10.18699/VJ18.397

# Invasive mosquito species *Aedes albopictus* and *Aedes aegypti* on the Black Sea coast of the Caucasus: genetics (*COI*, ITS2), *Wolbachia* and *Dirofilaria* infections

E.V. Shaikevich[1] ✉, I.V. Patraman[2], A.S. Bogacheva[3], V.M. Rakova[2], O.P. Zelya[2], L.A. Ganushkina[2]

[1] Vavilov Institute of General Genetics, RAS, Moscow, Russia
[2] Institute of Medical Parasitology and Tropical Medicine named after E. Martsinovskii, Sechenov First Moscow State Medical University, Moscow, Russia
[3] Moscow State University, Moscow, Russia

The area of invasive species *Aedes albopictus* and *Aedes aegypti* is expanding. Precise identification and understanding of the genetic diversity of invasive mosquito populations allows us to develop appropriate control methods. Endosymbiotic bacterium *Wolbachia pipientis* has different effects on their arthropod hosts and can influence the transmission and spread of the pathogens. The objective of the presented study was molecular-genetic identification of the *Aedes* mosquitoes collected in sampling sites on the Black Sea coast from 2007 to 2017; determination of genetic variability of *Ae. aegypti*, *Ae. albopictus* and their symbiotic bacteria *Wolbachia*; assessment of mosquitoes ability to be infected and to spread parasitic *Dirofilaria*. Another objective was obtaining the genetic characteristic of laboratory strain *Ae. aegypti* IMPITM. We investigated two markers of nuclear and mitochondrial DNA from *Ae. albopictus* and *Ae. aegypti* and compared them to DNA from *Ae. cretinus* and *Ae. koreicus* sympatrically inhabiting the territory, as well as to one of *Ae. aegypti* from a laboratory line. The study of nuclear and mitochondrial DNA revealed a low level of variability in the invasive mosquitoes *Ae. albopictus* and *Ae. aegypti* collected at different collection sites and in different years. More than a half of *Ae. albopictus* were infected with *Wolbachia*, two strains of bacteria, *w*AlbA and *w*AlbB, occur in the *Ae. albopictus* population on the Black Sea coast. Total infection of *Ae. aegypti* and *Ae. albopictus* with dirofilariae was 1.8 %. *Dirofilaria immitis* was found only in mosquito abdomen, larvae of infective stage L3 were not found. *D. repens* larvae developed to the infective stage in the mosquitoes of both species.

Key words: blood-sucking mosquitoes; *Aedes aegypti*; *Aedes albopictus*; invasion; population; Black Sea coast; ITS2; *COI*; *Wolbachia*; *Dirofilaria*.

# Инвазивные виды *Aedes albopictus* и *Aedes aegypti* на Черноморском побережье Краснодарского края: генетика (*COI*, ITS2), зараженность *Wolbachia* и *Dirofilaria*

Е.В. Шайкевич[1] ✉, И.В. Патраман[2], А.С. Богачева[3], В.М. Ракова[2], О.П. Зеля[2], Л.А. Ганушкина[2]

[1] Институт общей генетики им. Н.И. Вавилова РАН, Москва, Россия
[2] Институт медицинской паразитологии, тропических и трансмиссивных заболеваний им. Е.И. Марциновского Первого Московского государственного медицинского университета им. И.М. Сеченова (Сеченовский университет) Министерства здравоохранения РФ, Москва, Россия
[3] Московский государственный университет им. М.В. Ломоносова, Москва, Россия

Ареал инвазивных видов *Aedes aegypti* и *Aedes albopictus*, переносчиков ряда трансмиссивных инфекций, расширяется. Идентификация видов-переносчиков и понимание генетического разнообразия инвазивных популяций позволяют разработать соответствующие профилактические мероприятия. Эндосимбиотическая бактерия *Wolbachia pipientis* оказывает различные эффекты на своих хозяев-артропод и может влиять на процесс передачи и распространения возбудителей. Основной целью работы была молекулярно-генетическая идентификация видов комаров рода *Aedes*, собранных в населенных пунктах Черноморского побережья с 2007 по 2017 г.; определение генетической изменчивости *Ae. aegypti*, *Ae. albopictus* и их симбиотической бактерии *Wolbachia*; оценка способности *Ae. aegypti* и *Ae. albopictus* к заражению и распространению паразитических *Dirofilaria*. Отдельной задачей являлась генетическая характеристика лабораторной линии *Ae. aegypti* ИМПиТМ, которая поддерживается в лаборатории в течение 50 лет. Исследованы маркеры ядерной и митохондриальной ДНК у *Ae. albopictus* и *Ae. aegypti* и проведено их сравнение с *Ae. cretinus* и *Ae. koreicus*, симпатрически обитающими на данной территории, а также с *Ae. aegypti* лабораторной линии. Обнаружен низкий уровень изменчивости *Ae. albopictus* и *Ae. aegypti*, собранных в природе в разных точках сбора и в разное время. У *Ae. albopictus* выявлены четыре гаплотипа на основе сравнения вариабельной области внутреннего транскрибируемого спейсера (ITS2) кластера генов рРНК и два митохондриальных гаплотипа при сравнении последователь-

Received 30.03.2018
Accepted for publication 20.05.2018


✉ e-mail: elenashaikevich@mail.ru

ностей гена первой субъединицы цитохромоксидазы (*COI*). У *Ae. aegypti*, собранных в природе, обнаружены четыре гаплотипа ядерной ДНК и три митохондриальных гаплотипа. Более половины *Ae. albopictus* заражены *Wolbachia*. В популяции на Черноморском побережье Краснодарского края встречаются два штамма бактерии: *w*AlbA и *w*AlbB. Общая зараженность комаров *Ae. aegypti* и *Ae. albopictus* дирофиляриями составила 1.8 %. *Dirofilaria immitis* обнаружены только в брюшках комаров, развития личинок до инфекционной стадии L3 не выявлено. Личинки *D. repens* развились до инфекционной стадии в комарах обоих видов.

Ключевые слова: кровососущие комары; *Aedes aegypti*; *Aedes albopictus;* инвазия; популяция; Черноморское побережье Кавказа; ITS2; *COI*; *Wolbachia*; *Dirofilaria*.

**HOW TO CITE THIS ARTICLE:**
Shaikevich E.V., Patraman I.V., Bogacheva A.S., Rakova V.M ., Zelya O.P., Ganushkina L.A. Invasive mosquito species *Aedes albopictus* and *Aedes aegypti* on the Black Sea coast of the Caucasus: genetics (*COI*, ITS2), infection with *Wolbachia* and *Dirofilaria*. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):574-585. DOI 10.18699/VJ18.397

At the beggining of the 21st century, expansion of two mosquito species – *Aedes* (Stegomiya) *aegypti* (Linnaeus, 1796) and *Aedes* (Stegomiya) *albopictus* (Skuse, 1895) – was registered in Krasnodarskiy Region. These are associated with numerous epidemic outbreaks of denge, chikungunya, zika viral infections, etc. *Ae. aegypti* also acts as a primary carrier for yellow fever (Jeffries, Walker, 2016). Beside arboviruses, both mosquito species are able to transmit the threadworm larvae of *Dirofilaria* family responsible for dirofilariasis in humans and animals (Ganushkina et al., 2014a). This transmissible helmintosis demonstrates an expansive trend in the territory of Russia (Bogacheva et al., 2017).

*Ae. aegypti* had been absent in Russia for 50 years, until 2001 (Ryabova et al., 2005; Yunicheva et al., 2008) due to the special measures taken in the USSR in the 1920s and 1930s aimed at eradication of the aforementioned species (Martsinovsky, 1929). The measures were taken considering the extreme danger of this carrier and occurrence of substantial high-mortality dengue outbreaks in the countries of South Europe.

Later in the 20th century, greater concerns aroused worldwide about the expansion of *Ae. albopictus*, another effective carrier of dangerous arboviruses, which started to spread beyond its natural habitat in Southeast Asia. This mosquito species expanded into areas, previously inhabited by other endemic species, like morphologically similar *Ae. cretinus* Edwards (Patsoula et al., 2006). For the first time, this species was transferred to Albania from China in the middle of the 1970s (Adhami, Reiter, 1998). Currently this species of mosquito is found in more than 15 countries, and its habitat is growing persistently (Medlock et al., 2015). *Ae. albopictus* are able to cause outbreaks of dengue and chicagunya infections even in absence of *Ae. aegypti* (Delatte et al., 2008; Delisle et al., 2015; Calba et al., 2017; Chuchuy et al., 2018). In the Russian Federation these species were first found in 2011 (Ganushkina et al., 2012). Currently, the Caucasus coast of the Black Sea is the only territory in the WHO European region that geographically includes the Asian region of the Caucasus where the both mosquito species have been registered as active vectors for arbovirus infectious agents (http://www.who.int/about/regions/euro/).

Hematophagous mosquito species may vary in pathogen transmission capability. Their morphological markers used for authentication could be either unclear, or missed or erased during storage of adult specimens, which results in identification errors. For example, cytochrome gene sequence MF148262 annotated in Gene Bank as being derived from *Ae. albopictus* originated in Malaysia, was actually associated with *Ae. aegypti*. For that reason, DNA analysis for verification of the population's structure and diversity had to be performed to devise necessary preventive measures.

Dirofilariasis caused by *Dirofilaria immitis* and *D. repens* is endemic to the Southern parts of Russia (Sergiev et al., 2014; Kartashev et al., 2018). In assessment of epidemiological situation one uses xenomonitoring for entomological control over filariasis infections (Ganushkina et al., 2014a). Dirofilaries are identified by means of total DNA amplification of vector mosquito with specific primers. This method shows fillaries at all stages of development (L1, L2, L3), while not every filaria reaches the pathogenic L3 stage, which can be transmitted through salivary glands of a mosquito to human and animal hosts. Microfilaria development to stage L3 requires certain thermal conditions, therefore the disease is primarily limited to southern regions. *Dirofilaria* can be found in mosquitoes of the following genera: *Aedes*, *Anopheles*, *Ochlerotatus* and *Culex* (Bochková et al., 2013; Kronefeld et al., 2014; Bogacheva et al., 2017). During monitoring, taxonomical study of mosquitoes can be frequently limited to the genus level. *Dirofilaria* identification, specialization towards mosquito species and determining the invasive stage of *Dirofilaria* inside the mosquito are aimed towards discovering the actual epidemiological significance of various mosquito species (Ganushkina et al., 2014a). Major attention in latest research has been paid to discovering the role of endosymbiotic bacteria in the mosquito and other arthropods in the host survivability, as well as in processes of transmission

and spread of infectious diseases (Bourtzis et al., 2014; Jeffries, Walker, 2016).

The main objectives of the presented research were molecular genetics identification of *Ae. albopictus* and *Ae. aegypti* collected in the populated areas of the Black Sea coast from 2007 to 2017; determination of their gene diversity; characterization of sympatric species *Ae. cretinus* and *Ae. koreicus*; measuring of symbiotic bacteria *Wolbachia* appearence in investigated mosquito samples and of *Wolbachia* strains diversity within the population of *Ae. albopictus*; evaluation of the capability of *Ae. albopictus* and *Ae. aegypti* to vector and spread parasitic *Dirofilaria*. Within the framework of the study one also performed genetic sampling of the laboratory line of *Ae. aegypti* IMPITM cultivated for 50 years.

## Materials and methods

**Collecting the mosquitoes.** The mosquitoes were collected on the North Caucasus shore in years 2007, 2011–2013 and 2015–2017 (Table 1, Fig. 1). *Ae. albopictus* and *Ae. aegypti* larvae were picked in temporal water basins using a scoop-net. A part of the larvae was preserved in alcohol, while the others were raised to adults. The adult mosquitoes were caught either with an Electrofrog trap (LMD-Komplekt plus, Russia), or "on self" with an exhauster and preserved dry. More details on the mosquito collection protocol can be found in (Ganushkina et al., 2013, 2016). The total collection consisted of 3 005 specimens: 1 430 *Ae. aegypti* and 1 575 *Ae. albopictus*.

Molecular genetic analysis included mosquitoes from the IMPITM lab line, four dried adult *Ae. cretinus* from the IMPITM museum and five *Ae. koreicus* collected in Sochi in 2013. The *Ae. cretinus* and *Ae. koreicus* were used to determine the genetic differences between morphologically similar species of the *Aedes* genus, which could potentially present in the collections, collected in the territory in question. *Ae. cretinus* is endemic towards the Black Sea coast of the Caucasus. *Ae. koreicus* originating from Southeast Asia, have been found in the area since 2013 (Ganushkina et al., 2016).

**Determining species of mosquitoes.** Species identification of all the mosquito samples was conducted with account for the morphology data (Gutsevich et al., 1970) and using molecular genetic technologies. Up to 30 specimens from each place and year were used for PCR identification of second inner transcribable spacer of pRNA gene cluster (ITS2). Characteristic to *Ae. albopictus* is PCR sequence of 500 bp in size, for *Ae. koreicus* – 450 bp, for *Ae. cretinus* – 390 bp, and for *Ae. aegypti* – 340 bp.

**Mosquito identification by gene sampling.** DNA extraction from the mosquitoes was accomplished using DIAtom DNA Prep (Izogen, Moscow). PCR identification was conducted with the Evrogen Encyclo PCR kit (Evrogen, Moscow). For the ITS2 amplification primers 5,8S and 28S were used (Porter, Collins, 1991). A cytochrom oxidase I sequence (*COI*) of nearly 750 bp in length was built up using primers TY-J-1460 (Simon et al., 1994) and COIR (Shaikevich, 2007). Amplificates

**Table 1.** Years and points of *Ae. aegypti* and *Ae. albopictus* sampling

| Year | Sampling site | Geographical coordinates | Number of individuals | |
|---|---|---|---|---|
| | | | *Ae. aegypti* | *Ae. albopictus* |
| 2007 | Adler | 43°25′44″ N, 39°55′26″ E | 52 | 0 |
| | Sochi | 43°35′07″ N, 39°43′13″ E | 89 | 0 |
| | Lazarevskoye | 43°54′31″ N, 39°19′52″ E | 25 | 0 |
| | Tuapse | 44°06′19″ N, 39°04′48″ E | 23 | 0 |
| 2011 | Hosta | 43°30′53″ N, 39°52′05″ E | 1 | 16 |
| 2012 | Adler | 43°25′44″ N, 39°55′26″ E | 3 | 24 |
| | Hosta | 43°30′53″ N, 39°52′05″ E | 0 | 47 |
| | Sochi | 43°35′07″ N, 39°43′13″ E | 6 | 116 |
| | Mamaika | 43°38′35″ N, 39°42′34″ E | 48 | 406 |
| | Dagomys | 43°40′11″ N, 39°40′07″ E | 0 | 24 |
| | Lazarevskoye | 43°54′31″ N, 39°19′52″ E | 31 | 48 |
| | Tuapse | 44°06′19″ N, 39°04′48″ E | 566 | 28 |
| | New Afon | 43°04′50″ N, 40°50′17″ E | 45 | 6 |
| | Pizunda | 43°09′43″ N, 40°20′27″ E | 7 | 58 |
| 2013 | Adler | 43°25′44″ N, 39°55′26″ E | 6 | 164 |
| | Hosta | 43°30′53″ N, 39°52′05″ E | 0 | 23 |
| | Sochi | 43°35′07″ N, 39°43′13″ E | 3 | 46 |
| | Mamaika | 43°38′35″ N, 39°42′34″ E | 3 | 34 |
| | Lazarevskoye | 43°54′31″ N, 39°19′52″ E | 17 | 19 |
| | Tuapse | 44°06′19″ N, 39°04′48″ E | 394 | 21 |
| 2015 | Adler | 43°25′44″ N, 39°55′26″ E | 0 | 120 |
| | Tuapse | 44°06′19″ N, 39°04′48″ E | 30 | 20 |
| 2016 | Dagomys | 43°40′11″ N, 39°40′07″ E | 0 | 256 |
| 2017 | Adler | 43°25′44″ N, 39°55′26″ E | 81 | 32 |
| | Sochi | 43°35′07″ N, 39°43′13″ E | 0 | 67 |
| | Total | | 1430 | 1575 |

were visualized in 1–2 % agarous gel and purified using a clean-up extraction kit (Evrogen, Moscow) followed by sequencing with the BigDye Termination kit 3.1 (Applied Biosystems, USA). Thirteen ITS2 amplificates of *Ae. aegypti* и *Ae. albopictus* were sequenced (1–2 samples from 10 collection sites for various years), including *Ae. cretinus* and *Ae. aegypti* taken in the amount of one sample from each line. Mitochondrial DNA variability was studied using 28 sequenced sequences of the *COI* gene, 634 bp in length, from 1–4 specimens collected at 1–4 sites and four specimens from the *Ae. aegypti* lab line. The sequences were registered in Gene Bank. The *COI* gene: *Ae. aegypti* MG198586–MG198594, MH251909–MH251911; *Ae. albopictus* MG198595–MG198606; *Ae. aegypti* IMPITM MH023409 и ITS2: *Ae. aegypti* MH142316–MH142320; *Ae. albopictus* MH142321–MH142326; *Ae. aegypti* IMPITM MH142327; *Ae. cretinus* MH142328.

**Data analysis.** Gene sequence analysis was performed using the following software: ChromasPro, BLASTN, ClustalW, MAFFT v.6, MEGA v.6. Phylogenetic trees were built using the Neighbor-Joining technique, the evolution distances were calculated by maximum composite likelihood method using MEGA v.6 program (Tamura et al., 2013). The DNA sequences of *Ae. albopictus* and *Ae. aegypti* closest to the extracted ones, as well as ones characteristic for certain regions were selected in Gene Bank (https://www.ncbi.nlm.nih.gov/) for the purpose of comparative analysis. Their registry numbers can be found on the diagrams. Sample collection time periods are indicated in the annotations, where it is possible. The statistical reliability of filogenetic tree branches was analyzed using the bootstrap method (1 000 iterations). Evolutionary divergence level between the sequences was evaluated using the MEGA v.6 software (Tamura et al., 2013).

***Wolbachia* contamination.** Identification of the *Wolbachia* symbiotic bacteria was carried out using specific primers for bacterial surface protein gene *wsp* (81F and 691R, see Braig et al., 1998). In order to separate the two strains, multiprimer PCR was used (Zhou et al., 1998): primers 383F and 183F were paired with wsp-691R to separate the *w*AlbA and *w*AlbB strains of *Wolbachia* in *Ae. albopictus.* The PCR fragment corresponding to the *w*AlbA strain was one of 379 bp, and to the *w*AlbB strain – of 501 bp. Validity evaluation for the received data on bacterial contamination in the samples was carried out using the Fischer accuracy test with the error margin set for $N > 10$ (Tokarev et al., 2017).

**Susceptibility to *Dirofilaria*.** Only hemotrophic gonoactive female mosquitoes collected in the wild were used for analysis. In order to determine dirofilaria contamination in mosquito pools, the mosquito imagos were dissected, their abdomen and head-thorax parts separated, and 2 to 7 sample mosquitoes were grouped into pools based on their collection time and place. In the head-thorax parts, L3 larvae were registered. Contamination of mosquitoes with *Dirofilaria* larvae was revealed by the amplification of the ITS2 area using the DIDR-F1 and DIDR-R1 primers (Rishniw et al., 2006). PCR sequence size, specific


**Fig. 1.** Sampling points on the Caucasian coast of the Black Sea.

to *D. immitis* was 542 bp, for *D. repens* – 484 bp. For *Dirofilaria* DNA screening was conducted in pools and not individually, contamination was evaluated using the common MIR (minimum infection rate) method. The value was calculated based on the assumption that at least one mosquito specimen in the pool was infected with *Dirofilaria* with the minimum infection rate calculated as the number of positive pools divided by the total specimen quantity and expressed as a percentage (Cancrini et al., 2003).

## Results

### Geography of sampling

Mosquitoes *Ae. aegypti* and *Ae. albopictus* were collected on the coast of the Black Sea of Krasnodarskiy Region from Adler to Tuapse in years 2007, 2011–2013 and 2015–2017 (see Table 1). In 2012, the specimens of these species were additionally gathered in the Republic of Abkhazia near Pizunda and New Afon (see Fig. 1). In 2007 *Ae. albopictus* were absent in the region, while in four sampling sites from Adler to Tuapse *Ae. aegypti* were common (see Table 1). Starting 2011, *Ae. albopictus* were present in each pool from every populated area. Moreover, the quantity of collected *Ae. albopictus* was vastly superior to the one of *Ae. aegypti* in the area from Adler to Dagomys. In years 2012, 2013 and 2016 *Ae. aegypti* were not registered in Hosta and Dagomys. However, *Ae. aegypti* prevailed numerically in Tuapse, the northern part of the region, in years 2012 and 2013.

In year 2017 collection was conducted only in Sochi and Adler, and the larvae and imagos were attributed solely to *Ae. albopictus*. In the pools of adult mosquitoes no *Ae. aegypti* specimen was present, as no *Ae. aegypti* larvae were found in typical breeding grounds of this species (various small artificial basins filed with water: barrels, cans, decorative pools, old dishware, tires). A sample from a dried car tire found in Adler in August 2017 contained preserved viable eggs of *Ae. aegypti* and *Ae. albopictus* with the prevalence of *Ae. aegypti,* which later developed in the lab into larvae and imagos.

**Table 2.** Variable sites in the ITS2 region of *Ae. aegypti*

| GenBank annotation | Origin | Variable nucleotide sites* | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 5 3 | 2 5 4 | 2 5 5 | 3 4 9 | 3 5 1 | 3 5 2 | 3 5 3 | 3 5 4 | 3 5 5 | 3 5 6 | 3 5 7 | 3 5 8 | 4 9 2 | 4 9 4 |
| MH142327 | IMPITM | G | T | G | A | C | T | A | A | C | T | A | G | T | T |
| MH142320 | Sochi, 2007 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| MH142318 | Mamaika, 2012 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| MH142317 | Tuapse, 2012 | . | . | A | . | . | . | . | . | . | . | . | . | . | . |
| MH142326 | Adler, 2017 | – | – | . | C | . | . | . | . | . | . | . | A | G | . |
| MH142316 | New Afon, 2012 | – | – | . | . | – | – | – | – | – | – | – | – | . | C |
| MH142319 | Adler, 2013 | – | – | . | . | – | – | – | – | – | – | – | – | . | C |

Note: * The nucleotide positions are indicated relative to the MH142327 sequence. In Tables 2 and 3, the points denote the nucleotides identical to those indicated in the first line, dashes – the absence of nucleotides in the sequence (deletions).

**Table 3.** Variable sites in the ITS2 region of *Ae. albopictus*

| Gene Bank annotation | Origin | Variable nucleotide sites* | | | |
|---|---|---|---|---|---|
| | | 3 0 8 | 3 3 1 | 3 6 3 | 3 6 4 |
| MH142321 | Hosta, 2011 | G | T | G | C |
| MH142322 | Pizunda, 2012 | . | . | . | . |
| MH142323 | Pizunda, 2012 | A | C | – | – |
| MH142324 | Tuapse, 2013 | A | C | . | . |
| MH142325 | Adler, 2017 | . | C | . | . |

* The nucleotide positions are indicated relative to the MH142321 sequence.

**Table 4.** Estimates of average evolutionary divergence over ITS2/*COI* sequence pairs between and within the studied *Aedes*

| Species | *Ae. aegypti* | *Ae. albopictus* | *Ae. cretinus* |
|---|---|---|---|
| *Ae. aegypti* | 0.005* 0.0028* | | |
| *Ae. albopictus* | 0.14 0.37 | 0.004* 0.0127* | |
| *Ae. cretinus* | n. d. 0.30 | n. d. 0.38 | |
| *Ae. koreicus* | n. d. 0.28 | n. d. 0.46 | n. d. 0.39 |

Note: The number of base substitutions per site from averaging over all sequence pairs within and between the *Aedes* species is shown. Below the diagonal are differences in ITS2 sequences; above the diagonal – in *COI*.

* Intraspecific differences; n. d. – not defined.

## PCR analysis

***Ae. aegypti* lab line.** The *Ae. aegypti* mosquito lab line has been constantly sustained for more than five decades in the laboratory of the Martsinovskiy institute. The line was named "IMPITM" after the abbreviated name of the institute. Parts of the *COI* gene were sequenced for two mosquitoes from two different genres. DNA sequences from all four specimens were identical, one sequence was registered in Gene Bank under number MH023409. Two variable sites G148A and T624a (the site numbers correspond to MH023409) in *COI* separated IMPITM-line from lab cultures *Ae. aegypti* Liverpool (AY432648) and RED (AF390098). As a nuclear marker, ITS2 area from *Ae. aegypti* IMPITM (GenBank annotation MH142327) was sequenced. ITS2 DNA in *Ae. aegypti* IMPITM was identical to DNA of the Rockfeller strain (KF471588).

**Nuclear DNA variability in *Ae. aegypti* and *Ae. albopictus*.** In *Ae. aegypti* four haplotypes with different single-nucleotide substitutions and two deletions in their ITS2 areas were found (Table 2). The first type was *Ae. aegypti* collected in Sochi (2007) and Big Sochi – Mamaika (2012). The second type was from Tuapse (2012) and differed from the first by one G255A replacement. The third haplotype was discovered in the specimens from New Afon and Adler (2013). This haplotype differed by two deletions (two and eight nucleotides) and one T494C replacement. In the ITS2 area of *Ae. aegypti* collected in Adler (2017) two deleted nucleotides were found, which was similar to the first haplotype, and three nucleotide replacements – A349C, G358A and T492G.

Four variable haplotypes were found in *Ae. albopictus*, different by single-nucleotide mutations and one deletion (Table 3). The earliest in this respect is the Hosta collection (2011). We found an identical haplotype in *Ae. albopictus* from Pizunda (2012). The second similar haplotype was found in another specimen from the same collection (Pizunda, 2012), different by deletions of two nucleotides and replacements in G308A and T331C. The third haplotype of *Ae. albopictus* (Tuapse, 2013), unlike the second, had no deletion. The fourth one was found in *Ae. albopictus* (Adler, 2017) and had one T331C replacement.

Intraspecific ITS2 variability in wild *Ae. aegypti* was 0.3 %, while the variability in wild *Ae. albopictus* was 1.3 % (Table 4). The genetic differences between species

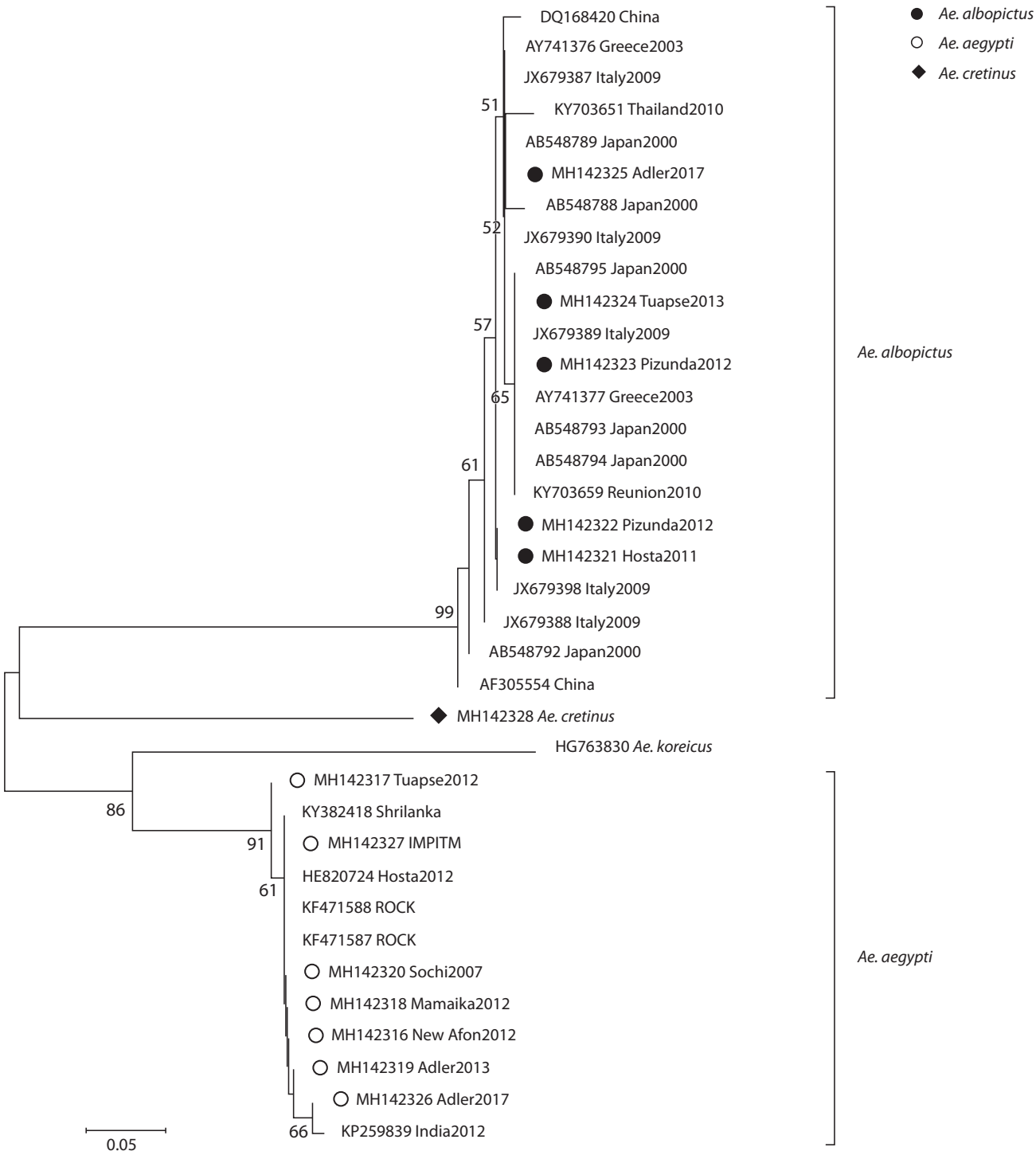



**Fig. 2.** Similarity dendrogram derived from comparative analysis of the ITS2 areas.
All the deletions were excluded from the analysis.

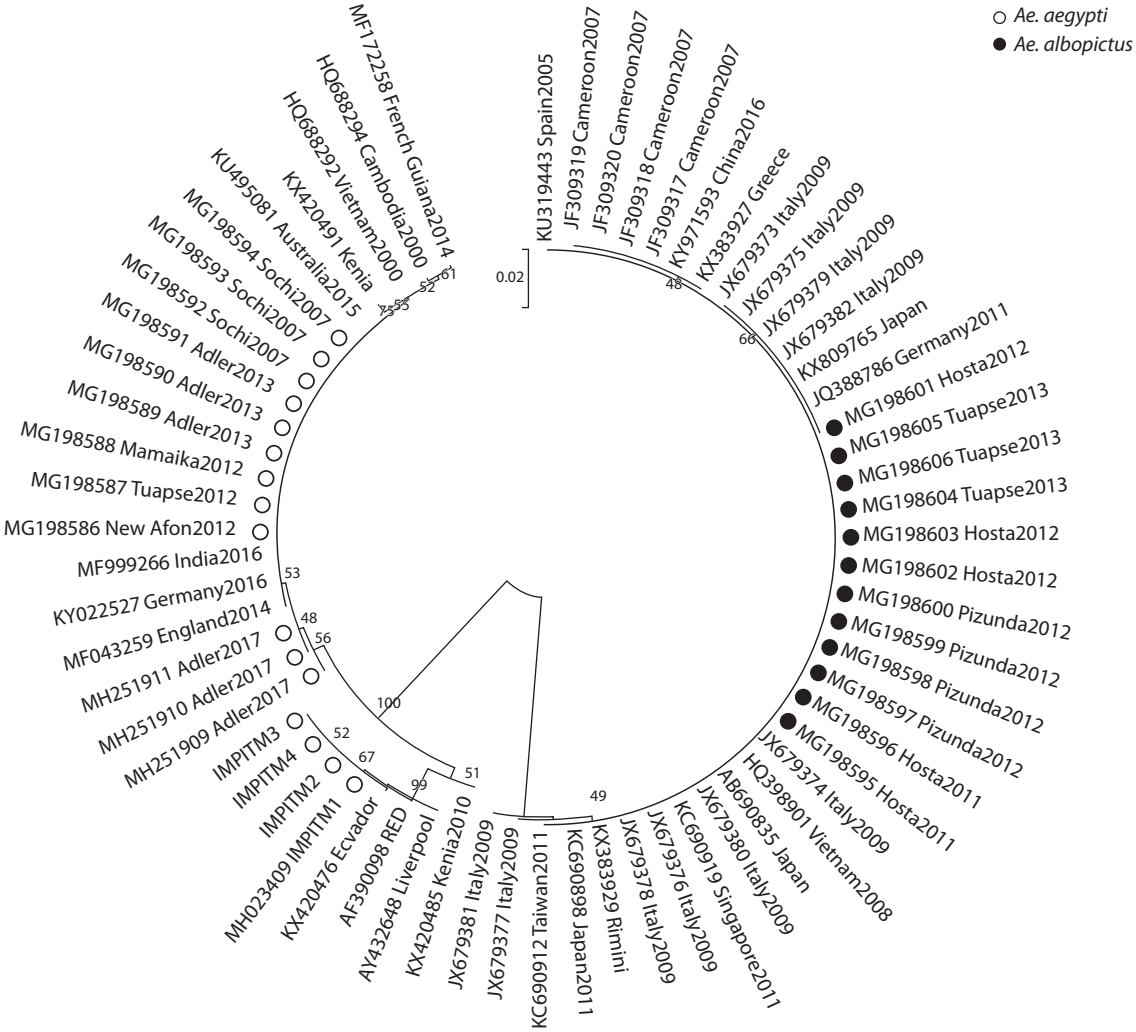



**Fig. 3.** Similarity dendrogram of the *COI* gene sequences in *Ae. aegypti* and *Ae. albopictus*.

after ITS2 comparison indicated that *Ae. aegypti* was 1.3 times closer to *Ae. koreicus* and *Ae. cretinus* than to *Ae. albopictus* (Table 4).

Comparison of the identified sequences between themselves and with their analogues from Gene Bank showed that, while the ITS2 region was variable in one mosquito species, its identical variants were present in specimens from geographically separated regions (Fig. 2). The ITS2 sequences of each species form different dendrogram clusters, their differences supported by the high values in bootstrap analyses.

**Mitochondrial DNA variability in *Ae. aegypti* and *Ae. albopictus*.** Nine *COI* sequences from wild *Ae. aegypty* (2007, 2011–2015) were identical. Same variant is present in Gene Bank – annotated *Ae. aegypty* from Cambodia (2000), India, England (2014), French Guyana (2014), Australia (2015), Germany (2016) (Fig. 3). *Ae. aegypti* (Adler, 2017) had two other different mitochondrial haplotypes (MH251909–MH251911) with synonymous nucleotide replacements C48T and additional T189C in MH251909. The genetic diversity between *Ae. aegypti* specimens from the Black sea coast for the *COI* gene was determined to be 0.5 %. The differences between the wild specimens and the lab line included 11 nucleotide substitutes, one of which (G148A) was non-synonymous (see Fig. 3).

Among *Ae. albopictus* species two mitochondrial haplotypes were found. The first one consisted of 11 identical sequences in the specimens (2011–2017) which were also registered in *Ae. albopictus* from Spain (2005), Italy (2009), China, Taiwan (2011), and Japan (2011) (see Fig. 3). The second mitochondrial haplotype that differed from the others by synonymic replacement A79G was aslo found in one mosquito from Hosta (2012) (MG198601). The haplotype identical to second one (79G) had been previously discovered in *Ae. albopictis* from northern Italy (2009), Japan and Germany (2011). The *COI* gene variability among *Ae. albopictus* was 0.4 % (see Table 4).

#### Endosymbiotic bacteria Wolbachia

The frequency *of Wolbachia* contamination was investigated in 411 specimens of *Ae. albopictus,* 50 specimens of *Ae. aegypti*, 4 – of *Ae. cretinus* and 5 – of *Ae. koreicus*.

**Table 5.** Prevalence of *w*AlbA and *w*AlbB strains of *Wolbachia* in *Ae. albopictus* from different sampling sites

| Sampling site, year | *N* | *w*AlbA | *w*AlbB | *w*AlbA + *w*AlbB | Infection±SD (%)* |
|---|---|---|---|---|---|
| Pizunda, 2012 | 16 | 0 | 10 | 4 | 87.5 ± 8.3 |
| Adler, 2013 | 30 | 0 | 18 | 0 | 60 ± 8.9 |
| Hosta, 2011 | 6 | 1 | 0 | 5 | 100 |
| Hosta, 2012 | 3 | 0 | 0 | 1 | 33.3 |
| Sochi, 2012 | 35 | 0 | 34 | 0 | 97.1 ± 2.8 |
| Mamaika, 2012 | 63 | 0 | 51 | 12 | 100 |
| Dagomys, 2012 | 27 | 0 | 24 | 3 | 100 |
| Dagomys, 2016 | 190 | 1 | 22 | 8 | 16.3 ± 2.7 |
| Lazarevskoye, 2012 | 30 | 0 | 24 | 6 | 100 |
| Tuapse, 2013 | 11 | 2 | 1 | 7 | 90.9 ± 8.7 |
| Total | 411 | 4 | 184 | 46 | 56.9 |

* Standard deviation (SD) was considered for samples of more than 10 individuals.

**Table 6.** Infection of *Ae. albopictus* and *Ae. aegypti* with *D. immitis* and *D. repens*

| Species | Number of specimens (pools) | Number of pools, infected with *D. repens* | | | Number of pools, infected with *D. immitis* | | |
|---|---|---|---|---|---|---|---|
| | | Head-thorax | Abdomen | MIR, % | Head-thorax | Abdomen | MIR,% |
| *Ae. albopictus* | 366 (74) | 1 | 0 | 0.3 | 0 | 5 | 1.4 |
| *Ae. aegypti* | 21 (4) | 1 | 0 | 4.8 | 0 | 0 | 0 |

All screened *Ae. aegypti*, *Ae. cretinus, Ae. koreicus* were not infected with *Wolbachia.* In *Ae. aibopictus* symbiotic bacteria *Wolbachia* were found in all the pools, the percentage of infected insects varied between 16.3 and 100 % (Table 5). On the Caucasus coast all possible variants of infected *Ae. albopictus* were found. Between 234 positive specimens 3 variants of infection were found: rare strain *w*AlbA (1.7 %), common strain *w*AlbB (78.6 %) and superinfection with both strains *w*AlbA and *w*AlbB (19.7 %); 177 specimens were not infected. What is especially interesting is the Dagomys population where in 2016 only 31 (16.3 %) out of the 190 mosquitoes were tested positively for *Wolbachia*. The infection rate of the Dagomys *Ae. albopictus* in 2016 was statistically different from the grand total (Fisher test, $p < 0.0001$). If the Dagomys 2016 pool were removed from analysis, the total contamination of *Ae. albopictus* would be 91.8 %.

The variations in total infected specimens in the pools can be explained by their small count, like in the Hosta case (2012) or, possibly, by the poor condition of bacterial DNA in the preserved mosquitoes. The low total of infected mosquitoes was mainly due to Dagomys (2016). In this case, 190 specimens were checked and considering good PCR results for other genes, poor DNA condition was hardly the reason. Most probably, the low rate of *Wolbachia* infection was consistent and the *Ae. albopictus* population requires further investigation (see Table 5).

### *Dirofilaria* contamination

Using PCR with specific primers for the DNA of two species of *Dirofilaria* 74 pools (366 specimens) of *Ae. albopictus,* and 4 pools (21 specimens) of *Ae. aegypti* were investigated (Table 6). Among 74 pools of *Ae. albopictus* one was infected with *D. epens* (MIR = 0.3 %), five pools were infected with *D. immitis* (MIR = 1.4 %). Only one pool of *Ae. aegypti* out of four was infected with *D. repens* (MIR = 4.8 %). *D. immitis* was found only in the abdomen pools of *Ae. albopictus*. Infective larvae L3 of *D. repens* were found in pools of head-thorax parts of the both species (see Table 6).

## Discussion

Stable, replenishing population of *Ae. albopictus* inhabits the territory of Krasnodarsky Region's Black Sea coast. Mosquitoes *Ae. albopictus,* first registered in 2011, have been expanding with the great speed presenting serious competition to *Ae. aegypti*.

Survey of the territory in years 2012–2013 showed presence in this area of consistent, replenishing populations of two dangerous mosquito species – *Ae. aegypti* and *Ae. albopictus* (Ganushkina et al., 2013). In 2012, the areas surrounding New Afon and Pizunda were additionlly investigated, with the same mosquito species discovered. In Russia *Ae. albopictus* were found only recently, in 2011, but they actively, as it is common for invasive species in a

new place, have taken their niche around the Big Sochi in the wet subtropical climate zone, pressing *Ae. aegypti* from Adler to Lazarevskoe. In typical semidry Mediterranean climate zone, Tuapse area, *Ae. aegypti* were predominant in 2011–2013, while the quantity of *Ae. albopictus* in this territory was negligible. Further to the north to Anapa, where semidry Mediterranean climate is also present, no mosquitoes of both species were found during July, September and October of 2013 (Ganushkina et al., 2013). In 2014–2015 the population of *Ae. albopictus* and *Ae. aegypti* on the Caucasus coast of the Black Sea developed in the way similar to the 2012–2013 trends (Ganushkina et al., 2016). In populated areas to the south of Tuapse, mainly *Ae. albopictus* were registered. In Tuapse *Ae. aegypti* were predominant (*Ae. aegypti* share 70 %, *Ae. albopictus* 30 %), however, the count of *Ae. albopictus* began to rise. No further advances of *Ae. albopictus* in 2012–2014 to the northwest of Jubga (57 km from Tuapse, the last locality, where *Ae. albopictus* were registered) were observed. However, as we had predicted (Ganushkina et al., 2014b). *Ae. albopictus* mosquitoes were able to expand in the northwest direction making it essential to investigate the Gelendjik area, where M.V. Zabshata (2016) had found this species in 2015.

Our collections from July and August 2016, as well as the data by Fedotova M.V. et al. (2017a, b) demonstrated that in Adler, Hosta and Sochi no *Ae. aegypti* had been found, and only *Ae. albopictus* had occupied the coast. However, during August 2017 in Adler the viable eggs of both *Ae. aegypti* and *Ae. albopictus* were found, which was confirmed by DNA tests. The *Ae. aegypti* eggs are able to withstand prolonged dehydration, but below-zero temperatures are lethal to them. Consequently, despite the drastic decrease in population, in positive conditions *Ae. aegypti* are able to regain abundance. The population decrease may be linked to the competition between *Ae. aegypti* and *Ae. albopictus* larvae for food resources and the fertility loss in *Ae. aegypti* due to possible interspecies copulation (Bargielowski et al., 2015; Carrasquilla, Lounibos, 2015). A major role in preserving the viability of diapausing eggs of *Ae. aegypti* may be attributed to winter temperatures. Theoretically, *Ae. aegypti* areal may correlate with the lowest rate of night temperature at ground surface (Tsai et al., 2018). Median borderline temperature of 13.8 degrees may play a critical role in limiting of *Ae. aegypti* expansion on the Caucasus Black Sea coast, where during some years it drops down from –3 to –13 centigrade.

Amplification of the ITS2 region enables for identification of *Ae. albopictus*, *Ae. aegypti*, *Ae. aretinus* and *Ae. koreicus* based on the size of the PCR product. This method can be used along with morphological criteria for the accurate species identification. Genetic divergence in the variable non-coding ITS2 area among *Ae. albopictus*, *Ae. cretinus*, *Ae. koreicus* and *Ae. aegypti* constitutes between 28 and 46 %. *Ae. albopictus* is genetically closer to *Ae. cretinus* and *Ae. koreicus* than to *Ae. aegypti*. The divergence between *Ae. albopictus* and *Ae. aegypti* is 37 %, and in the coding sequence of *COI* gene it reaches 14 %.

The methods of molecular genetics were first used to characterize the mosquitoes from the lab line of *Ae. aegypti* cultivated in a laboratory for more than 50 years. The origin of first mosquitoes of the IMPITM line is unknown. According to the nuclear marker, these mosquitoes are identical to the specimens in the Rockfeller line. The *COI* gene analysis has shown that mosquitoes from the *Ae. aegypti* IMPITM lab line contain DNA close to the one of African mosquitoes and the specimens from the Liverpool and RED lab lines. The databases contain neither *COI* sequences for the Rockfeller strain, nor ITS2 for Liverpool and RED. Therefore, no possibility exists to compare *Ae. aegypti* IMPITM with any of the known laboratory cultures of *Ae. aegypti* using the both markers. The origin of the laboratory strains of *Ae. aegypti*, bred in the laboratories of USA, England, France and other countries since 1940–1950s are often unknown (Kuno, 2010). The *Ae. aegypti* Rockfeller line originated from Cuba and the *Ae. aegypti* Liverpool line – possibly from West Africa (Kuno, 2010). The RED line heritage was undetermined, but we know that it is a variant of the Rex-D strain of *Ae. aegypti (*Costa-da-Silva et al., 2017).

*Ae. aegypti* originated from Africa, from where they expanded firstly to South and North Americas, and then towards Asia (Bennett et al., 2016). *Ae. aegypti* found on the Caucasus coast of the Black Sea have the *COI* gene identical to the one of invasive mosquitoes from Southeast Asia (India, Cambodia), America (French Guyana), Australia and single specimens carried to Europe (Kampen et al., 2016; Dallimore et al., 2017). These invasive *Ae. aegypti* are likely to have adaptive properties to subtropical and even moderate climates.

Southeast Asia is considered to be the place of *Ae. albopictus* origin, from where these mosquitoes expanded worldwide. Previous research of *Ae. albopictus* had shown a low level of diversity in mitochondrial DNA, but found differences in the *COI* gene between the populations that are present in countries with tropical or subtropical climate (Mousson et al., 2005; Patsoula et al., 2006; Kamgang et al., 2011; Zitko et al., 2011), where tropical populations carry 363C, and subtropical – 363T (relative to MG198595). No tropical *COI* gene haplotype was found on the Black Sea coast. One of the haplotypes found in *Ae. albopictus* during our research was typical to *Ae. albopictus* not only from Taiwan and Japan but also from Italy and Spain (see Fig. 3). The second haplotype found in *Ae. albopictus* from Hosta (2012) also presents in *Ae. albopictus* from Japan, Italy and Germany.

Investigation of *Ae. albopictus* and *Ae. aegypti* from the Black Sea coast using the markers of nuclear and mitochondrial DNA and comparison with existing databases has shown a low level of diversity among the mosquitoes of these species collected in various sites and at different times. It confirms that the worldwide expansion of invasive *Ae. aegypti,* and especially *Ae. albopictus,* has been happening very fast and no evolutionary changes have occurred so far.

Our findings confirmed the absence of symbiotic bacteria *Wolbachia* in wild *Ae. aegypti*. *Wolbachia* was not found in the specimens of *Ae. cretinus* и *Ae. koreicus*, but wider screening is required to make conclusive statements regarding symbiont presence in these species. We have revealed the circulation of two *Wolbachia* strains, *w*AlbA and *w*AlbB, in the *Ae. albopictus* population from the Caucasus Black Sea coast. The *w*AlbB strain is prevalent in our findings, which is similar for *Ae. albopictus* from different regions worldwide (Calvitti et al., 2015). It is known that the infection rate in the *Ae. albopictus* species is close to 100 %. Our values are lower, compared to what has usually been registered in *Ae. albopictus*. In the Dagomys collection (2016) less than 16 % of specimens were infected. An *Ae. albopictus* population totally free from *Wolbachia* was discovered in Vietnam in 2012 (Minard et al., 2017). Thorough investigation into *Wolbachia* infection and the genetic structure of *Ae. albopictus* should be continued in Dagomys using the markers of nuclear and mitochondrial DNA in order to determine the nature of infected and non-infected mosquitoes.

One of the goals of this study was revealing of invasive stages of microfilaria in *Ae. albopictus* and *Ae. aegypti* in order to determine their role as dirofilaria vectors. In *Ae. albopictus* total dirofilaria infestation with both species consisted of 6 infected pools out of 74 (MIR = 1.6 %). In *Ae. aegypti* one mosquito infected with *D. repens* was found in a pool of four (MIR = 4.8 %). Such high count should be attributed to low amount of *Ae. aegypti* specimens in the test. Discovery of *D. repens* DNA in the thorax part of mosquitoes points to microfilaria development at larvae stage L3 and that both *Ae. albopictus* and *Ae. aegypti* can infect humans or animals while sucking blood. Spread of dirofilariasis along the Black Sea coast is facilitated by optimal climate conditions for the development of infective agents, as well as intensive migration of people and dogs. Obligate carriers of *D. repens* and *D. immitis* are known to be carnivorous animals of feline and canine family. (Sergiev et al., 2014; Bogacheva et al., 2017). Southern Russia, due to its climate, is the region where consistent dirofilaria transmission has been taking place. In recent years, rising trend of dirofilariasis has been observed not only in animals, but also in humans (Ermakova et al., 2017; Kartashev et al., 2018). Considering the growth of invasive mosquito population (*Aedes* gene) on the Caucasus Black Sea coast, the presence of this suitable carriers may be the cause of dirofilariasis spreading.

## Acknowledgements

The authors would like to thank two anonymous review-ers for their valuable advice. The research was supported by the Russian Foundation for Basic Research, grant No. 16-04-00091.

## Conflict of interest

The authors declare no conflict of interest.

## References

Adhami J., Reiter P. Introduction and establishment of *Aedes* (*Stegomyia*) *albopictus* Skuse (Diptera: Culicidae) in Albania. J. Am. Mosq. Control Assoc. 1998;14(3):340-343.

Bargielowski I.E., Lounibos L.P., Shin D., Smartt C.T., Carrasquilla M.C., Henry A., Navarro J.C., Paupy C., Dennett J.A. Widespread evidence for interspecific mating between *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) in nature. Infect. Genet. Evol. 2015; 36:456-61. DOI 10.1016/j.meegid.2015.08.016.

Bennett K.L., Shija F., Linton Y.M., Misinzo G., Kaddumukasa M., Djouaka R., Anyaele O., Harris A., Irish S., Hlaing T., Prakash A., Lutwama J., Walton C. Historical environmental change in Africa drives divergence and admixture of *Aedes aegypti* mosquitoes: A precursor to successful worldwide colonization? Mol. Ecology. 2016;25:4337-4354. DOI 10.1111/mec.13762.

Bocková E., Rudolf I., Kočišová A., Betášová L., Venclíková K., Mendel J., Hubálek Z. *Dirofilaria repens* microfilariae in *Aedes vexans* mosquitoes in Slovakia. Parasitol. Res. 2013;112:3465-3470. DOI 10.1007/s00436-013-3526-9.

Bogacheva A.S., Shaikevich E.V., Rakova V.M., Ganushkina L.A. The fauna of bloodsucking mosquitoes in the Nizhny Novgorod Region, their infection with Dirofilaria and endosymbiotic bacteria. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2017;1:43-47. (in Russian)

Bourtzis K., Dobson S.L., Xi Z., Rasgon J.L., Calvitti M., Moreira L.A., Baton L.A., Hughes G.L., Mavingui P., Gilles J.R. Harnessing mosquito-*Wolbachia* symbiosis for vector and disease control. Acta Trop. 2014;132:150-163. DOI 10.1016/j.actatropica.2013.11.004.

Braig H.R., Zhou W., Dobson S.L., O'Neill S.L. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. J. Bacteriol. 1998;180:2373-2378.

Calba C., Guerbois-Galla M., Franke F., Jeannin C., Auzet-Caillaud M., Grard G., Pigaglio L., Decoppet A., Weicherding J., Savaill M.C., Munoz-Riviero M., Chaud P., Cadiou B., Ramalli L., Fournier P., Noël H., De Lamballerie X., Paty M.C., Leparc-Goffart I. Preliminary report of an autochthonous chikungunya outbreak in France, July to September 2017. Euro Surveill. 2017;22(39):17-00647. DOI 10.2807/1560-7917.ES.2017.22.39.17-00647.

Calvitti M., Marini F., Desiderio A., Puggioli A., Moretti R. *Wolbachia* density and cytoplasmic incompatibility in *Aedes albopictus*: Concerns with using artificial *Wolbachia* infection as a vector suppression tool. PLoS One. 2015;10(3):e0121813. DOI 10.1371/journal.pone.0121813.

Cancrini G., Frangipane di Regalbono A., Ricci I., Tessarin C., Gabrielli S., Pietrobelli M. *Aedes albopictus* is a natural vector of *Dirofilaria immitis* in Italy. Vet. Parasitol. 2003;118(3-4):195-202.

Carrasquilla M.C., Lounibos L.P. Satyrization without evidence of successful insemination from interspecific mating between invasive mosquitoes. Biol. Lett. 2015;11(9):20150527. DOI 10.1098/rsbl. 2015.0527.

Chuchuy A., Rodriguero M.S., Ferrari W., Ciota A.T., Kramer L.D., Micieli M.V. Biological characterization of *Aedes albopictus* (Diptera: Culicidae) in Argentina: implications for arbovirus transmission. Sci. Rep. 2018;8(1):5041. DOI 10.1038/s41598-018-23401-7.

Costa-da-Silva A.L., Ioshino R.S., Araújo H.R., Kojin B.B., Zanotto P.M., Oliveira D.B., Melo S.R., Durigon E.L., Capurro M.L.

Laboratory strains of *Aedes aegypti* are competent to Brazilian Zika virus. PLoS One. 2017;12(3):e0174081. DOI 10.1371/journal.pone. 0174081.

Dallimore T., Hunter T., Harbach R.E., Medlock J.M., Strode C., Vaux A.G. Discovery of a single male *Aedes aegypti* (L.) in Merseyside, England. Parasit. Vectors. 2017;10:309. DOI 10.1186/s13071-017-2251-0.

Delatte H., Paupy C., Dehecq J.S., Thiria J., Failloux A.B., Fontenille D. *Aedes albopictus*, vector of chikungunya and dengue viruses in Reunion Island: biology and control. Parasite. 2008;15(1):3-13. DOI 10.1051/parasite/2008151003.

Delisle E., Rousseau C., Broche B., Leparc-Goffart I., L'Ambert G., Cochet A., Prat C., Foulongne V., Ferre J.B., Catelinois O., Flusin O., Tchernonog E., Moussion I.E., Wiegandt A., Septfons A., Mendy A., Moyano M.B., Laporte L., Maurel J., Jourdain F., Reynes J., Paty M.C., Golliot F. Chikungunya outbreak in Montpellier, France, September to October 2014. Euro Surveill. 2015;20(17):pii=21108. DOI 10.2807/1560-7917.ES2015.20.17.21108.

Ermakova L.A., Tverdokhlebova T.I., Nagorny S.A., Pshenichnaya N.Yu., Boltachiev K.Kh. Analysis of incidence of larvae helminthiases (Echinococcosis, Toxocariasis, Dirofilariasis) in humans in the Russian Federation. Epidemiologiya i Vaktsynoprofilaktika = Epidemiology and Vaccinoprophylaxis. 2017;16(1):43-46. (in Russian)

Fedorova M.V., Ryabova T.E., Shaposhnikova L.I., Lopatina Yu.V., Sebentzova A.N., Yunicheva Yu.V. Invasive mosquito species in Sochi: larval development sites and counting methods. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2017a;4:9-15. (in Russian)

Fedorova M.V., Shvez O.G., Yunicheva Yu.V., Ryabova T.E., Medyanik I.M. Spreading of invasive mosquitoes *Aedes* (*Stegomyia*) *aegypti* (L., 1762) and *Aedes* (*Stegomyia*) *albopictus* (Skuse, 1895) (Diptera, Culicidae) in the southern Krasnodar region, Russia. Proc. of the II Symposium "Modern problems of general and special parasitology". St. Petersburg, 2017b;268-271. (in Russian)

Ganushkina L.A., Bezzhonova O.V., Patraman I.V., Tanygina E., Sergiev V.P. Distribution of *Aedes* (*Stegomyia*) *aegypti* L. and *Aedes* (*Stegomyia*) *albopictus* Skuse mosquitoes on the Black Sea coast of the Caucasus. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2013;1:45-46. (in Russian)

Ganushkina L.A., Morozova L.F., Patraman I.V., Sergiev V.P. Assessment of the risk of expansion of the habitats of the mosquitoes *Aedes aegypti* L. and *Aedes albopictus* Skuse in Russia. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2014b;4:8-10. (in Russian)

Ganushkina L.A., Patraman I.V., Rezza G., Migliorini L., Litvinov S.K., Sergiev V.P. Detection of *Aedes aegypti*, *Aedes albopictus*, and *Aedes koreicus* in the area of Sochi, Russia. Vector-Borne Zoonotic Dis. 2016;16(1):58-60. DOI 10.1089/vbz.2014.1761.

Ganushkina L.A., Rakova V.M., Ivanova I.B., Supryaga V.G., Sergiev V.P. Entomological monitoring of an area to assess Dirofilaria transmission risk. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2014a;3:9-12. (in Russian)

Ganushkina L.A., Tanygina E., Bezzhonova O.V., Sergiev V.P. Detection of *Aedes* (*Stegomyia*) *albopictus* Skuse mosquitoes in the Russian Federation. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2012;1:3-4. (in Russian)

Gutsevich A.V., Monchadskiy A.S., Shtakelberg A.A. Fauna of the USSR: Diptera. Mosquitoes. Vol. III, Iss. 4. Leningrad: Nauka Publ., 1970. (in Russian)

Jeffries C.L., Walker T. *Wolbachia* biocontrol strategies for arboviral diseases and the potential influence of resident *Wolbachia* strains in mosquitoes. Curr. Trop. Med. Rep. 2016;3:20-25. DOI 10.1007/s40475-016-0066-2.

Kamgang B., Brengues C., Fontenille D., Njiokou F., Simard F., Paupy C. Genetic structure of the tiger mosquito, *Aedes albopictus*, in Cameroon (Central Africa). PLoS One. 2011;6(5):e20257. DOI 10.1371/journal.pone.0020257.

Kampen H., Jansen S., Schmidt-Chanasit J., Walther D. Indoor development of *Aedes aegypti* in Germany, 2016. Euro Surveill. 2016;21(47):30407. DOI 10.2807/1560-7917.ES.2016.21.47.30407.

Kartashev V., Sagach O., Nikolaenko S., Chizh N., Korzan A., Ambalov Y., Bastrikov N., Ilyasov B., González-Miguel J., Morchón G., Siles-Lucas M., Simon F. Emerging Human dirofilariasis as a medical problem. Proc. of 28th Annual Meeting of the German Society for Parasitology. Berlin, 21–24 March. 2018;161.

Kronefeld M., Kampen H., Sassnau R., Werner D. Molecular evidence for the occurrence of *Dirofilaria immitis*, *Dirofilaria repens* and *Setaria tundra* in mosquitoes from Germany. Parasit. Vectors. 2014;7:30. DOI 10.1186/1756-3305-7-30.

Kuno G. Early history of laboratory breeding of *Aedes aegypti* (Diptera: Culicidae) focusing on the origins and use of selected strains. J. Med. Entomol. 2010;47:957-971. DOI 10.1603/ME10152.

Marzinovsky E.I. Measures against dengue fever in Russia. Russian Journal of Tropical Medicine, Medical and Veterinary Parazitology. 1929;7(3):162-165. (in Russian)

Medlock J.M., Hansford K.M., Versteirt V., Cull B., Kampen H., Fontenille D., Hendrickx G., Zeller H., Van Bortel W., Schaffner F. An entomological review of invasive mosquitoes in Europe. Bull. Entomol. Res. 2015;105(6):637-663. DOI 10.1017/S0007485315000103.

Minard G., Van V.T., Tran F.H., Melaun C., Klimpel S., Koch L.K., Ly Huynh Kim K., Huynh Thi Thuy T., Tran Ngoc H., Potier P., Mavingui P., Valiente Moro C. Identification of sympatric cryptic species of *Aedes albopictus* subgroup in Vietnam: new perspectives in phylosymbiosis of insect vector. Parasit. Vectors, 2017;10(1):276. DOI 10.1186/s13071-017-2202-9.

Mousson L., Dauga C., Garrigues T., Schaffner F., Vazeille M., Failloux A.-B. Phylogeography of *Aedes* (*Stegomyia*) *aegypti* (L.) and *Aedes* (*Stegomyia*) *albopictus* (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. Genet. Res., Camb. 2005;86(1):1-11. DOI 10.1017/S0016672305007627.

Patsoula E., Samanidou-Voyadjoglou A., Spanakos G., Kremastinou J., Nasioulas G., Vakalis N.C. Molecular and morphological characterization of *Aedes albopictus* in northwestern Greece and differentiation from *Aedes cretinus* and *Aedes aegypti*. J. Med. Entomol. 2006; 43(1):40-54.

Porter C.H., Collins F.H. Species-diagnostic differences in a ribosomal DNA internal transcribed spacer from the sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae). Am. J. Trop. Med. Hyg. 1991;45:271-279.

Riabova T.E., Yunicheva I.V., Markovich N.I., Ganushkina L.A., Orabey V.G., Sergiev V.P. Detection of *Aedes* (*Stegomyia*) *aegypti* L. mosquitoes in Sochi City. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2005;3:3-5 (in Russian)

Rishniw M., Barr S.C., Simpson K.W., Frongillo M.F., Franz M., Alpizar J.L. Discrimination between six species of canine mi-

crofilariae by a single polymerase chain reaction. Vet. Parasitol. 2006;135:303-314. DOI 10.1016/j.vetpar.2005.10.013.

Sergiev V.P., Supriaga V.G., Bronshtein A.M., Ganushkina L.A., Rakova V.M., Morozov E.N., Fedianina L.V., Frolova A.A., Morozova L.F., Ivanova I.B., Darchenkova N.N., Zhukova L.A. Results of studies of human dirofilariasis in Russia. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2014;3:3-9. (in Russian)

Shaikevich E.V. PCR-RFLP of the COI gene reliably differentiates *Cx. pipiens*, *Cx. pipiens* form molestus and *Cx. torrentium* of the Pipiens Complex. Eur. Mosq. Bull. 2007;23:25-30.

Simon C., Frati F., Beckenbach A., Crespi B., Liu H., Flook P. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 1994;87(6):651-701.

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 2013;30:2725-2729. DOI 10.1093/molbev/mst197.

Tokarev Y., Yudina M., Malysh J., Bykov R., Frolov A., Grushevaya I., Ilinsky Y. Prevalence rates of the *Wolbachia* endosymbiotic bacterium in natural populations of *Ostrinia nubilalis* and *Ostrinia scapulalis* (Lepidoptera: Pyraloidea: Crambidae) in South-Western Russia. Ekologicheskaya Genetika = Ecological Genetics (St. Petersburg). 2017;15(1):44-49. DOI 10.17816/ecogen15144-49. (in Russian)

Tsai P.-J., Lin T.-H., Teng H.-J., Yeh H.-C. Critical low temperature for the survival of *Aedes aegypti* in Taiwan. Parasit. Vectors. 2018; 11(1):22. DOI 10.1186/s13071-017-2606-6.

Yunicheva I.V., Ryabova T.E., Markovich N.I., Bezzhonova O.V., Ganushkina L.A., Semenov V.B., Tarkhov G.A., Vasilenko L.E., Guzeeva T.M., Sergiev V.P. First evidence for *Aedes aegypti* L. propagation in Greater Sochi and in some towns of Abkhasia. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2008;3:40-43. (in Russian)

Zabashta M.V. The expansion of *Aedes* (*Stegomia*) *albopictus* Skuse, 1885 on the Black Sea coast of Russia. Meditsinskaya Parazitologiya i Parazitarnye Bolezni = Medical Parasitology and Parasitical Disease. 2016;3:10-11. (in Russian)

Zhou W., Rousset F., O'Neil S. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc. R. Soc. B. 1998;265(1395):509-515.

Zitko T., Kovacic A., Desdevises Y., Puizina J. Genetic variation in East-Adriatic populations of the Asian tiger mosquito, *Aedes albopictus* (Diptera: Culicidae), inferred from *NADH5* and *COI* sequence variability. Eur. J. Entomol. 2011;108(4):501-508. DOI 10.14411/eje.2011.065.

Вавиловский журнал генетики и селекции. 2018;22(5):586-592
DOI 10.18699/VJ18.398

# The Demoiselle crane (*Anthropoides virgo*) population genetic structure in Russia

E.A. Mudrik[1] ✉, E.I. Ilyashenko[2], O.A. Goroshko[3, 4], T.A. Kashentseva[5], M.V. Korepov[6], I.A. Sikorskiy[7], G.S. Dzhamirzoev[8, 9], V.Yu. Ilyashenko[2], D.V. Politov[1]

[1] Vavilov Institute of General Genetics, RAS, Moscow, Russia
[2] A.N. Severtsov Institute of Ecology and Evolution, RAS, Moscow, Russia
[3] Daursky State Nature Biosphere Reserve, Nizhny Tsasuchey, Russia
[4] Institute of Nature Resources, Ecology and Cryology, RAS, Chita, Russia
[5] Oka Crane Breeding Center, Oka State Nature Biosphere Reserve, Brykin Bor, Russia
[6] Ulyanovsk State Pedagogical University named after I.N. Ulyanov, Ulyanovsk, Russia
[7] Nature Reserve "Opukskiy", Feodosia, Russia
[8] A.K. Tembotov Institute of Ecology of Mountain Territories, RAS, Nalchik, Russia
[9] Dagestan State Nature Reserve, Makhachkala, Russia

The Demoiselle crane (*Anthropoides virgo* Linneaus, 1758) is a widespread crane species of Eurasia distributed in the steppe and semi-desert zones from southeast Ukraine eastward to Northern China. The Demoiselle crane uses two wintering grounds in Africa and India corresponding to the European and Asian breeding parts of the range subdivided into several spatially separated breeding flocks. The first estimates of the genetic diversity and differentiation have been obtained from five of them: 1) Azov & Black Sea, 2) Caspian, 3) Volga & Ural, 4) South Siberian and 5) Eastern Asian sampled across the total breeding range in Russia using data from 10 microsatellite loci and the 1003-bp control region of mitochondrial DNA. In total, the Demoiselle crane demonstrates high level of observed ($H_O = 0.638 \pm 0.032$) and expected ($H_E = 0.657 \pm 0.023$) heterozygosity and haplotype diversity ($h = 0.960$). Genetic differentiation among populations has shown to be weak for both the microsatellite loci (Wright's $F_{ST} = 0.052$ or AMOVA estimate 0.016) and mtDNA ($F_{ST} = 0.040$). No evidence of significant population structuring of the Demoiselle crane has been found using the STRUCTURE analysis of multilocus microsatellite genotypes and the NETWORK grouping of control region haplotypes. Despite the haplotype diversity was high, the nucleotide diversity of the species was low ($0.0033 \pm 0.0003$). Negative but non-significant Tajima's and Fu's tests did not suggest the recent population expansion in the Demoiselle crane evolutionary history which contrasts to other cranes of the Palearctic (the Eurasian crane *Grus grus*, and the Hooded crane *G. monacha*). These data indicate more stable conditions for the Demoiselle crane breeding groups in the steppe zone in Pleistocene as compared to boreal and subarctic breeding grounds of other crane species.

Key words: *Anthropoides virgo*; microsatellite loci; Control Region; genetic variation; genetic differentiation; breeding groups.

## Популяционно-генетическая структура красавки *Anthropoides virgo* в России

Е.А. Мудрик[1] ✉, Е.И. Ильяшенко[2], О.А. Горошко[3, 4], Т.А. Кашенцева[5], М.В. Корепов[6], И.А. Сикорский[7], Г.С. Джамирзоев[8, 9], В.Ю. Ильяшенко[2], Д.В. Политов[1]

[1] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[2] Институт проблем экологии и эволюции им. А.Н. Северцова Российской академии наук, Москва, Россия
[3] Государственный природный биосферный заповедник «Даурский», Нижний Цасучей, Россия
[4] Институт природных ресурсов, экологии и криологии Сибирского отделения Российской академии наук, Чита, Россия
[5] Питомник редких видов журавлей Окского государственного биосферного заповедника, Брыкин Бор, Россия
[6] Ульяновский государственный педагогический университет им. И.Н. Ульянова, Ульяновск, Россия
[7] Природный заповедник «Опукский», Феодосия, Россия
[8] Институт экологии горных территорий им. А.К. Темботова Российской академии наук, Нальчик, Россия
[9] Государственный природный заповедник «Дагестанский», Махачкала, Россия

Красавка (*Anthropoides virgo* Linnaeus, 1758) – широко распространенный вид журавлей Евразии, гнездящийся в степной и полупустынной зонах от Юго-Восточной Украины до Северного Китая. Красавка, гнездящаяся в европейской и азиатской частях ареала, зимует в Северо-Восточной Африке и Индии соответственно. Вследствие фрагментации мест обитания, гнездовая часть ареала вида подразделена на несколько географических группировок. С использованием данных 10 микросателлитных локусов и контрольного региона митохондриальной ДНК длиной 1003 пар оснований были получены первые результаты оценки генетического разнообразия и дифференциации пяти гнездовых группировок на территории России: 1) азово-черноморской; 2) прикаспийской; 3) волго-уральской; 4) южно-сибирской и 5) восточноазиатской. В целом красавка демонстрирует высокий уровень наблюдаемой ($H_O = 0.638 \pm 0.032$) и ожидаемой ($H_E = 0.657 \pm 0.023$) гетерозиготности и гаплотипического разнообразия ($h = 0.960$). Генетическая дифференциация гнездовых группировок оказалась низкой как по микросателлитным локусам ($F_{ST}$ по Райту – 0.052, по данным AMOVA – 0.016), так и по митохондриальной ДНК ($F_{ST} = 0.040$). Не обнаружено очевидной значимой популяционной структуры *A. virgo* ни по многолокусным микросателлитным генотипам при анализе STRUCTURE, ни по гаплотипам контрольного региона в NETWORK. Несмотря на высокое гаплотипическое разнообразие, нуклеотидное разнообразие *A. virgo* оказалось низким ($0.0033 \pm 0.0003$).

Received 05.05.2018
Accepted for publication 07.06.2018


✉ e-mail: mudrik@vigg.ru

Отрицательные, но незначимые тесты Таджимы и Фу не подтвердили недавней популяционной экспансии красавки в ее эволюционной истории в отличие от других журавлей Палеарктики, например серого (*Grus grus*) и черного (*G. monacha*). Эти данные указывают на более стабильные условия для красавки в степной зоне в плейстоцене по сравнению с бореальными и субарктическими гнездовыми частями ареалов других видов журавлей.

Ключевые слова: *Anthropoides virgo*; микросателлитные локусы; контрольный регион; генетическая изменчивость; генетическая дифференциация; гнездовые группировки.

**HOW TO CITE THIS ARTICLE:**
Mudrik E.A., Ilyashenko E.I., Goroshko O.A., Kashentseva T.A., Korepov M.V., Sikorskiy I.A., Dzhamirzoev G.S., Ilyashenko V.Yu., Politov D.V. The Demoiselle crane (*Anthropoides virgo*) population genetic structure in Russia. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):586-592. DOI 10.18699/VJ18.398

The Demoiselle crane (*Anthropoides virgo* Linneaus, 1758) is one of the most abundant and the least genetically studied crane species in the world. It is a widespread species breeding in the steppe and semi-desert zone of Eurasia from southeast Ukraine eastwards through the south of Russia, Kazakhstan, Kyrgyzstan, and Mongolia to Northern China (Meine, Archibald, 1996). The main breeding range can be conditionally divided into three parts: European, South Siberian/Central Asian and East Asian. Besides, two isolated populations recently have inhabited Eastern Turkey and North Africa. The habitats in the European Russia are most fragmented and can be subdivided into Azov & Black Sea, Middle Don, Caspian and Volga & Ural/Western Kazakhstan breeding groups (Belik et al., 2011). The Demoiselle crane is phenotypically and ethologically uniform throughout the range, and variation in nesting time in the north and south can be explained by the environmental conditions.

The total population number is estimated at 200,000–220,000 individuals (Ilyashenko, 2016a), and approximately 60,000–65,000 of them inhabit Russia. Due to the total relatively high population size of the Demoiselle crane, its world conservation status as a species of the Least Concern tends to increase (www.iucnredlist.org). Red Data Book of the Russian Federation (2001) considers the Demoiselle crane as a recovering species. Despite the world high numbers, the Demoiselle crane experiences decline in several regions that led to its recognition as a locally threatened species tending to range contraction. During the 20th century it disappeared from most of territories westwards of the Black Sea as well as in Eastern Turkey and North Africa; its number continues to decline in Ukraine, Kyrgyzstan and Northern China as well as along the southern border of the range due to anthropogenic pressure, habitat degradation and long-term drought (Ilyashenko, 2016a, b).

The Demoiselle crane belongs to migratory species: birds from European and Asian parts of the breeding range use two wintering grounds in Africa and India, respectively. We are unaware of any data on genetic structure of the species in general or in particular localities including within-population levels of genetic diversity and spatial genetic differentiation. In this study, for the first time the genetic variation and differentiation of five Demoiselle crane breeding groups in Russia have been analyzed based on molecular data from nuclear microsatellite loci and mitochondrial DNA control region sequences. The objectives of our study were to compare parameters of population genetic structure of Demoiselle cranes from European and Asian parts differing by their breeding and wintering grounds and estimate the degree of genetic differentiation within and among them.

## Materials and methods

**Sample collection and DNA extraction.** We studied 115 individuals from five breeding groups of the Demoiselle crane in Russia representing almost the whole breeding range of the species. The number of studied birds from each group was the following: Azov & Black Sea (Republic of Crimea and Krasnodar Region) – 33 birds, Caspian (Republic of Kalmykia, Republic of Dagestan, Astrakhan Oblast, Stavropol Region, and the western part of Volgograd Oblast at the right Volga river bank) – 42, Volga & Ural (Saratov Oblast, Samara Oblast and eastern part of Volgograd Oblast in Transvolga) – 4, South Siberia (Republic of Khakasia and two birds from the Omsk zoo supposedly caught in the steppe zone east of Altai Mountains) – 14, and Eastern Asia (Transbaikalia) – 22 (Supplemantary Figure)[1]. Some birds were captive but they originated from the known breeding location. The biological samples from the most studied cranes were obtained from the natural populations of *A. virgo* during our own field work mainly in the Caspian and Transbaikalian regions. The biomaterial was taken partly from chicks of wild pairs caught and then released in several region of European Russia according to permits from the Federal Service for Supervision of Natural Resources (Rosprirodnadzor) No. 104, 105 and 106 from 13.06.2017. The chicks were caught and released immediately after taking their feathers later on used for DNA extraction. For DNA isolation in the wild cranes, moulted and plucked feathers were used, while in the captive birds we used the blood taken during a planned clinical examination zoos. In the blood and plucked feathers DNA was extracted using the DIAtom™ DNAPrep100 Kit (Isogen Laboratories Ltd., Russia), and in the calamus and blood clot of moulted feathers – the using innuPREP Forensic Kit (Analytik Jena, Germany) according to the manufacturers' protocols.

**Microsatellite genotyping.** Individual genotyping was performed by 10 preliminary selected heterologous polymorphic loci isolated from the genomes of the Red-crowned crane *Grus japonensis*: *Gj-M8*, *Gj-M15*, *Gj-M34* (Hasegawa et al., 2000), *Gj-4066*, *Gj-8077* (Zou et al., 2010), the Blue crane *Anthro-*

[1] Supplementary Figure is available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2018-22/appx10.pdf

*poides paradisea*: *Gpa12*, *Gpa38*, *Gpa39* (Meares et al., 2008) and the Whooping crane *Grus americana*: *Gram22*, *Gram30* (Jones et al., 2010). Polymerase chain reactions (PCR) were conducted using GenPak PCR Core Kit (Isogen Laboratories Ltd., Russia). PCR products electrophoresis was performed in 6 % polyacrylamide gel in Tris-EDTA-borate buffer system with subsequent gel staining with ethidium bromide and visualization in ultraviolet light using the Kodak Edas 290 gel documentation system (Kodak, USA). The size of observed alleles was defined by means of gel electrophoresis image analysis software GelAnalyzer (http://www.gelanalyzer.com).

**Control Region sequencing.** For the analysis of mitochondrial DNA (mtDNA) we amplified the Control Region with LC16575 (5′-ACA AAA GAA ACC CCC AAA CTC A-3′) and HC01342 (5′-AAG AAT TCT GCG GAT ACT TGC ATG T-3′) primers following the PCR procedures recommended in (Hasegawa et al., 1999). PCR products were detected by electrophoresis with 1.5 % agarose gel and then purified using Diatom DNA Clean-Up Kit (Isogen Laboratory, Russia). Subsequent sequencing was performed in both directions on ABI 3130 GeneticAnalyzer (Applied Biosystems, USA) at Evrogen (Moscow, Russia). Sequences have been deposited to GenBank under accession numbers MH286917–MH286933.

**Data analysis.** The parameters of genetic diversity, correspondence to Hardy–Weinberg equilibrium (HWE), *F*-statistics and analysis of molecular variance (AMOVA) by microsatellite loci were calculated using GenAlEx 6.5 (Peakall, Smouse, 2012). To reveal the population structure, Bayesian clustering analysis was implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000; Porras-Hurtado et al., 2013) using an admixture model. The probability of genetic clusters number K was determined in Structure Harvester (Earl, vonHoldt, 2011). Subsequent analysis of the relevant K value was done in CLUMPP v.1.1.2 (Jakobsson, Rosenberg, 2007). Visualization of the genetic structuring was realized in Distruct (Rosenberg, 2003). The alignment of the mitochondrial Control Region sequences was performed using MAFFT algorithm (Katoh et al., 2002) in Geneious 8.1.8 (Kearse et al., 2012). Alignment statistics and DNA polymorphism, Tajima's D (Tajima, 1989) and Fu's (Fu, 1997) tests were obtained from DnaSP v.5.10.01 (Librado, Rozas, 2009). Genetic subdivision ($F_{ST}$) for mtDNA data was calculated according to Hudson et al. (1992). A haplotype network diagram was constructed using the Median-joining method in Network v4 (Bandelt et al., 1999).

## Results

**Microsatellite analysis.** There were nine polymorphic microsatellite loci at Hardy–Weinberg equilibrium and only Gram30 showed the lack of heterozygotes most likely due to the presence of null-alleles. From two to ten alleles and high levels of observed ($H_O$ = 0.531–0.843) and expected ($H_E$ = 0.533–0.809) heterozygosity have been fixed in the Demoiselle crane by all studied loci (Table 1).

Generally, Demoiselle crane breeding groups demonstrate high level of genetic polymorphism by microsatellite loci: 5.2 alleles per locus ($N_A$), $H_O = 0.638 \pm 0.032$, $H_E = 0.657 \pm 0.023$ and low inbreeding coefficient ($F_{IS}$ = –0.023) insignificantly differed from zero. They are also characterized by weak genetic differentiation level by AMOVA ($F_{ST}$ = 0.016) and Wright's ($F_{ST}$ = 0.052) *F*-statistics (Table 2). Samples with low number of individuals (Volga & Ural and South Siberia) showed significant deviations in $H_O$ and $H_E$ levels that led to significant heterozygosity excess ($F_{IS}$ = –0.339) in the Volga & Ural location and deficiency of heterozygotes ($F_{IS}$ = 0.209) in the South Siberia sample. In all other samples, genotype distributions corresponded to Hardy–Weinberg equilibrium. Unbiased estimates of the expected heterozygosity were lower in European populations ($uH_E = 0.634 \pm 0.035$) as compared to Asian locations ($uH_E = 0.691 \pm 0.020$).

The European breeding groups of the Demoiselle crane were more genetically differentiated (AMOVA $F_{ST}$ = 0.021, Wright's $F_{ST}$ = 0.064) than the Asian locations (AMOVA $F_{ST}$ = 0.009, Wright's $F_{ST}$ = 0.012). The AMOVA analysis demonstrated that 94 % of Demoiselle crane genetic variation

**Table 1.** Characterization of 10 microsatellite loci in the Demoiselle crane

| Locus | Repeat motif | *A* | Observed alleles, bp | $H_O$ | $H_E$ | HWE deviation |
|---|---|---|---|---|---|---|
| *Gj-M8*£ | $(TC)_{10}$ | 2 | 106, 112 | 0.598 | 0.617 | ns |
| *Gj-M15*£ | $(GT)_{11}$ | 2 | 112, 116 | 0.643 | 0.633 | ns |
| *Gj-M34*£ | $(CA)_7$ | 4 | 126, 130, 132, 134 | 0.606 | 0.608 | ns |
| *Gpa12*† | $(GATA)_{11}$ | 7 | 230, 234, 238, 242, 246, 250, 254 | 0.678 | 0.711 | ns |
| *Gpa38*† | $(CTAT)_{13}$ | 6 | 186, 190, 194, 198, 202, 206 | 0.713 | 0.747 | ns |
| *Gpa39*† | $(GA)_2$ $(GATA)_{13}$ | 10 | 104, 112, 116, 120, 124, 128, 132, 136, 140, 144 | 0.843 | 0.809 | ns |
| *Gram22*§ | $(AAAC)_9$ | 8 | 152, 156, 160, 164, 168, 172, 176, 180 | 0.574 | 0.537 | ns |
| *Gram30*§ | $(AAGG)_7$ | 10 | 154, 158, 166, 170, 174, 178, 182, 186, 190, 194 | 0.617 | 0.726 | *(0.027) |
| *Gj4066*¥ | $(ATAG)_9$ | 4 | 133, 137, 141, 149 | 0.531 | 0.533 | ns |
| *Gj8077*¥ | $(CTG)_{13}$ | 3 | 172, 175, 178 | 0.603 | 0.606 | ns |

Note: *A* – number of alleles; bp – base pairs; $H_O$ – observed heterozygosity; $H_E$ – expected heterozygosity; HWE deviation – deviation from Hardy–Weinberg equilibrium: ns – nonsignificant, * significant at the 0.05 % probability level. Loci taken from: £ (Hasegawa et al., 2000); † (Meares et al., 2008); § (Jones et al., 2010); ¥ (Zou et al., 2010).

**Table 2.** Parameters of population genetic structure for five breeding populations of the Demoiselle crane in Russia estimated by 10 microsatellite loci

| Breeding groups | $N$ | $N_A$ | $H_O$ | $H_E$ | u$H_E$ | $F_{IS}$ | $F_{ST}$ (Wright's/AMOVA) |
|---|---|---|---|---|---|---|---|
| Azov & Black Sea | 33 | 5.3 | 0.598 ± 0.046 | 0.628 ± 0.050 | 0.637 ± 0.051 | 0.035 | |
| Caspian | 42 | 6.3 | 0.643 ± 0.057 | 0.628 ± 0.044 | 0.636 ± 0.044 | –0.025 | |
| Volga & Ural | 4 | 3.5 | 0.750 ± 0.116 | 0.551 ± 0.077 | 0.629 ± 0.088 | –0.339 | |
| Average by the European group | 79 | 5.1 | 0.664 ± 0.051 | 0.602 ± 0.033 | 0.634 ± 0.035 | –0.110 | 0.064/0.021 |
| South Siberian | 14 | 5.2 | 0.527 ± 0.070 | 0.665 ± 0.032 | 0.690 ± 0.033 | 0.209 | |
| Eastern Asian | 22 | 5.3 | 0.670 ± 0.027 | 0.677 ± 0.026 | 0.693 ± 0.027 | 0.004 | |
| Average by the Asian group | 36 | 5.3 | 0.599 ± 0.046 | 0.671 ± 0.020 | 0.691 ± 0.020 | 0.106 | 0.012/0.009 |
| Total for five breeding groups | 115 | 5.2 | 0.638 ± 0.032 | 0.630 ± 0.022 | 0.657 ± 0.023 | –0.023 | 0.052/0.016 |

Note: $N$ – sample size; $N_A$ – allele number per locus; $H_O$ – observed heterozygosity; $H_E$ – expected heterozygosity; u$H_E$ – unbiased expected heterozygosity; $F_{IS}$ – intrapopulation coefficient of inbreeding; $F_{ST}$ – among population coefficient of inbreeding.

was concentrated within individuals, 5 % – among individuals, and 1 % – among populations. No evident population structure among different geographical breeding populations of this species has been revealed by Bayesian clustering analysis in STRUCTURE based on microsatellite loci with most likely estimated cluster number K = 4 (Fig. 1). However, some increase in proportion of 'green', 'red' and 'yellow' clusters in eastern samples can be considered as a trend to starting process of differentiation among Demoiselle cranes wintering in Africa and India.

**MtDNA analysis.** The Control Region sequences of the Demoiselle Crane with full-length 1 003 bp were obtained for 23 birds from different populations. This fragment contained 20 variable sites including ten singleton sites, nine parsimony-informative sites and one inserted site. Among the 23 studied individuals, a total of 17 haplotypes were defined: 12 and 5 in European and Asian parts of the breeding range, respectively. Among European birds, eight haplotypes (H2, H3, H5, H6, H7, H9, H10, H11) were unique to the Caspian breeding group, one haplotype (H12) was unique to Volga & Ural location, and three haplotypes were shared by Azov & Black Sea and Volga & Ural (H1) and Caspian and Volga & Ural (H4, H8) breeding groups (Table 3, see Suppl. Figure). All five Asian haplotypes were unique: one to the South Siberian (H13) and four to the Eastern Asian (H14, H15, H16, H17) parts of the breeding range so as European and Asian breeding groups did not share any mitotypes.

In general, the haplotype diversity ($h$) of the mitochondrial Control Region in the Demoiselle crane was exclusively high ($h$ = 0.960 ± 0.026). The overall nucleotide diversity was low ($\pi$ = 0.0033 ± 0.0003). The small number of analyzed haplotypes did not allow us to compare all breeding populations but the genetic differentiation between European and Asian parts of the breeding range by mtDNA sequences also was low ($F_{ST}$ = 0.040) corresponding to the microsatellite data. The test for deviation from selective neutrality of Tajima ($D$ = −1.255) and Fu ($F_S$ = −9.712) were negative but non-significant and did not show the evidence of population expansion of *A. virgo*. Lack of typical star-like structure of haplotype network also suggests no recent bottlenecks and subsequent expansion (Avise, 2000). Despite five identified Asian haplotypes were unique, they are incorporated in general network and do not form a separate haplogroup. Median haplotype network did not reveal highly diverged haplogroups that can be attributed for European and Asian lineages (Fig. 2, see Suppl. Figure).

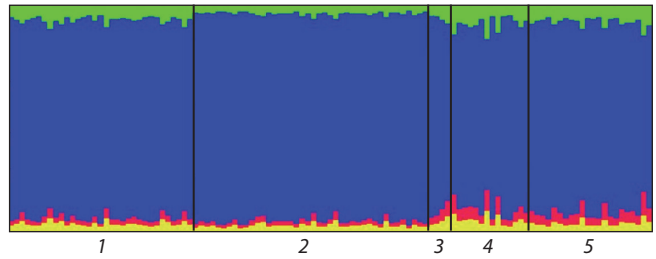



**Fig. 1.** Model-based clustering results for multi-loci individual genotypes of the Demoiselle crane iteratively assigned to four genetic groupings (K = 4).

*1* – Azov & Black Sea, *2* – Caspian, *3* – Volga & Ural, *4* – South Siberian, *5* – Eastern Asian breeding groups.

## Discussion

The Demoiselle crane demonstrates a high level of genetic diversity by nuclear microsatellite loci ($H_O$ = 0.638 ± 0.032, $H_E$ = 0.657 ± 0.023) and Control Region of mtDNA ($h$ = 0.960). We did not find a significant genetic subdivision of the species across its breeding range from the Azov & Black Sea coast to the Transbaikalia ($F_{ST}$ = 0.016 and 0.040 by microsatellite loci and mtDNA, respectively). In general, the weak genetic differentiation in bird populations is common for the migratory and especially widespread species. The Gruidae family includes the migratory and non-migratory species as well as species consisting of migratory and non-migratory subspecies or populations. Significant level of genetic differentiation due to the gene flow limitation has been revealed among the subspecies of the non-migratory Sarus crane *Grus antigone* by microsatellite loci ($F_{ST}$ = 0.210) (Jones et al., 2005) and isolated populations of the non-migratory Wattled crane *Bugeranus carunculatus* both by microsatellite loci ($F_{ST}$ = 0.100) and control region of mtDNA ($F_{ST}$ = 0.450) (Jones et al., 2006). In migratory crane species, the genetic differentia-

**Table 3.** Distribution of 17 haplotypes of the control region among five breeding groups of the Demoiselle crane

| Haplotype | GenBank accession number | Breeding group | | | | | Total |
|---|---|---|---|---|---|---|---|
| | | Azov & Black Sea | Caspian | Volga & Ural | South Siberian | Eastern Asian | |
| H1 | MH286917 | 1 | | 1 | | | 2 |
| H2 | MH286918 | | 3 | | | | 3 |
| H3 | MH286919 | | 2 | | | | 2 |
| H4 | MH286920 | | 1 | 1 | | | 2 |
| H5 | MH286921 | | 1 | | | | 1 |
| H6 | MH286922 | | 1 | | | | 1 |
| H7 | MH286923 | | 1 | | | | 1 |
| H8 | MH286924 | | 1 | 1 | | | 2 |
| H9 | MH286925 | | 1 | | | | 1 |
| H10 | MH286926 | | 1 | | | | 1 |
| H11 | MH286927 | | 1 | | | | 1 |
| H12 | MH286928 | | | 1 | | | 1 |
| H13 | MH286929 | | | | 1 | | 1 |
| H14 | MH286930 | | | | | 1 | 1 |
| H15 | MH286931 | | | | | 1 | 1 |
| H16 | MH286932 | | | | | 1 | 1 |
| H17 | MH286933 | | | | | 1 | 1 |
| Total | | 1 | 13 | 4 | 1 | 4 | 23 |

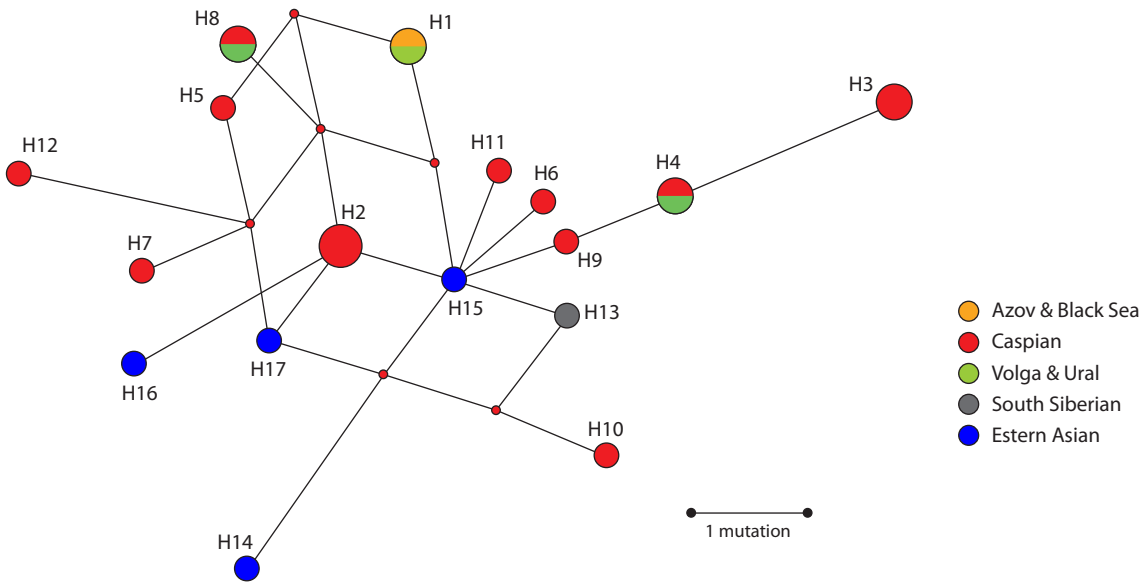



**Fig. 2.** Median-joining network of the Demoiselle crane haplotypes named as in Table 3.

The circle size is proportional to the number of individuals; the black spots represent interior nodes; connector length is proportional to the number of mutations between haplotypes.

tion of populations including isolated ones was usually low by different molecular markers not only in the wide-spread Eurasian crane *G. grus* (Haase, Ilyashenko, 2012; Mudrik et al., 2015) but also in the rare species like the Red-crowned crane *G. japonensis* having the migratory and non-migratory populations (Hasegawa et al., 1999, 2000; Sugimoto et al., 2015), the Siberian crane *G. leucogeranus* having two isolated populations (Ponomarev et al., 2004), and the Hooded crane *G. monacha* (Zhang et al., 2012). As for the Sandhill crane *G. canadensis*, the species divided to six subspecies, shows strong genetic differentiation ($F_{ST}$ = 0.480) between two mitochondrial lineages: the first one was composed of only one

migratory arctic subspecies, and the second one combined both the remaining non-migratory and migratory subspecies differentiation among which was low $F_{ST}$ = 0.066 (Rhymer et al., 2001).

We can conclude that for the Demoiselle crane Pleistocene glaciation did not cause a significant disruption of the initial common range located in the steppe zone, and they were relatively slightly affected by ice age events. Thus, fragmentation of the species range could be likely attributed to in late Holocene/Anthropocene. The Demoiselle crane did not experience substantial demographic changes like bottlenecks and expansions, retaining a typical for migratory birds level of genetic diversity and differentiation. Despite the idea of complete isolation of western breeding groups wintering in Africa from Asian breeding groups migrating for winter to India, gene flow between them may have ceased recently and/or incompletely. In other words, ever continuous range of the Demoiselle crane disrupted so late in its evolutionary history that differences has not accumulated yet as it has happened in some other species cited above. Nevertheless, the already started process of differentiation between European and Asian breeding groups of the Demoiselle crane is becoming evident from the presented multilocus analysis of genotypic variation, and the main cause of such subdivision is using different migration flyways and wintering sites by birds of these groupings. Another reason for the observed level of differentiation could be low sensitivity of the selected genetic markers (microsatellite loci and control region) so that efforts should be made to develop and use additional marker types such as nuclear gene sequences and SNPs.

## Acknowledgments

The study was supported by the project of Russian Foundation of Basic Research No. 17-04-01287, programs of Presidium RAS No. 41 “Biological diversity and biological resources of Russia” (state assignment 0112-2018-0025) and No. 32 “Evolution of organic world. Role and influence of planetary processes” (state assignment 0112-2018-0027), program of Eurasian Regional Association of Zoos and Aquariums “Conservation of cranes of Eurasia” and Budget state assignment 0112-2016-0002.

We thank K.M. Kuniev, V.G. Kalmykov, G.A. Kalmykova, G.N. Moskov, V.M. Muzaev, Yu.N. Arylov, E.V. Gugueva, T.V. Selezneva, I.P. Ariulina, A.S. Urusova, A.G. Grinko for their invaluable assistance during our fieldwork. The birds were handled in accordance with the approved national guidelines for the care and use of laboratory animals. All the experiments were approved by Ethics Committee on Animal Care and Use of Vavilov Institute of General Genetics of RAS (approval No. 1 of 15.05.2017).

## Conflict of interest

The authors declare they have no conflict of interest.

## References

Avise J.C. Phylogeography: The history and formation of species. Cambridge: Harvard Univ. Press, 2000.

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

Belik V.P., Guguyeva E.V., Vetrov V.V., Milobog Y.V. The Demoiselle crane in the northwestern Caspian lowland: distribution, number, and breeding success. Cranes of Eurasia (Biology, Distribution, Migrations, Management). 2011;4:157-174.

Earl D.A., vonHoldt B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Res. 2011;4(2):359-361. DOI 10.1007/s12686-011-9548-7.

Fu Y.X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147:915-925.

Haase M., Ilyashenko V. A glimpse on mitochondrial differentiation among four currently recognized subspecies of the common crane *Grus grus*. Ardeola. 2012;59(1):131-136.

Hasegawa O., Ishibashi Y., Abe S. Isolation and characterization of microsatellite loci in the red-crowned crane *Grus japonensis*. Mol. Ecol. 2000;9:1677-1678.

Hasegawa O., Takada S., Yoshida M.C., Abe S. Variation of mitochondrial control region sequences in three crane species, the red-crowned crane *Grus japonensis*, the common crane *G. grus* and the hooded crane *G. monacha*. Zool. Sci. 1999;16:685-692.

Hudson R.R., Slatkin M., Maddison W.P. Estimation of levels of gene flow from DNA sequence data. Genetics. 1992;132(2):583-589.

Ilyashenko E.I. Changes in Demoiselle crane status over the last 20 years. Proc. of the VIIIth Europ. Crane Conf. Gallocanta, Spain, 10–14 Nov. 2014. Association Amigos de Gallocanta, 2016a;80-88.

Ilyashenko E.I. Estimated numbers of cranes (Gruiformes, Gruidae) in Northern Eurasia at the beginning of the twenty-first century. Biol. Bull. 2016b;43(9):1048-1051. DOI 10.1134/S1062359016090119.

Jakobsson M., Rosenberg N.A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007;2(14):1801-1806. DOI 10.1093/bioinformatics/btm233.

Jones K.L., Barzen J., Ashley M.V. Geographic partitioning of microsatellite variation in the sarus crane. Anim. Conserv. 2005;8:1-8. DOI 10.1017/S1367943004001842.

Jones K.L., Henkel J.R., Howard J.J., Lance S.L., Hagen C., Glenn T.C. Isolation and characterization of 14 polymorphic microsatellite DNA loci for the endangered Whooping crane (*Grus americana*) and their applicability to other crane species. Conserv. Gen. Res. 2010;2(1):251-254. DOI 10.1007/s12686-010-9196-3.

Jones K.L., Rodwell L., McCann K.I., Verdoorn J.H., Ashley M.V. Genetic conservation of South African wattled cranes. Biol. Conserv. 2006;127:98-106. DOI 10.1016/j.biocon.2005.07.016.

Katoh K., Misawa K., Kuma K., Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059-3066. DOI 10.1093/nar/gkf436.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Mentjies P., Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28(12):1647-1649. DOI 10.1093/bioinformatics/bts199.

Librado P., Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451-1452. DOI 10.1093/bioinformatics/btp187.

Meares K., Dawson D., Horsburgh G.J., Perri M.R., Burke T., Taylor T.D. Characterisation of 14 blue crane *Grus paradisea* (Gruidae, AVES) microsatellite loci for use in detecting illegal trade. Conserv. Genet. 2008;9:1363-1367. DOI 10.1007/s10592-007-9490-0.

Meine C.D., Archibald G.W. (Eds.) The Cranes: Status Survey and Conservation Action Plan. IUCN, Gland, Switzerland, and Cambridge, U.K., 1996.

Mudrik E.A., Kashentseva T.A., Redchuk P.S., Politov D.V. Microsatellite variability data confirm low genetic differentiation of Western and Eastern subspecies of Common crane *Grus grus* L. (Gruidae, Aves). Mol. Biol. 2015;49(2):260-266. DOI 10.1134/S0026893315020090.

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.

Ponomarev A., Tatarinova T., Bubyakina V., Smagulova F., Kashentseva T., Morozov I. Variation of mitochondrial DNA D-loop sequences in the endangered Siberian crane *Grus leucogeranus* Pallas. Conserv. Genet. 2004;5:847-851.

Porras-Hurtado L., Ruiz Y., Santos C., Phillips C., Carracedo Á., Lareu M.V. An overview of STRUCTURE: applications, parameter settings, and supporting software. Front. Genet. 2013;4:98. DOI 10.3389/fgene.2013.00098.

Pritchard J.K., Matthew S., Peter D. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2000;164(4):1567-1587. DOI 10.3410/f.1015548.197423.

Red Data Book of the Russian Federation. Animals. Moscow: Astrel, 2001.

Rhymer J.M., Fain M.G., Austin J.E., Johnson D.H., Krajewski C. Mitochondrial phylogeography, subspecific taxonomy, and conservation genetics of sandhill cranes (*Grus canadensis*; Aves: Gruidae). Conserv. Gen. 2001;2:203-218.

Rosenberg N.A. Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes. 2003;4(1):137-138. DOI 10.1046/j.1471-8286.2003.00566.x.

Sugimoto T., Hasegawa O., Azuma N., Masatomi H., Sato F., Matsumoto F., Masatomi Y., Izumi H., Abe S. Genetic structure of the endangered red-crowned cranes in Hokkaido, Japan and conservation implications. Conserv. Genet. 2015;16:1395-1401. DOI 10.1007/s10592-015-0748-7.

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585-595.

Zhang L., Zhou L., Dai Y. Genetic structure of wintering Hooded Cranes (*Grus monacha*) based on mitochondrial DNA D-loop sequences. Chinese Birds. 2012;3(2):71-81. DOI 10.5122/cbirds.2012.0012.

Zou H.F., Dong H.Y., Kong W.Y., Ma J., Liu J. Characterization of 18 polymorphic microsatellite loci in the red-crowned crane (*Grus japonensis*), an endangered bird. Anim. Sci. J. 2010;81(4):519-522. DOI 10.1111/j.1740-0929.2010.00779.x.

Вавиловский журнал генетики и селекции. 2018;22(5):593-599
DOI 10.18699/VJ18.399



# Экспрессия пальмитилтрансфераз в структурах мозга мышей с генетически детерминированной предрасположенностью к депрессивно-подобному поведению

Е.М. Кондаурова[1, 2], Т.В. Ильчибаева[1], А.С. Цыбко[1], Е.Г. Понимаскин[3], В.С. Науменко[1] ✉

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Институт клеточной нейрофизиологии, Высшая медицинская школа Ганновера, Германия, Ганновер

Большинство сопряженных с G-белками мембранных рецепторов подвергается различным посттрансляционным модификациям. Среди таких модификаций – осуществляемое специализированными ферментами пальмитилтрансферазами S-пальмитилирование. Оно представляет собой ковалентное присоединение длинной цепи жирной кислоты пальмитата к цистеиновым аминокислотным остаткам. Пальмитилирование может существенно влиять на функцию рецепторов, сопряженных с G-белками, модифицируя их стабильность, транспортировку и функциональную активность. Очевидно, что нарушения в работе этого класса рецепторов могут приводить к возникновению самых разнообразных психопатологий, включая депрессию. Тем не менее в настоящее время связи между пальмитилтрансферазами и депрессивно-подобным поведением не установлено. Нет данных и о регион-специфических особенностях экспрессии пальмитилтрансфераз в структурах мозга. В настоящей работе исследована экспрессия пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21 в структурах мозга мышей линии ASC с генетической предрасположенностью к депрессивно-подобному поведению по сравнению с мышами родительской линии CBA, у которых не наблюдается депрессивно-подобного поведения. Продемонстрированы регион-специфические особенности при иммунодетекции белков пальмитилтрансфераз. При детекции белка ZDHHC5 в среднем мозге выявлено две полосы массой 75 и 55 кДа. При иммунодетекции белка ZDHHC21 обнаружено две полосы: первая массой 27 кДа во фронтальной коре и среднем мозге, тогда как в гиппокампе антитела визуализировали полосу белка массой 32 кДа. При иммунодетекции белка ZDHHC9 выявлено по две полосы белка в среднем мозге и гиппокампе: первая полоса массой 46 кДа, вторая – 41 кДа. Однако мыши линии ASC практически не отличались от мышей CBA по экспрессии исследуемых пальмитилтрансфераз. Таким образом, нами впервые показаны регион-специфические особенности экспрессии исследованных пальмитилтрансфераз в структурах мозга. В то же время установлено, что генетически детерминированное депрессивно-подобное поведение у мышей линии ASC не связано с изменениями экспрессии пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21.

Ключевые слова: ZDHHC5; ZDHHC9 и ZDHHC21 пальмитилтрансферазы; гены *Zdhhc5*, *Zdhhc9*, *Zdhhc21*; экспрессия генов; уровень белка; мыши ASC; депрессивно-подобное поведение.

## Expression of palmitoyl transferases in brain structures of mice genetically predisposed to depressive-like behavior

E.M. Kondaurova[1, 2], T.V. Ilchibaeva[1], A.S. Tsybko[1], E.G. Ponimaskin[3], V.S. Naumenko[1] ✉

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Institute of Cellular Neurophysiology, Hannover Medical High School, Hannover, Germany

Most G-coupled receptors undergo posttranslational modifications. Among these modifications is S-palmitoylation, carried out by specialized enzymes palmitoyl transferases. Palmitoylation is the covalent attachment of a long-chain fatty acid, palmitate, to cysteine residues. It can influence receptor stability, transportation, and function. Obviously, malfunction of G-protein coupled receptors can cause various psychic disorders, including depression. However, no association between palmitoyl transferases and depressive-like behavior has been found hitherto. There is no information on brain structure specific features of palmitoyl transferase expression either. Here we investigate the expression of ZDHHC5, ZDHHC9, and ZDHHC21 palmitoyl transferases in brain structures of ASC mice with genetic predisposition to depressive-like behavior in comparison with "nondepressive" CBA mice. Several brain region-specific features were detected in the immunodetection of palmitoyl transferase proteins. Western blot of the ZDHHC5 protein in the midbrain revealed two bands at 75 kDa and 55 kDa. Immunodetection of ZDHHC21 palmitoyl transferase revealed two bands. One of them was visualized at 27 kDa in the frontal cortex and midbrain. The other, at 32 kDa in the hippocampus. Probing for ZDHHC9 also showed two bands in each of the midbrain and hippocampus, at 46 and 41 kDa. However, the expression of all investigated palmitoyl transferases in ASC mice with depressive-like behavior was almost identical to those in CBA mice. Thus, it was the first detection of brain region-specific features of the expression of investigated palmitoyl transferases. However, the study demonstrates that

Received 22.02.2018
Accepted for publication 22.05.2018


✉ e-mail: naumenko2002@mail.ru

the genetic predisposition to depression-like behavior in ASC mice is not associated with changes in ZDHHC5, ZDHHC9, or ZDHHC21 palmitoyl transferase expression.

Key words: ZDHHC5; ZDHHC9 and ZDHHC21 palmitoyl transferases; *Zdhhc5*, *Zdhhc9*, *Zdhhc21* genes; gene expression; protein level; ASC mice; depressive-like behavior.

**КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:**

Кондаурова Е.М., Ильчибаева Т.В., Цыбко А.С., Понимаскин Е.Г., Науменко В.С. Экспрессия пальмитилтрансфераз в структурах мозга мышей с генетически детерминированной предрасположенностью к депрессивно-подобному поведению. Вавиловский журнал генетики и селекции. 2018;22(5):593-599. DOI 10.18699/VJ18.399

**HOW TO CITE THIS ARTICLE:**

Kondaurova E.M., Ilchibaeva T.V., Tsybko A.S., Ponimaskin E.G., Naumenko V.S. Expression of palmitoyl transferases in brain structures of mice genetically predisposed to depressive-like behavior. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):593-599. DOI 10.18699/VJ18.399 (in Russian)

Различные посттрансляционные модификации могут существенно влиять на функцию рецепторов, сопряженных с G-белком. Например, липидные модификации служат механизмом, посредством которого нейроны обеспечивают распределение белков. Эти модификации увеличивают гидрофобность белков, облегчая их встройку во внутриклеточную или плазматическую мембрану. Миристилирование, пренилирование и пальмитилирование – наиболее широко распространенные липидные модификации. Пальмитилирование часто наблюдается у нейрональных белков. Это уникальная, зачастую обратимая, посттрансляционная модификация трансмембранных белков потенциально регулирует их процессинг и локализацию на определенных мембранных микродоменах, например липидных рафтах, или изменяет конформацию белка, что оказывает влияние на его функциональную активность и возможность взаимодействовать с другими белками (Fukata Y., Fukata M., 2010).

Участие пальмитилирования в транспортировке и/или локализации рецептора на мембранных субдоменах показано в эксперименте, продемонстрировавшем, что серотониновые 5-$HT_{1A}$ рецепторы дикого типа остаются на липидных рафтах, тогда как количество мутантного непальмитилируемого 5-$HT_{1A}$ рецептора в этих микродоменах было значительно снижено (Renner et al., 2007). Липидные рафты, как известно, часто выполняют роль платформы, обеспечивающей совместную локализацию GPCRs (рецепторов, сопряженных с G-белком (G-protein-coupled receptors)) с определенными G-белками. Эта особенность позволяет предположить, что зависимая от пальмитилирования локализация 5-$HT_{1A}$ рецепторов на липидных рафтах важна для опосредованной рецептором передачи сигнала (Gorinski, Ponimaskin, 2013). Для того же серотонинового 5-$HT_{1A}$ рецептора показано, что замена необходимых для пальмитилирования рецептора аминокислотных остатков цистеина Cys417 и Cys420 на серин приводит к нарушению взаимодействия рецептора и его G-белка (с Gαi-субъединицей). Непальмитилируемый мутантный 5-$HT_{1A}$ рецептор демонстрировал также неспособность ингибировать продукцию цАМФ (циклический аденозинмонофосфат). Полученные результаты говорят о том, что пальмитилирование 5-$HT_{1A}$ рецептора важно для сопряжения рецептора с G-белком и трансдукции сигнала. Более того, нарушение пальмитилирования 5-$HT_{1A}$ рецептора приводило к нарушению рецептор-зависимой активации ERK киназ, что указывает на важность пальмитилирования 5-$HT_{1A}$ рецептора для трансдукции сигнала по Gβγ-опосредованному пути (Papoucheva et al., 2004).

Таким образом, пальмитилирование может существенно влиять на функциональную активность рецепторов, сопряженных с G-белками, а нарушение пальмитилирования этих рецепторов в свою очередь может привести к возникновению самого широкого спектра поведенческих и психических нарушений. Действительно, показано вовлечение пальмитилтрансфераз, ответственных за пальмитилирование различных белков, в механизмы таких психопатологий как, например, болезнь Альцгеймера, болезнь Хантингтона, а также в механизмы шизофрении и умственных отклонений (Cho, Park, 2016). Так, подавление пальмитилирования белка-предшественника амилоида и белка хантингтина приводит к нарушению их процессинга (Cho, Park, 2016). Тем не менее сведения о роли пальмитилирования определенных нейрональных рецепторов в регуляции различных форм поведения крайне скудны. Однако имеются данные о вовлечении ZDHHC9 в механизмы умственной отсталости (Fukata Y., Fukata M., 2010), а ZDHHC5 – в механизмы условно-рефлекторного замирания и гиппокамп-зависимого обучения (Li et al., 2010). Показано также, что ZDHHC21 принимает участие в регуляции воспалительного ответа (Beard et al., 2016).

Многие современные гипотезы о патогенезе депрессии указывают на ключевую роль серотониновой системы мозга (Maes, Meltzer, 1995; Harro, Oreland, 1996; Duman et al., 1997). Серотониновые нейроны, локализованные в ядрах шва (raphe nuclei) среднего мозга, дают обширные проекции во многие области мозга (Jacobs, Azmitia, 1992), которые вовлечены в развитие депрессии, в том числе в кору и гиппокамп (Neumeister et al., 2005; Ressler, Mayberg, 2007).

В связи с вышесказанным целью нашей работы было изучение экспрессии пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21 как на уровне мРНК, так и на уровне белка в структурах мозга (средний мозг, кора и гиппокамп) мышей с генетической предрасположенностью к депрессивно-подобному поведению (мыши линии ASC), по сравнению с мышами родительской линии CBA, не проявляющими депрессивно-подобного поведения.

## Материалы и методы

**Животные.** Опыты проводили на взрослых самцах мышей линий CBA и ASC. Мыши линии ASC (Antidepressant Sensitive Catalepsy) были созданы в лаборатории нейро-

геномики поведения Института цитологии и генетики Сибирского отделения Российской академии наук (ИЦиГ СО РАН) в результате длительной селекции гибридов каталептической линии CBA и некаталептической линии AKR на повышенную предрасположенность к каталепсии (Базовкина и др., 2005; Kondaurova et al., 2006). Полученные мыши демонстрировали ряд депрессивно-подобных характеристик, включая повышенную неподвижность в тестах принудительного плавания и подвешивания за хвост (Базовкина и др., 2005), нарушения иммунной системы (Альперина и др., 2007). При этом введение экзогенного BDNF приводило у мышей ASC к нормализации поведенческих показателей и активации серотониновой системы мозга (Naumenko et al., 2012). Мышей содержали в пластиковых клетках размером 40 × 30 × 15 см в стандартных условиях (температура 20–22 °С, относительная влажность 50–60 %, регулируемый световой режим (12 ч света и 12 ч темноты)) со свободным доступом к стандартной пище и воде. Все процедуры выполняли в соответствии с международными правилами обращения с животными (National Institute of Health Guide for the Care and Use of Laboratory Animals, NIH Publications No. 80023, 1996). Количество животных в каждой группе: $n \geq 7$.

**Экспрессию генов** определяли с помощью количественного метода ОТ-ПЦР (ОТ – обратная транскрипция; ПЦР – полимеразная цепная реакция), разработанного в лаборатории нейрогеномики поведения ИЦиГ СО РАН (Kulikov et al., 2005; Науменко, Куликов, 2006; Naumenko et al., 2008). Использовали два типа стандартов: внутренний и внешний. Внутренний стандарт (мРНК *Gapdh* – glyceraldehyde 3-phosphate dehydrogenase (глицеральдегид-3-фосфат дегидрогеназа) применяли для контроля обратной транскрипции в качестве основы для расчета уровня мРНК исследуемых генов. В предварительных опытах не были выявлены различия в уровне мРНК *Gapdh* в исследуемых структурах мозга. Внешним стандартом служила геномная ДНК мыши известной концентрации, что позволило контролировать ПЦР и определять число копий мРНК исследуемых генов и *Gapdh* в образцах.

*Выделение общей РНК* осуществляли с помощью TRIzol Reagent (“Lifetechnologies”, США) в соответствии с инструкцией производителя. Общую РНК обрабатывали ДНКазой без РНКазной активности (RQ1RNase-Free DNase cat. # M6101, Promega Corporation, США) в соответствии с протоколами производителей. Присутствие примесей геномной ДНК в препаратах РНК определяли в соответствии с протоколом, описанным ранее (Науменко, Куликов, 2006; Naumenko et al., 2008). РНК разводили водой до концентрации 0.125 мкг/мкл и хранили при –70 °С.

*Реакция обратной транскрипции.* Общую РНК (8 мкл, или 1 мкг) смешивали со 180 нг статистического праймера длиной 6 нуклеотидов (конечная концентрация праймера – 5 мкМ) и 2.25 мкМ стерильного KCl в объеме 16 мкл, денатурировали при 94 °С в течение 5 мин на амплификаторе «БИС» М-120 (БИС-Н, Россия), после чего проводили отжиг при 41 °С в течение 15 мин, затем добавляли 15 мкл смеси, содержащей обратную транскриптазу M-MLV (200 ед.), Трис-HCl (pH = 8.3, 0.225 мкМ), смесь dNTP (0.015 мкМ каждого), DTT (0.225 мкМ) и $MnCl_2$ (0.03 мкМ). Полученную смесь (конечный объем – 31 мкл) инкубировали при 41 °С в течение 60 мин. Синтезированная кДНК хранилась при температуре –20 °С.

**Table 1.** Primer sequences, annealing temperatures, and amplicon length

| Gene | Sequence | Annealing temperature, °C | Amplicon length, bp |
|---|---|---|---|
| *Gapdh* | F5′-tttgaagacgccagggaaatg-3′<br>R5′-tgtccagaatcaaccaccaag-3′ | 63 | 242 |
| *Zdhhc21* | F5′-aacagatggatggtgggcac-3′<br>R5′-gtttcatcccaatcactgccttc-3′ | 63 | 289 |
| *Zdhhc9* | F5′-gagtagtcccaaagcccatc-3′<br>R5′-aggatgaggaagagggtcag-3′ | 61 | 177 |
| *Zdhhc5* | F5′-ccacacctctgcacctacac-3′<br>R5′-agggatcaaggaagggaaggtc-3′ | 65 | 217 |

*Полимеразная цепная реакция в реальном времени*. Праймеры, используемые для амплификации кДНК исследуемых генов (табл. 1), разработаны на основе последовательностей, опубликованных в базе данных EMBL Nucleotide database, при помощи инструментов Oligoanalizer (https://eu.idtdna.com/calc/analyzer) и Ensemble (https://www.ensembl.org/index.html) и синтезированы в компании «Биосан» (Новосибирск). 1 мкл кДНК смешивали с 19 мкл Мастер Микс (R-412, «Синтол», Россия), содержащего интеркалирующий краситель SYBR Green I. ПЦР проводили на амплификаторе LightCycler 480 System (Roche, Швейцария) в соответствии со следующим протоколом: 3 мин 94 °С, 1 цикл; 10 с при 94 °С, 30 с при соответствующей температуре отжига (см. табл. 1), 30 с при 72 °С 40 циклов. Серию разведений геномной ДНК с концентрацией 0.125, 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64 и 128 нг/мкл амплифицировали одновременно в отдельных пробирках и использовали как внешний экзогенный стандарт для построения калибровочной кривой. Калибровочная кривая в координатах Ct – log P количества стандарта ДНК была построена автоматически программным обеспечением LightCycler 480 System. Для контроля специфичности амплификации использовали анализ кривых плавления ПЦР-продукта для каждой реакции с каждой парой праймеров. Экспрессия генов представлена как отношение количества кДНК исследуемых генов к 100 копиям внутреннего стандарта – кДНК *Gapdh*.

**Определение количества белка.** Уровни белков пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21 определяли при помощи Вестерн-блот анализа. Для получения препаратов плазматических мембран белка соответствующие структуры мозга (кора, гиппокамп, средний мозг) гомогенизировали в трис-хлоридном буфере, содержащем 300 мМ сахарозы, 10 мМ Трис HCl, pH 7.2, 1 мМ ЕДТА, 5 мМ β-меркаптоэтанол и ингибиторы протеаз (ThermoFisher Scientific Inc., cat. No. 88265). Ингибиторы протеаз использовали в рекомендованных производителем концентрациях. Гомогенат центрифугировали на 500 g 15 мин при 4 °С, отбирали супернатант и центрифугировали на 20 000 g при 4 °С в течение одного часа. Отбирали супернатант, оставшийся в пробирке осадок ресуспен-

**Table 2.** Antibodies for probing and immunoassay conditions

| Antibodies, manufacturer | Dilution | Incubation time, conditions |
|---|---|---|
| Primary antibodies | | |
| chicken against ZDHHC5 protein, ProSci, cat. No 54-211 | 1:400 | 2 h, room temperaturev |
| rabbit against ZDHHC9 protein, Sigma-Aldrich, cat. No SAB4502104 sigma | 1:200 | » |
| rabbit ZDHHC21 protein, ThermoFisher Scientific Inc. cat. No PA5-25096 | 1:200 | Overnight, 4 °C |
| against GAPDH conjugated with horseradish peroxidase, Santa Cruz, США, cat. No sc 25778 | 1:500 | 2 h, room temperature |
| Secondary antibodies | | |
| rabbit anti-chicken antibodies conjugated with horseradish peroxidase, ThermoFisher Scientific, Inc. cat. No 31401 | 1:15000 | 1 h, room temperature |
| goat anti-rabbit antibodies conjugated with horseradish peroxidase, Santa Cruz, США, cat. No sc 2004 | 1:10000 | » |

зировали в гомогенизирующем буфере (Каткова и др., 2009). Количество общего белка оценивали, используя коммерческий набор Pierce BCA Protein Assay kit (ThermoFisher Scientific Inc., США). Образцы приводили к равной концентрации (1мг/мл) с помощью Леммли буфера (Трис 62 мМ, сахароза 10 %, SDS 2 %, β-меркаптоэтанол 5 %) и денатурировали в течение 15 мин на 42 °С.

Белок разделяли с помощью SDS-PAGE гель-электрофореза, используя 10 %-й разделяющий гель для пальмитилтрансфераз ZDHHC5 и ZDHHC9, 12 %-й – для пальмитилтрансферазы ZDHHC21, и переносили с помощью полусухого электроблоттинга на нитроцеллюлозную мембрану (Bio-Rad, США) в течение ночи при силе тока 50 мА. Для переноса использовали буфер, содержащий 190 мМ глицина, 25 мМ Трис pH 8.3 и 20 % метанола. В качестве маркера использовали Broad Range Markers (Santa Cruz, США, sc-2361). Для иммунодетекции белка мембрану блокировали с 5 %-м молоком, разведенным в TBS-T буфере (Tris Buffered Saline with Tween 20, Santa Cruz, США), в течение одного часа при комнатной температуре и инкубировали с первичными антителами для ZDHHC5, ZDHHC9 и ZDHHC21 (табл. 2). После детекции белков (как описано ниже) все мембраны обрабатывали с помощью Restore Plus Western Blot Stripping Buffer (ThermoFisher Scientific Inc., cat. No. 46430) для дальнейшей реиммунодетекции внутреннего контроля GAPDH (см. табл. 2). Для детекции всех белков мембрану отмывали 5 × 5 мин буфером TBS-T, добавляли вторичные поликлональные антитела (см. табл. 2). Повторяли отмывку мембраны.

Все связанные антитела визуализировали с помощью Super Signal TM West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific Inc., США) в соответствии с инструкцией производителя и гельдокументирующей системы Fusion FX7-820 System (Vilber Lourmat, Франция). Полученное изображение денситометрировали, и количественно оценивали содержание белка при помощи программы Scion Image. Экспрессию белка выражали в относительных единицах и нормировали на экспрессию GAPDH, которая конститутивна для клеток мозга.

Анализ выявленных вариантов белков осуществляли при помощи сравнения с уже известными изоформами, опубликованными в базе данных http://www.uniprot.org (http://www.uniprot.org/uniprot/Q8VDZ4).

**Статистика.** Результаты представляли как m ± SEM (среднее ± стандартная ошибка среднего) и сравнивали их с использованием однофакторного дисперсионного анализа ANOVA с последующим множественным сравнением по Фишеру. Критерий значимости результатов – $p < 0.05$.

## Результаты

**Исследование уровня мРНК.** Однофакторный дисперсионный анализ не выявил различий в экспрессии генов пальмитилтрансфераз *Zdhhc5* ($F_{1.17} = 0.50$, $p > 0.5$) и *Zdhhc9* ($F_{1.17} = 1.6$, $p > 0.2$) в гиппокампе мышей исследованных линий (рис. 1, *а*, *б*). Однако экспрессия гена *Zdhhc21* в гиппокампе у мышей линии ASC была статистически достоверно ниже по сравнению с экспрессией этого гена у мышей линии СВА ($F_{1.17} = 5.40$, $p < 0.05$) (рис. 1, *в*). Во фронтальной коре и среднем мозге мышей исследованных линий различий по экспрессии генов *Zdhhc5*, *Zdhhc9* и *Zdhhc21* не обнаружено (см. рис. 1).

**Исследование уровня белка.** При оценке уровня белка ZDHHC5 в гиппокампе различий между изучаемыми линиями животных не установлено ($F_{1.14} = 0.14$, $p = 0.72$) (рис. 2). Интересно отметить, что в среднем мозге при иммунодетекции этой пальмитилтрансферазы было обнаружено две полосы массой 75 и 55 кДа (см. рис. 2, *в*). Тем не менее достоверной разницы между линиями ASC и СВА не выявлено как для полосы массой 75 кДа ($F_{1.14} = 0.03$, $p = 0.86$) (см. рис. 2, *а*), так и для второй полосы массой 55 кДа ($F_{1.13} = 1.36$, $p = 0.26$) (см. рис. 2, *б*). Во фронтальной коре различий по уровню белка ZDHHC5 между мышами ASC и СВА также не обнаружено (см. рис. 2).

При иммунодетекции белка ZDHHC21 также было выявлено две полосы, детектируемые антителами: первая полоса массой 27 кДа (рис. 3, *б*) во фронтальной коре и среднем мозге, а вторая – массой 32 кДа в гиппокампе (см. рис. 3, *б*). Однако уровень экспрессии этих бэндов не отличался между исследованными линиями мышей во всех структурах мозга (фронтальная кора – $F_{1.14} = 0.01$, $p = 0.98$, средний мозг – $F_{1.15} = 0.43$, 4.38 $p = 0.54$ и гиппокамп – $F_{1.16} = 2.00$, $p = 0.18$) (см. рис. 3, *а*).

Определение количества белка ZDHHC9 не выявило различий между мышами линий ASC и СВА во всех исследованных структурах мозга. Однако при иммунодетекции

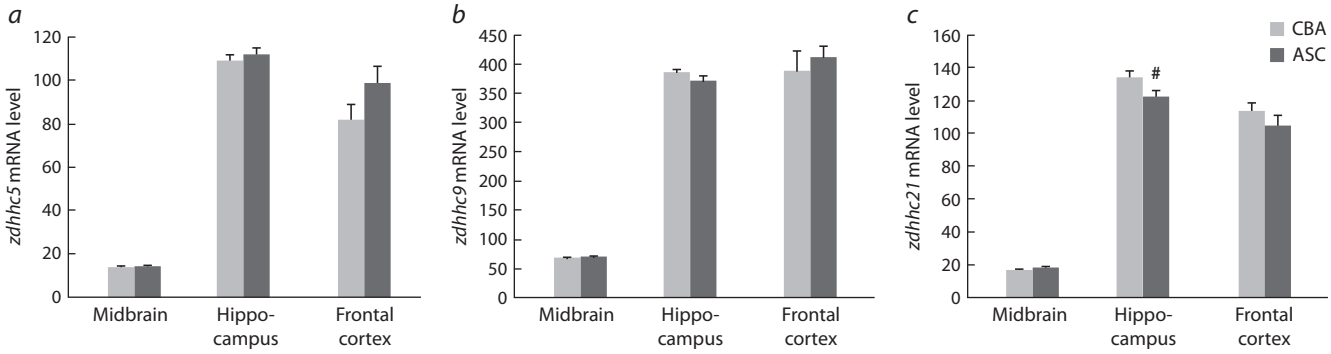



**Fig. 1.** Levels of mRNAs of the genes (*a*) *Zhhc5*; (*b*) *Zhhc9*; (*c*) *Zhhc21* in brain divisions of CBA and ASC mice.

Gene expression is presented as the number of copies of the cDNA of the corresponding gene per 100 copies of Gapdh cDNA. # p <0.05 compared to CBA. The numbers of samples in groups are $N \geq 9$.

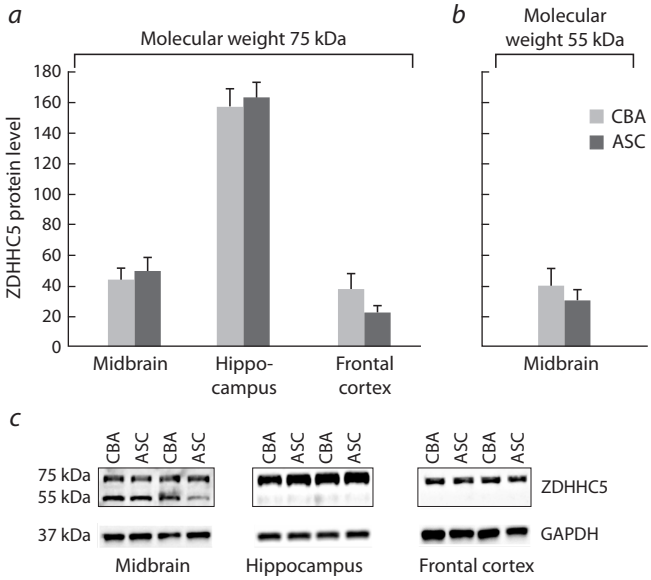



**Fig. 2.** The level of ZDHHC5 protein in brain divisions of CBA and ASC mice.

(*a*) Quantitative evaluation of the intensity of the chemiluminescent signal at 75 kDa. The ZDHHC5 protein level is presented in relative units normalized to the corresponding level of GAPDH. (*b*) Quantitative evaluation of the intensity of the chemiluminescent signal at 55 kDa in the midbrain of CBA and ASC mice. The ZDHHC5 protein level is presented in relative units normalized to the corresponding level of GAPDH. (*c*) Western blotting on the membrane. The numbers of samples in groups are $N \geq 7$.

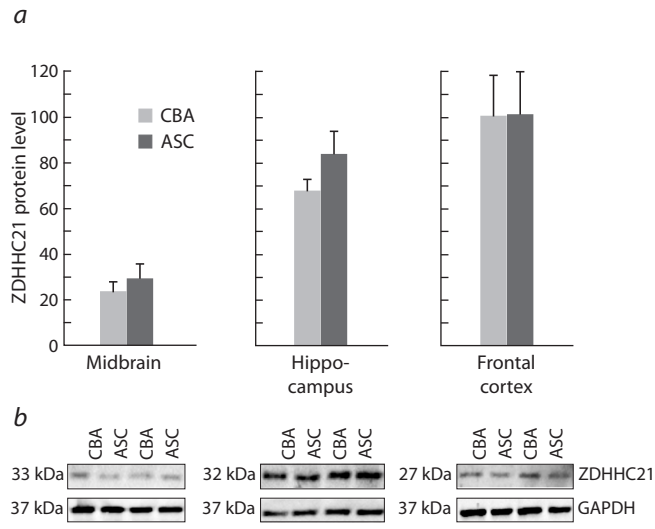



**Fig. 3.** The levels of ZDHHC21 protein in brain divisions of CBA and ASC mice.

(*a*) Quantitative evaluation of the intensity of the chemiluminescence signal at 27 kDa in the midbrain and in the frontal cortex and at 32 kDa in the hippocampus. The ZDHHC21 protein level is presented in relative units normalized to the corresponding level of GAPDH. (*b*) Western blotting on the membrane. The numbers of samples in groups are $N \geq 7$.

этой пальмитилтрансферазы было обнаружено по две полосы белка в среднем мозге и гиппокампе: первая полоса массой 46 кДа (рис. 4, *в*) и вторая, которой не наблюдалось во фронтальной коре, массой 41 кДа (см. рис. 4, *в*). Следует отметить, что в среднем мозге интенсивность нижней полосы (41 кДа) намного сильнее целевой полосы массой 46 кДа (см. рис. 4, *в*). Разницы по уровню белка как массы 46 кДа (средний мозг – $F_{1.15} = 0.22$, $p = 0.64$ и гиппокамп – $F_{1.16} = 2.17$, $p = 0.16$) (см. рис. 4, *а*), так и массы 41 кДа (средний мозг – $F_{1.16} = 0.39$, и гиппокамп – $F_{1.13} = 0.02$, $p = 0.89$) (см. рис. 4, *б*) обнаружено не было.

## Обсуждение

В настоящей работе впервые исследована экспрессия пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21 на уровне мРНК и белка во фронтальной коре, гиппокампе и среднем мозге двух линий мышей: не предрасположенной к депрессивно-подобному поведению CBA и ASC, характеризующейся генетической предрасположенностью к депрессивно-подобному поведению. Показано, что в гиппокампе мышей обеих линий белок ZDHHC21 визуализировался в виде полосы массой 32 кДа, что отличает ее от основной полосы массой 27 кДа, определяемой как основной бэнд белка ZDHHC21, экспрессируемый во фронтальной коре и среднем мозге. Можно предположить, что в гиппокампе происходят постранслящионные модификации ZDHHC21, что может играть существенную роль в регуляции функции белка (Varki, 1993; Dutta et al., 2017). В базе данных http://www.uniprot.org есть информация о постранслящионных модификациях только для пальмитилтрансферазы ZDHHC5 (http://www.uniprot.org/uniprot/Q8VDZ4). Для двух других исследованных в на-

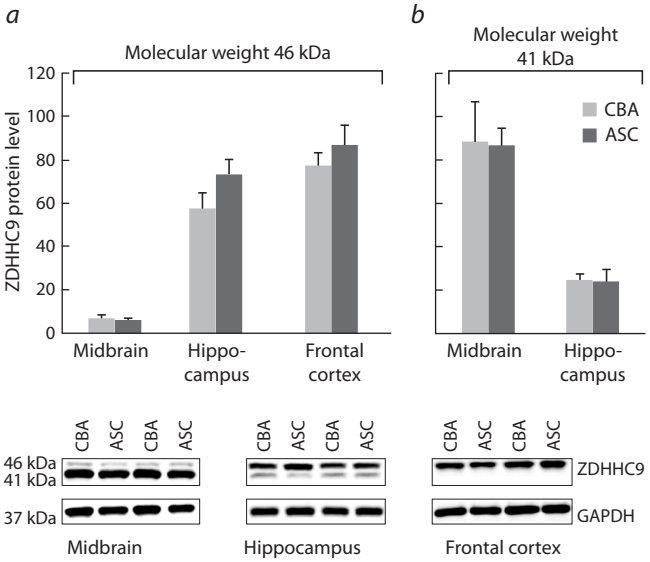



**Fig. 4.** The levels of ZDHHC9 protein in brain divisions of CBA and ASC mice.

(*a*) Quantitative evaluation of the intensity of the chemiluminescent signal at 26 kDa. The ZDHHC9 protein level is presented in relative units normalized to the corresponding level of GAPDH; (*b*) Quantitative evaluation of the intensity of the chemiluminescent signal at 41 kDa in the midbrain and in hippocampus of CBA and ASC mice. The ZDHHC5 protein level is presented in relative units normalized to the corresponding level of GAPDH. (*c*) Western blotting on the membrane. The numbers of samples in groups are $N \geq 9$.

шей работе пальмитилтрансфераз таких данных нет. Тем не менее, основываясь на наблюдаемых нами регион-специфических различиях массы исследуемых белков в сравнении с заявленными производителем антител, можно предположить, что пальмитилтрансферазы ZDHHC21 и ZDHHC9, как и пальмитилтрансфераза ZDHHC5, могут подвергаться посттрансляционным модификациям. Нельзя также исключать и возможность существования регион-специфических вариантов исследуемых ферментов, однако эта гипотеза требует дальнейших исследований с использованием целевого белка или нокаутных животных.

При иммунодетекции белка ZDHHC5 в среднем мозге мышей обеих линий обнаружено две полосы массой 75 кДа (ожидаемая масса) и 55 кДа, тогда как в коре и гиппокампе вторая полоса не обнаружена. В работе по изучению этой пальмитилтрансферазы в среднем мозге мышей линии 129/Ola (Li et al., 2010), как и в исследованиях, выполненных на культурах клеток (Brigidi et al., 2015; Badawy et al., 2017), полосу ZDHHC5 весом 55 кДа не обнаружили.

Однако аналогичная выявленной для белка ZDHHC5 картина наблюдалась и при визуализации пальмитилтрансферазы ZDHHC9 в среднем мозге и гиппокампе. При иммунодетекции белка ZDHHC9 в этих структурах мозга также наблюдалось по две полосы. В среднем мозге полоса массой 41 кДа намного ярче, чем полоса с заявленным производителем антител массой 46 кДа. В гиппокампе нижняя полоса, наоборот, менее интенсивна. Интересно, что для этой пальмитилтрансферазы расчетная масса составляет именно 41 кДа (http://www.uniprot.org/uniprot/P59268), тогда как ожидаемая, согласно описанию производителя антител, – 46 кДа. Следовательно, можно предположить, что более тяжелая полоса – это посттрансляционно модифицированная форма пальмитилтрансферазы ZDHHC9.

Нужно отметить крайне интересный факт, заключающийся в отсутствии во фронтальной коре мышей обеих линий альтернативных бэндов, исследуемых пальмитилтрансфераз, кроме ожидаемых. В первую очередь, это указывает на особенность этой структуры мозга для функции исследованных пальмитилтрансфераз. Кроме того, принимая во внимание выявленный факт, можно сделать вывод, что наблюдаемые нами в гиппокампе и среднем мозге дополнительные полосы для всех исследуемых пальмитилтрансфераз не являются артефактом иммунодетекции.

Таким образом, можно предположить, что наблюдаемая нами при иммунодетекции пальмитилтрансферазы ZDHHC5 легкая дополнительная полоса – изоформа фермента, экспрессирующаяся только в определенных структурах мозга. Для ZDHHC9 более вероятно, что дополнительная легкая полоса может представлять собой посттрансляционно немодифицированную форму фермента. В то же время тяжелая дополнительная полоса для пальмитилтрансферазы ZDHHC21 может быть как регион-специфической изоформой фермента, так и посттрансляционно модифицированным ферментом. Однако для подтверждения или опровержения таких предположений требуются дальнейшие исследования.

Следует отметить, что существенных различий по экспрессии исследованных пальмитилтрансфераз между мышами линии ASC с генетической предрасположенностью к депрессивно-подобному поведению и мышами родительской линии CBA обнаружено не было.

## Заключение

Таким образом, полученные в настоящей работе данные указывают на то, что генетическая предрасположенность к депрессивно-подобному поведению у мышей линии ASC не связана с экспрессией пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21 во фронтальной коре, гиппокампе и среднем мозге. В то же время нами впервые были выявлены регион-специфические особенности экспрессии пальмитилтрансфераз ZDHHC5, ZDHHC9 и ZDHHC21. Результаты позволяют предположить существование нескольких изоформ или посттрансляционных модификаций исследованных ферментов, дифференциально экспрессирующихся в различных структурах мозга, однако эта гипотеза требует дальнейших исследований.

## Acknowledgments

This work was supported by the Russian Foundation for Basic Research, project 16-54-12016. Animal propagation and maintenance were supported by State Budgeted Project 0324-2018-0016. The authors are grateful to the staffs of the Shared Access Center for Genetic Resources of Laboratory Animals (ICG SB RAS) (supported by project RFMEFI62117X0015) and the Vivarium for Conventional Animals (ICG SB RAS) for providing access to the equipment .

## Conflict of interest

The authors declare no conflict of interest..

## References

Alperina E.L., Kulikov A.V., Popova N.K., Idova G.V. Immnune response in mice of a new strain ASC (antidepressant-sensitive catalepsy). Byulleten Eksperimentalnoy Biologii i Meditsiny = Bulletin of Experimental Biology and Medicine. 2007; 144(8):188-190. (in Russian)

Badawy S.M.M., Okada T., Kajimoto T., Ijuin T., Nakamura S.I. DHHC5-mediated palmitoylation of S1P receptor subtype 1 determines G-protein coupling. Sci. Rep. 2017;7(1):16552. DOI 10.1038/s41598-017-16457-4.

Bazovkina D.V., Kulikov A.V., Kondaurova E.M., Popova N.K. Selection for the predisposition to catalepsy enhances depressive-like traits in mice. Genetika = Genetics (Moscow). 2005; 41(9):1222-1228. (in Russian)

Beard R.S.Jr., Yang X., Meegan J.E., Overstreet J.W., Yang C.G., Elliott J.A., Reynolds J.J., Cha B.J., Pivetti C.D., Mitchell D.A., Wu M.H., Deschenes R.J., Yuan S.Y. Palmitoyl acyltransferase DHHC21 mediates endothelial dysfunction in systemic inflammatory response syndrome. Nat. Commun. 2016;7(12823). DOI 10.1038/ncomms12823.

Brigidi G.S., Santyr B., Shimell J., Jovellar B., Bamji S.X. Activity-regulated trafficking of the palmitoyl-acyl transferase DHHC5. Nat. Commun. 2015;6(8200). DOI 10.1038/ncomms9200.

Cho E., Park M. Palmitoylation in Alzheimer's disease and other neurodegenerative diseases. Pharmacol. Res. 2016;111(133-151). DOI 10.1016/j.phrs.2016.06.008.

Duman R.S., Heninger G.R., Nestler E.J. A molecular and cellular theory of depression. Arch. Gen. Psychiatry. 1997;54(7):597-606.

Dutta D., Mandal C., Mandal C. Unusual glycosylation of proteins: Beyond the universal sequon and other amino acids. Biochim. Biophys. Acta. 2017;1861(12):3096-3108. DOI 10.1016/j.bbagen.2017.08.025.

Fukata Y., Fukata M. Protein palmitoylation in neuronal development and synaptic plasticity. Nat. Rev. Neurosci. 2010;11(3):161-175. DOI 10.1038/nrn2788.

Gorinski N., Ponimaskin E. Palmitoylation of serotonin receptors. Biochem. Soc. Trans. 2013;41(1):89-94. DOI 10.1042/BST20120235.

Harro J., Oreland L. Depression as a spreading neuronal adjustment disorder. Eur. Neuropsychopharmacol. 1996;6(3):207-223.

Jacobs B.L., Azmitia E.C. Structure and function of the brain serotonin system. Physiol. Rev. 1992;72(1):165-229.

Katkova L.E., Solenov E.I., Ivanova L.N. The role of protein kinase C in the formation of the mechanism of vasopressin antidiuretic action in the rat kidney during mammalian postnatal development. Ontogenez = Ontogenesis (Moscow). 2009;40(6):442-448. (in Russian)

Kondaurova E.M., Bazovkina D.V., Kulikov A.V., Popova N.K. Selective breeding for catalepsy changes the distribution of microsatellite D13Mit76 alleles linked to the 5-HT serotonin receptor gene in mice. Genes Brain Behav. 2006;5(8):596-601. DOI GBB212.

Kulikov A.V., Naumenko V.S., Voronova I.P., Tikhonova M.A., Popova N.K. Quantitative RT-PCR assay of 5-HT1A and 5-HT2A serotonin receptor mRNAs using genomic DNA as an external standard. J. Neurosci. Meth. 2005;141(1):97-101. DOI S016502700400216X.

Li Y., Hu J., Hofer K., Wong A.M., Cooper J.D., Birnbaum S.G., Hammer R.E., Hofmann S.L. DHHC5 interacts with PDZ domain 3 of post-synaptic density-95 (PSD-95) protein and plays a role in learning and memory. J. Biol. Chem. 2010;285(17):13022-13031. DOI 10.1074/jbc.M109.079426.

Maes M., Meltzer H.Y. The serotonin hypothesis of major depression. Psychopharmacology: The Fourth Generation of Progress. Eds. E.E. Bloom, N.N. Kupfer. New York, 1995;933-944.

Naumenko V.S., Kondaurova E.M., Bazovkina D.V., Tsybko A.S., Tikhonova M.A., Kulikov A.V., Popova N.K. Effect of brain-derived neurotrophic factor on behavior and key members of the brain serotonin system in genetically predisposed to behavioral disorders mouse strains. Neuroscience. 2012;214:59-67. DOI S0306-4522(12)00390-9.

Naumenko V.S., Kulikov A.V. Quantitative assay of 5-HT1A receptor gene expression in the brain. Molekulyarnaya Biologiya = Molecular Biology (Moscow). 2006;40(1):37-44. (in Russian)

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. Meth. 2008;170(2):197-203. DOI S0165-0270(08)00044-7.

Neumeister A., Wood S., Bonne O., Nugent A.C., Luckenbaugh D.A., Young T., Bain E.E., Charney D.S., Drevets W.C. Reduced hippocampal volume in unmedicated, remitted patients with major depression versus control subjects. Biol. Psychiatry. 2005;57(8):935-937. DOI 10.1016/j.biopsych.2005.01.016.

Papoucheva E., Dumuis A., Sebben M., Richter D.W., Ponimaskin E.G. The 5-hydroxytryptamine(1A) receptor is stably palmitoylated, and acylation is critical for communication of receptor with Gi protein. J. Biol. Chem. 2004;279(5):3280-3291. DOI 10.1074/jbc.M308177200.

Renner U., Glebov K., Lang T., Papusheva E., Balakrishnan S., Keller B., Richter D.W., Jahn R., Ponimaskin E. Localization of the mouse 5-hydroxytryptamine(1A) receptor in lipid microdomains depends on its palmitoylation and is involved in receptor-mediated signaling. Mol. Pharmacol. 2007;72(3):502-513. DOI 10.1124/mol.107.037085.

Ressler K.J., Mayberg H.S. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. Nat. Neurosci. 2007;10(9):1116-1124. DOI 10.1038/nn1944.

Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology. 1993;3(2):97-130.

Вавиловский журнал генетики и селекции. 2018;22(5):600-605
DOI 10.18699/VJ18.400

# Влияние однократного введения стрептозотоцина на метаболиты гиппокампа мышей линии NODSCID

Д.А. Тур, О.Б. Шевелев, М.Б. Шарапова, М.А. Золотых, А.Е. Акулов ✉

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

Значимое увеличение за последние годы числа людей с установленным диагнозом «сахарный диабет» выводит исследования, посвященные этой проблеме, в число наиболее актуальных. Продолжительная гипергликемия, сопровождающая развитие и течение сахарного диабета 1-го типа (СД1), может отразиться на функциональном и структурном уровне организации работы головного мозга. В основе подобных реакций может лежать изменение метаболизма. Общепринятым методом прижизненного выявления метаболических реакций в организме служит магнитно-резонансная спектроскопия (МРС). В настоящей работе для оценки влияния стрептозотоцина (СТЗ) и хронической гипергликемии, обусловленной отсроченным эффектом СТЗ, реализованным через гибель β-клеток поджелудочной железы, проведена МРС гиппокампа мышей линии NOD.CB17-Prkdc$^{scid}$/NcrCrl (NODSCID) через 4 и 60 дней после введения СТЗ. Модель СД1 с введением СТЗ – самая распространенная в мировой практике. Вместе с тем остается открытым вопрос – существует ли краткосрочный эффект введения СТЗ на уровень детектируемых с помощью МРС метаболитов гиппокампа животных. В результате сравнения опытной группы животных с контролем выявлено отсутствие влияния СТЗ на метаболиты гиппокампа мышей NODSCID на 4-й день после его введения. Однако в другом сравнении животных опыта и контроля через 60 дней после введения СТЗ отмечаются увеличение содержания аланина и таурина и снижение содержания лактата. Таким образом, введение самого СТЗ не сказывается на метаболизме гиппокампа. Использование МРС является перспективным методом для оценки влияния СД1 на метаболизм головного мозга животных.

Ключевые слова: сахарный диабет 1-го типа; мыши линии NOD.CB17-Prkdc$^{scid}$/NcrCrl; магнитно-резонансная спектроскопия; стрептозотоцин; гиппокамп.

## The effect of a single administration of streptozotocin on hippocampus metabolites in NODSCID mice

D.A. Tur, O.B. Shevelev, M.B. Sharapova, M.A. Zolotykh, A.E. Akulov ✉

Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

The significant increase in the number of people diagnosed with diabetes mellitus in recent years makes studies of this problem topical. The persistent hyperglycemia accompanying the development and course of type 1 diabetes mellitus (T1DM) can affect the functional and structural levels of the organization of the central nervous system. These changes may be mediated by metabolic aberrations. Magnetic resonance spectroscopy (MRS) is a common method of intravital detection of metabolic reactions. In this study, MRS of the hippocampus of NOD.CB17-Prkdc$^{scid}$/NcrCrl mice (NODSCID) was performed 4 days after the administration of streptozotocin (STZ) to assess the effect of STZ itself, and 60 days after the administration of STZ to another group of animals to assess the effect of chronic hyperglycemia caused by the delayed effect of STZ, involving the death of pancreatic β-cells. The simulation of T1DM by STZ administration is used worldwide. Nevertheless, the question remains whether there is a short-term effect of the introduction of STZ at the level of hippocampal metabolites recorded by MRS. The comparison of experimental and control animal groups revealed no effect of STZ on metabolites in the hippocampus of NODSCID mice on day 4 after its administration. In contrast, another comparison of the experimental and control animals on day 60 after STZ administration showed elevated contents of alanine and taurine, and a reduced lactate content. Thus, the introduction of STZ itself does not affect the metabolism of the hippocampus, and MRS is a promising method for assessing the effect of T1DM on brain metabolism in animals.

Key words: type 1 diabetes mellitus; NOD.CB17-Prkdc$^{scid}$/NcrCrl mice; magnetic resonance spectroscopy; streptozotocin; hippocampus.

**КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:**
Тур Д.А., Шевелев О.Б., Шарапова М.Б., Золотых М.А., Акулов А.Е. Влияние однократного введения стрептозотоцина на метаболиты гиппокампа мышей линии NODSCID. Вавиловский журнал генетики и селекции. 2018;22(5):600-605. DOI 10.18699/VJ18.400

**HOW TO CITE THIS ARTICLE:**
Tur D.A., Shevelev O.B., Sharapova M.B., Zolotykh M.A., Akulov A.E. The effect of a single administration of streptozotocin on hippocampus metabolites in NODSCID mice. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):600-605. DOI 10.18699/VJ18.400 (in Russian)

Received 16.04.2018
Accepted for publication 10.05.2018


✉ e-mail: akulov_ae@ngs.ru

Сахарный диабет 1-го типа (СД1) – распространенное хроническое заболевание, в основе которого лежит нарушение секреции инсулина и обмена веществ. СД1, особенно на фоне слабого гликемического контроля, неблагоприятно влияет на мозг (van Harten et al., 2006), может приводить к многочисленным осложнениям, в частности к развитию диабетической энцефалопатии. Ее признаки могут проявляться в атрофии белого и/или серого вещества как всего головного мозга, так и отдельных его структур. При этом структурные и метаболические отклонения зачастую носят сопряженный характер.

Клинические и экспериментальные исследования показали, что диабетическая энцефалопатия связана с нарушениями мозгового метаболизма (Northam et al., 2009), а протонная магнитно-резонансная спектроскопия (МРС) успешно используется для оценки подобных нарушений (Sarac et al., 2005; Heikkila et al., 2009; Mangia et al., 2013). Большинство исследований *in vivo* нарушений мозгового метаболизма на животных моделях сосредоточено на МРС гиппокампа (Biessels et al., 2001; Duarte et al., 2009; Wang et al., 2012; Moshkin et al., 2014), что обусловлено особенностями его организации и повышенной реакцией на воздействия (Revsin et al., 2009). И несмотря на то, что МРС головного мозга обладает относительно слабой селективностью, определяемые в ходе исследования метаболиты регистрируются как в нейронах и клетках глии, так и в межклеточном пространстве, результаты исследований актуальны для решения многих практических задач. С этой точки зрения полученная в ходе МРС информация об уровне метаболитов, участвующих в разных процессах функционирования головного мозга, позволяет качественно оценить изменение метаболизма при СД1 или влиянии веществ, моделирующих его.

Как известно, основной характеристикой СД1 является разрушение β-клеток поджелудочной железы с последующей недостаточной продукцией инсулина. В моделях на животных такое состояние достигается посредством введения ряда химических агентов, в частности стрептозотоцина (СТЗ). Химически индуцированная модель диабета относительно проста – диабет обычно индуцируют за пять-семь дней (King, 2012), и дешева (Dufrane et al., 2006) – СТЗ синтезируется *Streptomycetes achromogenes*. После введения СТЗ в организм он с помощью белка-переносчика глюкозы GLUT2 накапливается в β-клетках поджелудочной железы и вызывает алкилирование ДНК (Szkudelski et al., 2001). Происходит истощение NAD+ и снижение клеточного АТФ (Sandler, Swenne, 1983). Кроме того, СТЗ провоцирует появление свободных радикалов, которые способствуют повреждению ДНК и гибели β-клеток (Lenzen et al., 2008). Селективность СТЗ к клеткам поджелудочной железы не является абсолютной, и хотя СТЗ не способен проникать через гематоэнцефалический барьер, его высокие дозы могут привести к повышенной проницаемости барьера (Huber et al., 2006). В связи с этим в исследованиях, посвященных оценке метаболизма тканей головного мозга, представляется необходимым проверка краткосрочных эффектов СТЗ на метаболиты головного мозга, особенно структур, имеющих наиболее уязвимый гематоэнцефалический барьер. Наряду с этим некоторые линии мышей имеют различную чувствительность к СТЗ, например, DBA/2 более чувствительны, чем C57BL/6, а последние – восприимчивее линии BALB/c (Gurley et al., 2006). Использование конкретных линий мышей целесообразно при исследованиях осложнений, вызванных СД1 и ориентированных на изучение метаболических основ формирования диабетической энцефалопатии. В работе (Schmidt et al., 2003) выявлено, что для изучения нейропатий подходящей моделью могут стать мыши линии NODSCID (non-obese diabetes severe combined immunodeficiency). Введение СТЗ мышам NODSCID приводит к устойчивой гипергликемии, которая провоцирует дистрофические изменения в нейронах, что может быть следствием генетического дефекта этой линии, так как животные имеют пониженную способность к восстановлению разрывов двуцепочечных ДНК. В то же время остается открытым вопрос об изменениях концентрации нейрометаболитов в ранние и поздние сроки после введения СТЗ, т. е. при эугликемии и устойчивой гипергликемии.

В настоящей работе у мышей линии NODSCID посредством однократной прижизненной МРС исследованы метаболические изменения в гиппокампе на разных сроках после введения СТЗ.

## Материалы и методы

**Объект исследования и условия содержания.** В качестве объекта исследования использовали самцов мышей линии NOD.CB17-Prkdc$^{scid}$/NcrCrl (NODSCID). Возраст животных в начале эксперимента был равен восьми неделям. Животные были выращены в SPF-виварии Центра генетических ресурсов лабораторных животных Института цитологии и генетики СО РАН.

На протяжении всего исследования животных содержали в индивидуально-вентилируемых клетках (OptiMice, США) группами по два животных при следующих условиях: освещение: 14С:10Т; температура 22–24 °С; относительная влажность 40–50 %; режим питания и потребления воды *ad libitum*. В качестве питания использовали стандартный автоклавированный корм ssniff® R/M-H autoclavable V1534-3 для содержания грызунов (Sniff, Германия), питья – очищенную стерильную воду с добавлением необходимого количества минеральных солей (минеральная добавка для воды «Северянка», Россия).

**Схема исследования и манипуляции с животными.** Проведена МРС гиппокампа одной группы животных через четыре дня после введения СТЗ – для оценки влияния самого СТЗ и другой группы через 60 дней после введения СТЗ – для оценки влияния хронической гипергликемии, обусловленной введением СТЗ и гибелью β-клеток поджелудочной железы.

В первой группе были контрольные и опытные животные по шесть особей. Перед исследованием у всех животных определяли массу тела и, исходя из нее, контрольным особям внутрибрюшинно вводили физиологический раствор, а опытным – СТЗ (Sigma, St. Louis, США) в дозе 150 мг/кг. Через четыре дня после введения всех животных подвергали процедуре МРС-сканирования.

Во второй группе были также контрольные и опытные животные по восемь особей. На животных проводили такие же манипуляции и тесты, как и в первой группе,

только МРС была проведена через 60 дней после введения СТЗ.

У всех животных измеряли концентрацию глюкозы в крови посредством электрохимического метода при помощи глюкометра Diacont (ООО «Диаконт», Москва, Россия) и индивидуальных тест-полосок этой же фирмы. Измерения проводили у всех животных до начала эксперимента, затем на 4-й день для первой группы животных и на 60-е сутки после введения СТЗ для второй группы. Забор крови осуществляли в искусственно созданный период отсутствия потребления корма из кончика хвоста животного, таким образом, она была смешанного типа, взятая натощак. По окончании МРС животных выводили из эксперимента методом цервикальной дислокации.

Все экспериментальные процедуры выполнены в соответствии с Директивой 2010/63/EU Европейского парламента и совета Европейского Союза от 22 сентября 2010 года по охране животных, используемых в научных целях, и одобрены Комиссией по биоэтике ИЦиГ СО РАН.

**Магнитно-резонансная спектроскопия.** Исследования проводили на горизонтальном томографе с напряженностью магнитного поля 11.7 Тесла (Bruker, BioSpec 117/16 USR, Германия). Животных наркотизировали с помощью изофлюрана в смеси кислорода (1.5 %, скорость потока 300 мл/мин). Температуру тела поддерживали теплой водой (36 °C), циркулирующей через настольный лоток томографического сканера. Для контроля глубины анестезии использовали пневматический датчик дыхания (SA Instruments, Stony Brook, NY, США). Все изображения мозга и спектры были получены с использованием объемной $^{1}$H радиочастотной катушки (500.3 МГц, диаметром 23 мм). Для правильного позиционирования спектроскопических вокселей, размер которых составлял 1.3 × 2.5 × 2.5 мм – 8.1 мм$^3$, были получены $T_2$-взвешенные изображения (импульсная последовательность TurboRARE с параметрами: TE = 8 мс; TEeff = 24 мс; TR = 2500 мс; RARE factor = 8, сканирование проведено в двух проекциях: аксиальной и сагиттальной, параметры изображений: толщина среза = 0.5 мм; расстояние между центрами срезов = 0.9 мм; поле зрения = 2 × 2 см; матрица = 256 × 256 пикселей; число усреднений = 5. Продолжительность сканирования для каждой проекции составила 6 мин 40 с. Размер вокселя позволил получить спектры с относительно высоким разрешением за приемлемое время регистрации – 10 мин. Все протонные спектры получены с помощью одновоксельной спектроскопии методом STEAM (Stimulated Echo Acquisition Mode Spectroscopy) с TE = 3 мс, TM = 20 мс и TR = 5000 мс, спектральной шириной 4000 Гц, числом точек 2048. Перед каждым спектроскопическим измерением сигнал воды подавляли с помощью переменной мощности импульса и оптимизированной задержки релаксации – VAPOR (Gruetter, 1993); проводили также настройку однородности магнитного поля в пределах выбранного вокселя с использованием методики FastMap (Bruker).

Спектры были обработаны специализированной программой, разработанной сотрудниками ИЦиГ СО РАН (Moshkin et al., 2014), по принципам аналогичной программному обеспечению LCModel (Provencher, 1993). Полученные спектры содержали пики следующих соединений: мио-инозитол, глицин, креатин, глутамат/глутамин, таурин, холиновые компоненты, аспартат, N-ацетиласпартат, гамма-аминомасляная кислота (ГАМК), лактат, аланин, фосфорилэтаноламин.

**Статистическая обработка данных.** Все значения исследуемых параметров представлены в виде средней и стандартной ошибки (M ± SE). Нормальность распределения параметров проверяли критерием Колмогорова–Смирнова. Достоверность различий между группами: 1) контроль – опыт на 4-е сутки после введения СТЗ и 2) контроль – опыт на 60-е сутки после введения СТЗ, ввиду отсутствия множественного сравнения средних, оценивали с помощью *t*-критерия Стьюдента. Расчеты были выполнены с использованием программного пакета STATISTICA 6.0.

## Результаты

В ходе проведения МРС получены данные по 12 метаболитам головного мозга. У животных на 4-е сутки после введения СТЗ уровни исследованных метаболитов достоверно не отличались от значений, зарегистрированных у контрольных особей (таблица). В начале исследования, перед введением СТЗ, концентрация глюкозы в крови у мышей в контроле составила 5.7 ± 0.2 ммоль/л, в опыте – 5.6 ± 0.2 ммоль/л. На 4-е сутки эти показатели были следующие: 6.0 ± 0.9 ммоль/л в контроле и 5.8 ± 0.9 ммоль/л в опыте.

Во второй группе животных у самцов опыта в сравнении с контролем обнаружены достоверные отличия: более высокие значения аланина и таурина и пониженное значение лактата (см. таблицу). У этих животных так же в начале исследования, перед введением СТЗ, произведены замеры концентрации глюкозы в крови: у мышей в контроле она составила 5.4 ± 0.3 ммоль/л, в опыте 5.6 ± 0.2 ммоль/л. На 60-е сутки эти показатели составляли: 6.8 ± 0.2 ммоль/л в контроле и 22.5 ± 2.0 ммоль/л в опыте.

## Обсуждение

Полученные в результате исследования данные указывают на отсутствие эффектов СТЗ на метаболиты гиппокампа на 4-й день после введения, но наблюдаются эффекты на фоне развития гипергликемии. Так, в первой группе животных, в которой оценивалось непосредственное влияние СТЗ на метаболиты гиппокампа, не выявлено ни одного достоверного отличия в уровнях метаболитов между контрольными и опытными особями. Известно, что введение СТЗ в мировой практике – наиболее простой и удобный подход для получения устойчивой гипергликемии (King, 2012) и целесообразный для тестирования лекарств, методов лечения, а также работ, связанных с диабетическими осложнениями, включая диабетическую энцефалопатию (Jederstrom et al., 2005; Sheshala et al., 2009). Устойчиво гипергликемия формируется с 5–7-го дня после введения СТЗ. Несмотря на это, остается необходимость контроля возможного эффекта СТЗ на метаболизм головного мозга еще до наступления гипергликемии. Потенциально к таким эффектам может привести краткосрочное формирование гипогликемии в начальной стадии моделирования СД1 или неполное аккумулирование СТЗ в панкреатических β-клетках. Однако в сравнении с аллоксаном – другим

Levels of hippocampus metabolites in NOD SCID mice determined by magnetic resonance spectroscopy *in vivo*

| Metabolites | Day 4 after STZ | | | Day 60 after STZ | | |
|---|---|---|---|---|---|---|
| | Control | STZ | *p* | Control | STZ | *p* |
| N-acetylaspartate | 12.62 ± 0.29 | 13.11 ± 1.08 | 0.69 | 13.46 ± 0.49 | 13.08 ± 0.45 | 0.57 |
| GABA | 5.11 ± 0.61 | 5.08 ± 0.62 | 0.96 | 7.45 ± 0.26 | 7.07 ± 0.55 | 0.52 |
| Alanine | 3.17 ± 1.07 | 2.08 ± 1.55 | 0.59 | 4.04 ± 0.70 | 7.12 ± 0.95 | 0.01 |
| Aspartate | 0.38 ± 0.30 | 0.30 ± 0.07 | 0.79 | 1.69 ± 0.62 | 0.95 ± 0.31 | 0.31 |
| Choline compounds | 1.72 ± 0.17 | 1.73 ± 0.16 | 0.96 | 1.50 ± 0.15 | 1.47 ± 0.32 | 0.93 |
| Creatine | 10.73 ± 0.78 | 10.99 ± 2.39 | 0.92 | 12.03 ± 0.82 | 12.60 ± 1.17 | 0.69 |
| Glutamate/Glutamine | 11.45 ± 0.62 | 11.64 ± 1.51 | 0.91 | 8.69 ± 0.90 | 8.17 ± 0.66 | 0.64 |
| myo-Inositol | 1.54 ± 0.09 | 6.03 ± 4.61 | 0.38 | 6.71 ± 2.07 | 7.77 ± 2.09 | 0.72 |
| Taurine | 7.44 ± 0.35 | 7.37 ± 0.51 | 0.91 | 7.96 ± 0.58 | 10.25 ± 0.86 | 0.03 |
| Glycine | 18.70 ± 1.07 | 17.17 ± 6.65 | 0.83 | 13.92 ± 3.31 | 15.20 ± 3.14 | 0.78 |
| Lactate | 5.55 ± 1.04 | 3.56 ± 0.98 | 0.23 | 9.33 ± 1.37 | 4.23 ± 0.97 | < 0.01 |
| Phosphorylethanolamine | 10.11 ± 1.94 | 9.28 ± 2.46 | 0.80 | 13.18 ± 1.46 | 12.07 ± 1.85 | 0.63 |

распространенным агентом для химически-индуцированного СД1 (у которого в фазу I уже в первые минуты после введения наблюдается гипогликемическая реакция как результат кратковременной повышенной стимуляции секреции инсулина) – при введении СТЗ эта гипогликемическая реакция не наблюдается, так как СТЗ не ингибирует глюкокиназу. Кроме того, СТЗ не имеет непосредственного прямого ингибирующего воздействия на транспорт глюкозы (Elsner et al., 2000) или фосфорилирование глюкокиназы (Lenzen et al., 1987). Вместе с тем СТЗ аналогичен по своему строению с глюкозой и переносится в клетку белком транспорта глюкозы GLUT2, но не распознается другими переносчиками глюкозы. Это объясняет его относительную селективную токсичность для β-клеток, поскольку они имеют высокие уровни GLUT2 (Schnedl et al., 1994). Поэтому наличие переносчика глюкозы GLUT2 в других органах приобретает особую значимость. Показано, что СТЗ потенциально способен приводить к повреждению почек или печени (Weiss, 1982; Qinna, Badwan, 2015), где также содержится GLUT2. Однако обеспечение GLUT2 увеличения пассивного транспорта СТЗ в этих органах ограничено из-за их низкой аффинности к глюкозе. Важно отметить, что головной мозг не имеет высокого уровня содержания GLUT2.

Результаты во второй группе животных во многом определены СД1 и формированием хронической гипергликемии, поскольку введение СТЗ со временем приводит как к необратимому некрозу β-клеток, так и к развитию периферической резистентности к инсулину. Метаболизм головного мозга при СД1 претерпевает серьезные изменения, которые наглядно демонстрируют данные МРС. По результатам нашей работы, у мышей опыта через 60 дней после введения СТЗ наблюдается достоверное повышение уровня таурина. Считается, что таурин вместе с креатином и мио-инозитолом участвует в осмотической регуляции мозга (Lien et al., 1990, 1991). Увеличение уровня таурина отмечается также в головном мозге крыс с СТЗ-индуцированным сахарным диабетом (Rose et al., 2000). Подобная картина наблюдается и в клинической практике у пациентов с СД1 (Kreis, Ross, 1992; Geissler et al., 2003). Помимо этого, таурин, являясь одной из наиболее распространенных свободных аминокислот в центральной нервной системе, способствует улучшению энергетических процессов, стимулируя регенерацию при дистрофических заболеваниях и процессах, сопровождающихся значительным нарушением метаболизма (Timbrell et al., 1995; Hussy et al., 2000; Tanabe et al., 2010). Некоторые исследования показывают, что таурин может предотвратить или обратить церебральные и нейронные дисфункции, вызванные гипергликемией (Obrosova et al., 2001; Terada et al., 2011; Ito et al., 2012).

У животных после 60 дней с момента введения СТЗ происходит также снижение уровня лактата по отношению к контрольным особям. В литературе существуют противоречивые данные по уровню лактата на фоне СД1 и хронической гипергликемии. Некоторые исследователи отмечают повышение уровня лактата вплоть до возникновения гиперлактацидемической комы (Salceda et al., 1998). В других работах, наоборот, показано снижение уровня лактата (Lapidot, Haber, 2001; Wang et al., 2012). В нашем исследовании на фоне продолжительной гипергликемии мы наблюдаем, что аланин и лактат находятся в обратной связи.

По результатам полученных данных, можно отметить отсутствие влияния введения СТЗ на уровень метаболитов гиппокампа исследованных животных на 4-й день после введения. Наблюдается изменение в уровне метаболитов гиппокампа при продолжительном влиянии СТЗ, что реализуется не за счет самого СТЗ, но за счет токсичности СТЗ по отношению к β-клеткам поджелудочной железы, их гибели и формирования хронической гипергликемии. Таким образом, в модели химически-индуцированного сахарного диабета 1-го типа введение самого СТЗ не приводит к значительным изменениям уровня метаболитов

гиппокампа, а наблюдаемые на 60-й день после введения СТЗ изменения сопряжены с развитием хронической гипергликемии.

## Acknowledgments

This study was supported by the Russian Science Foundation, project 17-75-10029. Use of the equipment of the Shared Access Center for Genetic Resources of Laboratory Animals (ICG SB RAS) was supported by project RFMEFI62117X0015.

## Conflict of interest

The authors declare no conflict of interest.

## References

Biessels G.J., Braun K.P., de Graaf R.A., van Eijsden P., Gispen W.H., Nicolay K. Cerebral metabolism in streptozotocin-diabetic rats: an *in vivo* magnetic resonance spectroscopy study. Diabetologia. 2001; 44:346-353. DOI 10.1007/s001250051625.

Duarte J.M., Carvalho R.A., Cunha R.A., Gruetter R. Caffeine consumption attenuates neurochemical modifications in the hippocampus of streptozotocin-induced diabetic rats. J. Neurochem. 2009;111: 368-379. DOI 10.1111/j.1471-4159.2009.06349.x.

Dufrane D., van Steenberghe M., Guiot Y., Goebbels R.M., Saliez A., Gianello P. Streptozotocin-induced diabetes in large animals (pigs/primates): role of GLUT2 transporter and beta-cell plasticity. Transplantation. 2006;15;81(1):36-45. DOI 10.1097/01.tp.0000189712.74495.82.

Elsner M., Guldbakke B., Tiedge M., Munday R., Lenzen S. Relative importance of transport and alkylation for pancreatic beta-cell toxicity of streptozotocin. Diabetologia. 2000;43:1528-1533. DOI 10.1007/s001250051564.

Geissler A., Frund R., Scholmerich J., Feuerbach S., Zietz B. Alterations of cerebral metabolism in patients with diabetes mellitus studied by proton magnetic resonance spectroscopy. Exp. Clin. Endocrinol. Diabetes. 2003;111(7):421-427. DOI 10.1055/s-2003-44289.

Gruetter R. Automatic, localized *in vivo* adjustment of all first-and second-order shim coils. Magn. Reson. Med. 1993;29:804-811. DOI 10.1002/mrm.1910290613.

Gurley S.B., Clare S.E., Snow K.P., Hu A., Meyer T.W., Coffman T.M. Impact of genetic background on nephropathy in diabetic mice. Am. J. Physiol. 2006;290(1):F214-F222. DOI 10.1152/ajprenal.00204.2005.

Heikkila O., Lundbom N., Timonen M., Groop P.H., Heikkinen S., Makimattila S. Hyperglycaemia is associated with changes in the regional concentrations of glucose and myo-inositol within the brain. Diabetologia. 2009;52:534-540. DOI 10.1007/s00125-008-1242-2.

Huber J.D., VanGilder R.L., Houser K.A. Streptozotocin-induced diabetes progressively increases blood-brain barrier permeability in specific brain regions in rats. Am. J. Physiol. Heart Circ. Physiol. 2006;291(6):H2660-H2668. DOI 10.1152/ajpheart.00489.2006.

Hussy N., Deleuze C., Desarménien M.G., Moos F.C. Osmotic regulation of neuronal activity: a new role for taurine and glial cells in a hypothalamic neuroendocrine structure. Prog. Neurobiol. 2000; 62(2):113-134. DOI 10.1016/S0301-0082(99)00071-4.

Ito T., Schaffer S.W., Azuma J. The potential usefulness of taurine on diabetes mellitus and its complications. Amino Acids. 2012;42(5): 1529-1539. DOI 10.1007/s00726-011-0883-5.

Jederstrom G., Grasjo J., Nordin A., Sjoholm I., Andersson A. Blood glucose-lowering activity of a hyaluronan-insulin complex after oral administration to rats with diabetes. Diabetes Technol. Ther. 2005; 7(6):948-957. DOI 10.1089/dia.2005.7.948.

King A.J. The use of animal models in diabetes research. Br. J. Pharmacol. 2012;166(3):877-894. DOI 10.1111/j.1476-5381.2012.01911.x.

Kreis R., Ross B.D. Cerebral metabolic disturbances in patients with subacute and chronic diabetes mellitus: detection with proton MR spectroscopy. Radiology. 1992;184(1):123-130. DOI 10.1148/radiology.184.1.1319074.

Lapidot A., Haber S. Effect of endogenous beta-hydroxybutyrate on glucose metabolism in the diabetic rabbit brain: A C-13-magnetic resonance spectroscopy study of [U-C-13] glucose metabolite. J. Neurosci. Res. 2001;64(2):207-216. DOI 10.1002/jnr.1067.

Lenzen S. The mechanisms of alloxan- and streptozotocin-induced diabetes. Diabetologia. 2008;51(2):216-226. DOI 10.1007/s00125-007-0886-7.

Lenzen S., Tiedge M., Panten U. Glucokinase in pancreatic B-cells and its inhibition by alloxan. Acta Endocrinol. 1987;115:21-29.

Lien Y.H., Shapiro J.I., Chan L. Effects of hypernatremia on organic brain osmoles. J. Clin. Invest. 1990;85(5):1427-1435. DOI 10.1172/JCI114587.

Lien Y.H., Shapiro J.I., Chan L. Study of brain electrolytes and organic osmolytes during correction of chronic hyponatremia. Implications for the pathogenesis of central pontine myelinolysis. J. Clin. Invest. 1991;88(1):303-309. DOI 10.1172/JCI115292.

Mangia S., Kumar A.F., Moheet A.A., Roberts R.J., Eberly L.E., Seaquist E.R., Tkáč I. Neurochemical profile of patients with type 1 diabetes measured by 1H-MRS at 4T. J. Cereb. Blood Flow Metab. 2013;33:754-759. DOI 10.1038/jcbfm.2013.13.

Moshkin M.P., Akulov A.E., Petrovski D.V., Saik O.V., Petrovskiy E.D., Savelov A.A., Koptyug I.V. Proton magnetic resonance spectroscopy of brain metabolic shifts induced by acute administration of 2-deoxy-D-glucose and lipopolysaccharides. NMR Biomed. 2014;27:399-405. DOI 10.1002/nbm.3074.

Northam E.A., Rankins D., Lin A., Wellard R.M., Pell G.S., Finch S.J., Werther G.A., Cameron F.J. Central nervous system function in youth with type 1 diabetes 12 years after disease onset. Diabetes Care. 2009;32:445-450. DOI 10.2337/dc08-1657.

Obrosova I.G., Fathallah L., Stevens M.J. Taurine counteracts oxidative stress and nerve growth factor deficit in early experimental diabetic neuropathy. Exp. Neurol. 2001;172(1):211-219. DOI 10.1006/exnr.2001.7789.

Provencher S.W. Estimation of metabolite concentrations from localized *in vivo* proton NMR spectra. Magn. Reson. Med. 1993;30(6):672-679. DOI 10.1002/mrm.1910300604.

Qinna N.A., Badwan A.A. Impact of streptozotocin on altering normal glucose homeostasis during insulin testing in diabetic rats compared to normoglycemic rats. Drug Des. Devel. Ther. 2015;9:2515-2525. DOI 10.2147/DDDT.S79885.

Revsin Y., Rekers N.V., Louwe M.C., Saravia F.E., De Nicola A.F., de Kloet E.R., Oitzl M.S. Glucocorticoid receptor blockade normalizes hippocampal alterations and cognitive impairment in streptozotocin-induced type 1 diabetes mice. Neuropsychopharmacology. 2009;34(3):747-758. DOI 10.1038/npp.2008.136.

Rose S.J., Bushi M., Nagra I., Davies W.E. Taurine fluxes in insulin dependent diabetes mellitus and rehydration in streptozotocin treated rats. Adv. Exp. Med. Biol. 2000;483:497-501. DOI 10.1007/0-306-46838-7_55.

Salceda R., Vilchis C., Coffe V., Hernandez-Munoz R. Changes in the redox state in the retina and brain during the onset of diabetes in rats. Neurochem. Res. 1998;23(6):893-897. DOI 10.1023/A:1022467230259.

Sandler S., Swenne I. Streptozotocin, but not alloxan, induces DNA repair synthesis in mouse pancreatic islets *in vitro*. Diabetologia. 1983; 25(5):444-447.

Sarac K., Akinci A., Alkan A., Aslan M., Baysal T., Ozcan C. Brain metabolite changes on proton magnetic resonance spectroscopy in children with poorly controlled type 1 diabetes mellitus. Neuroradiology. 2005;47:562-565. DOI 10.1007/s00234-005-1387-3.

Schmidt R.E., Dorsey D.A., Beaudet L.N., Frederick K.E., Parvin C.A., Plurad S.B., Levisetti M.G. Non-obese diabetic mice rapidly develop

dramatic sympathetic neuritic dystrophy a new experimental model of diabetic autonomic neuropathy. Am. J. Pathol. 2003;163(5):2077-2091. DOI 10.1016/S0002-9440(10)63565-1.

Schnedl W.J., Ferber S., Johnson J.H., Newgard C.B. STZ transport and cytotoxicity. Specific enhancement in GLUT2-expressing cells. Diabetes. 1994;43(11):1326-1333. DOI 10.2337/diab.43.11.1326.

Sheshala R., Peh K.K., Darwis Y. Preparation, characterization, and in vivo evaluation of insulin-loaded PLA-PEG microspheres for controlled parenteral drug delivery. Drug Dev. Ind. Pharm. 2009;35(11):1364-1374. DOI 10.3109/03639040902939213.

Szkudelski T. The mechanism of alloxan and streptozotocin action in B cells of the rat pancreas. Physiol. Res. 2001;50(6):537-546. DOI 10.1097/01.tp.0000189712.74495.82.

Tanabe M., Nitta A., Ono H. Neuroprotection via strychnine-sensitive glycine receptors during post-ischemic recovery of excitatory synaptic transmission in the hippocampus. J. Pharmacol. Sci. 2010;113(4):378-386. DOI 10.1254/jphs.10150FP.

Terada T., Hara K., Haranishi Y., Sata T. Antinociceptive effect of intrathecal administration of taurine in rat models of neuropathic pain. Can. J. Anaesth. 2011;58(7):630-637. DOI 10.1007/s12630-011-9504-8.

Timbrell J.A., Seabra V., Waterfield C.J. The *in vivo* and *in vitro* protective properties of taurine. Gen. Pharmacol. 1995;26(3):453-462. DOI 10.1016/0306-3623(94)00203-Y.

van Harten B., de Leeuw F.E., Weinstein H.C., Scheltens P., Biessels G.J. Brain imaging in patients with diabetes: a systematic review. Diabetes Care. 2006;29:2539-2548. DOI 10.2337/dc06-1637.

Wang W.T., Lee P., Yeh H.W., Smirnova I.V., Choi I.Y. Effects of acute and chronic hyperglycemia on the neurochemical profiles in the rat brain with streptozotocin-induced diabetes detected using *in vivo* 1H MR spectroscopy at 9.4T. J. Neurochem. 2012;121:407-417. DOI 10.1111/j.1471-4159.2012.07698.x.

Weiss R.B. Streptozocin: a review of its pharmacology, efficacy, and toxicity. Cancer Treat. Rep. 1982;66:427-438.

Вавиловский журнал генетики и селекции. 2018;22(5):606-610
DOI 10.18699/VJ18.401

# Регуляция ацетилирования гистона H4 в центральной нервной системе и командных нейронах оборонительного поведения моллюска *Helix* серотонином и нейропептидом FMRFамидом

Л.Н. Гринкевич ✉, Т.Г. Зачепило

Институт физиологии им. И.П. Павлова Российской академии наук, Санкт-Петербург, Россия

Вовлеченность эпигенетических механизмов в формирование долговременной памяти не вызывает сомнений. В настоящее время среди этих механизмов наиболее активно исследуются изменения уровня различных гистоновых модификаций (в первую очередь, ацетилирования и метилирования) в составе хроматина клеток центральной нервной системы (ЦНС) на различных экспериментальных моделях. Одна из наиболее удобных моделей – моллюски, их ЦНС относительно просто устроена и для ряда видов достаточно хорошо охарактеризована. В работе в качестве объекта исследования использована ЦНС виноградной улитки (*Helix lucorum*), для которой ранее была выявлена группа нейронов, участвующих в формировании различных типов поведения, включая сохраняющийся во времени ответ на различные стимулы. Целью работы было изучение влияния известных эффекторов: серотонина и FMRFамида, связанных в ЦНС с активаторными и тормозными путями соответственно, на ацетилирование гистона H4 в подглоточном комплексе ганглиев, а также в командных нейронах оборонительного поведения правого (ППа3/2) и левого (ЛПа3/2) париетальных ганглиев улитки. Исследование проводилось методом Вестерн-блот гибридизации. Полученные результаты указывают на сильную зависимость эффектов изучаемых нейромедиаторов от структур ЦНС, которые подвергались воздействию этих веществ. Так, оказалось, что в подглоточном комплексе ганглиев под действием серотонина происходило усиление суммарного ацетилирования гистона H4, а FMRFамид подавлял его эффект. В противоположность этому, в командных нейронах правого париетального ганглия серотонин и FMRFамид усиливали действие друг друга, что приводило к существенному повышению уровня ацетилирования гистона H4. Однако в симметричных нейронах левого париетального ганглия никаких изменений в уровне ацетилирования под действием обоих веществ не наблюдалось, что служит новым свидетельством наличия функциональной асимметрии у *Helix*. Результаты исследования позволяют сделать заключение о двоякой роли тормозных путей, опосредуемых FMRFамидом, в зависимости от контекста нейрональных комплексов, они могут как подавлять действие активаторных путей, что было зафиксировано нами в подглоточном комплексе ганглиев улитки, так и выступать в роли их синергистов, как в командных нейронах оборонительного поведения правого париетального ганглия.

Ключевые слова: моллюск *Helix*; иммуноблоттинг; эпигенетика; долговременная память; ацетилирование гистона H4; серотонин; FMRFамид; командные нейроны.

**КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:**

Гринкевич Л.Н., Зачепило Т.Г. Регуляция ацетилирования гистона H4 в центральной нервной системе и командных нейронах оборонительного поведения моллюска *Helix* серотонином и нейропептидом FMRFамидом. Вавиловский журнал генетики и селекции. 2018;22(5):606-610. DOI 10.18699/VJ18.401

**HOW TO CITE THIS ARTICLE:**

Grinkevich L.N., Zachepilo T.G. Regulation of histone H4 acetylation in the CNS and defensive behavior command neurons of the mollusk *Helix* mediated by serotonin and neuropeptide FMRFamide. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):606-610. DOI 10.18699/VJ18.401 (in Russian)

Received 07.03.2018
Accepted for publication 28.05.2018


✉ e-mail: Larisa_Gr_spb@mail.ru

# Regulation of histone H4 acetylation in the CNS and defensive behavior command neurons of the mollusk *Helix* mediated by serotonin and neuropeptide FMRFamide

L.N. Grinkevich ✉, T.G. Zachepilo

Pavlov Institute of Physiology, RAS, St. Petersburg, Russia

Epigenetic mechanisms are commonly known to underlie memory formation. Presently, scientists' attention is focused on changes in the levels of histone modifications (mainly acetylation and methylation) in the chromatin of CNS cells tested in various experimental models. Owing to their relatively simple CNSs, mollusks are among the most popular models. Our experiments were conducted with the mollusk *Helix lucorum* because its CNS had been investigated in detail and most of its neurons had been proven to participate in the formation of different behavior patterns, including the prolonged response to various stimuli. This work concerns the influence of various effectors (serotonin and FMRFamide, associated with CNS activator and inhibitory pathways, respectively) on the acetylation of H4 histone in the subesophageal ganglion complex and in defensive behavior command neurons of the right and left parietal ganglia (RPa3/2 and LPa3/2) in the snail. Western blot analysis showed that FMRFamide inhibited histone H4 acetylation induced by serotonin in the subesophageal complex of CNS ganglia. However, serotonin and FMRFamide cooperatively enhanced the induction of histone H4 acetylation in RPa3/2 defensive behavior command neurons. No changes were found in the counterpart LPa3/2. It is a new piece of evidence for functional asymmetry in *Helix*. The inhibitory pathways mediated by FMRFamide not only inhibit the activatory intracellular processes in the entire CNS but can also enhance them, as in RPa3/2 defensive behavior command neurons.

Key words: mollusk *Helix*, Western blotting, epigenetics, long-term memory, histone H4 acetylation, serotonin, FMRFamide, command neurons.

Изучение механизмов формирования долговременной памяти (ДП) – одна из актуальнейших задач современной нейробиологии. В последние годы наиболее активно развивающимся направлением исследований в этой области стало изучение изменений в эпигенетическом ландшафте хроматина в различных клетках и структурах центральной нервной системы (ЦНС). При этом особое внимание уделяется посттрансляционным модификациям гистонов, регулирующим активность промоторных и энхансерных районов генов и обеспечивающим долгосрочные изменения в уровне их экспрессии (Zovkic et al., 2013; Kim, Kaang, 2017). Ацетилирование и метилирование гистонов относятся к числу наиболее «популярных модификаций». Это связано как со значительным вкладом этих модификаций в определение уровня генной экспрессии, так и с существованием терапевтических способов корректировки статуса ацетилирования и метилирования гистонов, потенциально способных восстанавливать ментальные характеристики при ряде патологий (Abel, Zukin, 2008; Gräff, Tsai, 2013). Тем не менее следует отметить, что и механизмы регуляции процессов ацетилирования и метилирования гистонов, и способы воздействия на эффективность этих процессов изучены далеко недостаточно.

Моллюски – широко используемая экспериментальная модель в исследованиях молекулярных механизмов ДП (Balaban, 2002; Гринкевич, 2012; Kandel, 2012), в первую очередь, в связи с относительной простотой организации их ЦНС и наличием чрезвычайно удобных для исследования гигантских нейронов, позволяющих вести работу с отдельными нервными клетками. Именно на моллюске (*Aplysia*) впервые была показана ключевая роль ацетилирования гистонов в формировании ДП (Guan et al., 2002), что в дальнейшем было подтверждено и на моделях позвоночных животных (Levenson, Sweatt, 2006). Ранее нами была проведена серия исследований роли ацетилирования гистона H3 в процессе обучения на модели условного оборонительного рефлекса пищевой аверсии у моллюска *Helix*. Результаты этих исследований не только выявили повышение уровня ацетилирования гистона H3 в результате обучения, но и показали, что этот процесс запускается серотонином при участии киназного пути MAPK/ERK (Grinkevich et al., 2008; Danilova et al., 2010; Kharchenko et al., 2010; Danilova, Grinkevich, 2012; Гринкевич, Воробьева, 2014).

Однако хорошо известно, что наряду с активаторными процессами для формирования ДП необходимы также процессы торможения, осуществление которых у моллюсков связано с нейропептидом FMRFамидом (Phe-Met-Arg-Phe-NH2) (Guan et al., 2002). В недавнем исследовании влияния серотонина и FMRFамида на содержание активаторной (H3K4me3) и ингибиторной (H3K9me2) модификаций гистона H3 в ЦНС *Helix* (Гринкевич, Воробьева, 2016) мы попытались подойти поближе к выяснению механизма этого явления и обнаружили противоположные эффекты указанных соединений.

Задачей настоящей работы было изучение влияния этих эффекторов на ацетилирование гистона H4 в подглоточном комплексе ганглиев, а также командных нейронах оборонительного поведения правого (ППа3/2) и левого (ЛПа3/2) париетальных ганглиев улитки.

## Материалы и методы

Эксперименты проводили на взрослых виноградных улитках *Helix lucorum*. Выделение ЦНС, обработку серотонином и/или FMRF-амидом, выделение подглоточных комплексов ганглиев и отдельных командных нейронов осуществляли, как описано (Danilova et al., 2010; Гринкевич, Воробьева, 2016). Для каждого независимого эксперимента нейроны из трех животных объединяли (отдельно из левых и правых париетальных ганглиев) и получали клеточные экстракты, согласно (Monsey et al., 2011). В качестве контроля использовали ЦНС или командные нейроны из ЦНС, которую инкубировали в физиологическом растворе без добавления изучаемых препаратов. Разделение белков электрофорезом, Вестерн-блот анализ и обработку результатов также выполняли согласно описанным методам и протоколам (Гринкевич, Воробьева, 2016).

## Результаты

Изучение влияния серотонина и FMRFамида на ацетилирование гистона H4 проводили в подглоточном комплексе ганглиев (ЦНС), а также в изолированных командных нейронах оборонительного поведения ЛПа3/2 и ППа3/2, расположенных в левых и правых париетальных ганглиях ЦНС *Helix* (рис. 1).

Основная функция подглоточного комплекса ганглиев заключается в формировании оборонительного поведения. Ключевую роль в этом процессе играют командные нейроны оборонительного поведения ЛПа3/2 и ППа3/2, размер которых составляет около 200 мкм.

Для анализа влияния серотонина и нейропептида FMRFамида на уровень тотального ацетилирования гистона H4 ЦНС улиток инкубировали 1.5 часа в физиологическом растворе для беспозвоночных, содержащем или серотонин (0.2 мМ), или FMRF-амид (10 мкМ), или оба вещества. Это время инкубации с серотонином приводит к долговременной сенситизации нейронов, вовлеченных в формирование оборонительных рефлексов улитки, а с FMRFамидом – вызывает развитие привыкания (Balaban, 2002; Guan et al., 2002). Оба эти процесса относятся к неассоциативным формам долговременной памяти. В качестве контроля использовали экстракты, полученные из ЦНС, инкубируемой в физиологическом растворе без добавления серотонина или FMRFамида.

**Анализ влияния серотонина и нейропептида FMRF-амида на ацетилирование гистона H4 в подглоточном комплексе ганглиев *Helix*.** Результаты проведенных методом Вестерн-блот гибридизации экспериментов показывают, что инкубация ЦНС с серотонином приводит к достоверному увеличению уровня ацетилирования гистона H4 ($p < 0.001$) в суммарных гомогенатах подглоточного комплекса ганглиев спустя один час после обучения, а добавление в инкубационную среду нейропептида FMRFамида эффект серотонина полностью нивелирует (рис. 2).

Отличие уровня ацетилирования гистона H4 при инкубации ЦНС с серотонином и смесью серотонин + FMRFамид достоверно при $p < 0.001$. Нейропептид FMRFамид, примененный независимо, достоверного влияния на уровень ацетилирования не оказывает. Отличие ацетилирования гистона H4 при инкубации ЦНС с серотонином от контроля, серотонина с добавлением FMRFамида и просто

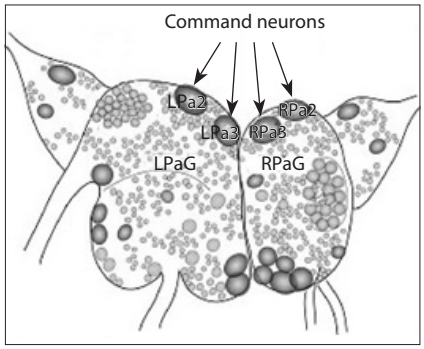



**Fig. 1.** Schematic location of identified neurons in the subesophageal ganglion complex (dorsal surface) of the *Helix lucorum* CNS.

Left and right cerebral ganglia with the LPa(2/3) and RPa(2/3) giant defensive behavior command neurons, symmetrically located in the left (LPaG) and right (RPaG) parietal ganglia, respectively, are shown, after (Kharchenko et al., 2010) with modifications.

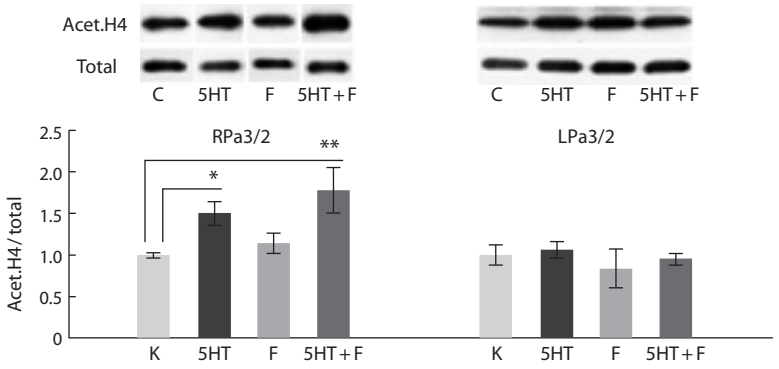



**Fig. 3.** Effect of FMRFamide and serotonin on the acetylation of histone H4 in the command neurons of the right and left parietal ganglia of the *Helix* CNS.

Y-axis: Amounts of acetylated histone H4 forms relative to the total amount of histone H4 and to the control. C, control; 5HT, incubation of CNS with serotonin; 5HT + F, incubation with serotonin and FMRFamide; Acet.H4, acetylated histone H4; Total, total histone H4. * $p < 0.03$, post hoc LSD; ** $p < 0.01$, post hoc LSD, Scheffe, Tukey HSD. Numbers of independent experiments: command neurons RPa 3/2: C ($n$ = 4), 5HT ($n$ = 5), F ($n$ = 3), 5HT + F ($n$ = 3); command neurons LPa 3/2: C ($n$ = 4), 5HT ($n$ = 4), F ($n$ = 3), 5HT + F ($n$ = 3). Command neurons from three individual animals were combined for analysis.

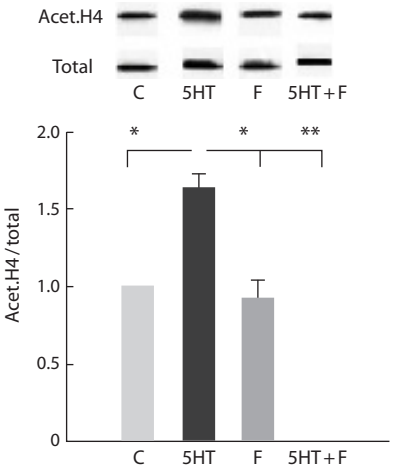



**Fig. 2.** Histone H4 acetylation in the subesophageal ganglion complex after incubation of the *Helix* CNS with FMRFamide, serotonin, and the combination of serotonin and FMRFamide.

Y-axis: Amounts of acetylated histone H4 forms relative to total H4 and to the control. C, control; 5HT, incubation of CNS with serotonin; F, incubation with FMRFamide; 5HT + F, incubation with serotonin and FMRFamide; Acet.H4, acetylated histone H4; Total, total histone H4. C/5HT * $p < 0.003$; 5HT/(5HT + F) * $p < 0.003$; 5HT/F ** $p < 0.001$; F (3.24) = = 13.803, $p$ = 0.0002. (ANOVA). Numbers of independent experiments: C ($n$ = 8); 5HT ($n$ = 9); F ($n$ = 5); 5HT + F ($n$ = 6).

The data are shown as mean ± SEM normalized ratios. Upper panel: an exemplary Western blot with antibodies against the acetylated form of histone H4 and against the total histone H4.

FMRFамида подтверждается post-hoc LSD, Scheffe, Tukey HSD (ANOVA). Полученные результаты позволяют прийти к заключению о диаметрально противоположном действии серотонина и FMRFамида на интегральный уровень ацетилирования H4 в суммарной ЦНС виноградной улитки.

**Анализ влияния серотонина и нейропептида FMRFамида на ацетилирование гистона H4 в командных нейронах оборонительного поведения ЦНС *Helix*.** Для более детального изучения влияния серотонина и FMRFамида на ацетилирование гистона H4 в ЦНС улитки на следующем этапе работы нами были исследованы изолированные командные нейроны ЛПа3/2 и ППа3/2, симметрично расположенные в левых и правых париетальных ганглиях (см. рис. 1).

Показано, что под действием серотонина возрастает уровень ацетилирования гистона H4 в командных нейронах ППа3/2 (достоверно при $p < 0.03$), а в ЛПа3/2 он не меняется (рис. 3). Добавление только FMRFамида в культуральную среду достоверного изменения ацетилирования не вызывает (см. рис. 3). Однако обнаружена аддитивность в действии серотонина и FMRFамида в командных нейронах ППа3/2, в случае если оба вещества присутствовали в культуральной среде. Показано, что совместное внесение в среду серотонина и FRMFамида вызывает значительное увеличение ацетилирования гистона H4 относительно контроля ($1.78 \pm 0.27$ $p < 0.01$, ANOVA). При этом на уровень ацетилирования гистона H4 в нейронах ЛПа3/2 смесь серотонина и FRMFамида влияния не оказывает ($0.92 \pm 0.03$).

Следовательно, серотонин и FMRFамид оказывают аддитивный положительный эффект на тотальный уровень ацетилирования гистона H4 в командных нейронах оборонительного поведения правых париетальных ганглиев, не влияя при этом на уровень ацетилирования гистона H4 в симметрично расположенных командных нейронах левых париетальных ганглиев.

## Обсуждение

Полученные нами результаты четко показывают влияние сочетания серотонина и FMRFамида на ацетилирование гистона H4 как в суммарной ЦНС, так и в командных нейронах оборонительного поведения ППа3/2 виноградной улитки. Однако эффекты взаимодействия этих веществ кардинально различны. В суммарной ЦНС (подглоточный комплекс ганглиев) индуцированное серотонином возрастание уровня ацетилирования H4 снимается нейропептидом FMRFамидом, а в командных нейронах оборонительного поведения правых париетальных ганглиев наблюдается синергизм в действии этих медиаторов на ацетилирование гистона H4.

Таким образом, оказывается, что FMRFамид способен не только тормозить действие серотонина на ацетилирование гистона H4 (сходный результат был получен и в суммарных экстрактах ЦНС моллюска *Aplysia*) (Guan et al.,

2002), но и в отдельных функционально важных нейронах оборонительной сети усиливать его. С одной стороны, подавление FMRFамидом ацетилирования, индуцируемого серотонином в ЦНС моллюсков, свидетельствует о возможности тормозного влияния FMRFамида на нервные сети, в дальнейшем не вовлекаемые в пластические перестройки, что в конечном итоге тормозит поведенческие реакции на стимулы, не значимые для изучаемого типа памяти. С этим положением хорошо согласуются данные о том, что FMRFамид представляет собой тормозной медиатор, играющий существенную роль в формировании депрессии или привыкания (Guan et al., 2002). С другой стороны, как показано нами, в отдельных идентифицированных нейронах FMRFамид способен выступать в качестве синергиста серотонина и оказывать активирующее влияние на так называемые зоны пластичности. Наши результаты согласуются с работой (Дьяконова, Ш.-Рожа, 1986), в которой показано, что серотонин и FMRFамид могут оказывать однонаправленное действие на пластические характеристики изучаемых командных нейронов.

На существование зависящих как от структуры головного мозга, так и от формы обучения различий в уровне ацетилирования гистонов H3 и H4 при формировании долговременной памяти указывают также результаты исследований этих процессов, проведенных на позвоночных животных (Levenson et al., 2004; Bredy et al., 2007; Takase et al., 2013). Такая специфика, вероятнее всего, обусловлена клеточной индивидуальностью экспрессии различных гистон-ацетилаз, а также особенностями регуляторных путей, связанных с модуляцией активности этих ферментов и их привлечением к определенным участкам хроматина.

Кроме того, в настоящей работе мы обнаружили, что серотонин и FMRFамид изменяют уровнь ацетилирования гистона H4 в командных нейронах оборонительного поведения правых ганглиев ППа3, но не оказывают никого действия на аналогичные нейроны в левых (ЛПа3). Ранее такая асимметрия была зафиксирована и при изучении влияния FMRFамида на пластические процессы в командных нейронах (Дьяконова, Ш.-Рожа, 1986). Показано, что FMRFамид вызывает блокирование привыкания к ритмической внутриклеточной стимуляции импульсами тока в нейроне ЛПа3 и не вызывает блокирования в клетке ППа3. На наличие функциональной асимметрии в этих нейронах свидетельствует и работа В.А. Дятлова (1988), в которой выявлено усиление ацетилхолинового ответа в ППа3 под действием серотонина и ослабление его в ЛПа3. Можно предполагать, что различия связаны с тем, что исследуемые нейроны содержат индивидуальные подтипы серотониновых и/или FMRFамидных рецепторов, однако этот вопрос требует специального изучения.

В наших предыдущих работах было установлено наличие асимметрии в командных нейронах ППа3/2 и ЛПа3/2 также на уровне активации серотонин-индуцируемого каскада MAPK/ERK и ацетилирования гистона H3 при формировании у *Helix* условного оборонительного рефлекса пищевой аверсии (Danilova et al., 2010; Kharchenko et al., 2010) и выдвинуто предположение о том, что, возможно, такая асимметрия способна отражать латеризацию памяти у беспозвоночных. В последние годы показано, что асимметрия играет важную роль не только у человека и позвоночных, но и у беспозвоночных животных. Восстановление симметрии приводит к значительным функциональным нарушениям (Hobert et al., 2002; Rogers, Vallortigara, 2008). Однако как генетические, так и эволюционные механизмы такой организации все еще остаются мало исследованнными.

На феноменологическом уровне роль FMRFамида в функционировании ЦНС ряда других беспозвоночных и позвоночных животных уже довольно хорошо изучена. Эта роль многогранна и включает участие в процессах развития и апоптоза, формирования памяти и снижения болевой чувствительности, а также стимуляции сна (Telegdy, Bollók, 1987; Raffa, 1988; Rőszer, Bánfalvi, 2012; Zatylny-Gaudin, Favrel, 2014; Lenz et al., 2015). В то же время молекулярные механизмы действия этого нейропептида пока изучены недостаточно.

В ганглиях виноградной улитки содержится около 1100 FMRFамид-содержащих нейронов. К ним относятся и исследуемые нами командные нейроны (Elekes, Ude, 1993; Balaban, 2002). Регуляторное действие серотонина и FMRFамида на геном может опосредоваться через внутриклеточные сигнальные каскады MAPK/ERK (Гринкевич, 2012; Гринкевич, Воробьева, 2016) и p38 MAPK (Guan et al., 2003; Гринкевич, 2017). При этом известно, что p38 MAPK участвует в механизмах синаптической депрессии у позвоночных животных (Zhen et al., 2001), что делает возможным проводить параллели с нашей экспериментальной моделью.

В целом полученные в настоящей и предыдущей (Гринкевич, Воробьева, 2016) работах данные позволяют заключить, что запускаемые нейромедиатором серотонином и нейропептидом FMRFамидом сигнальные пути имеют точки пересечения на уровне эпигенетических модификаций гистонов H3 и H4 (с последующими изменениями транскриптомов) и что синергизм или антогонизм действия изученных эффекторов на этом уровне зависит от контекста нейрональных структур. Дальнейшее развитие таких исследований в отдельных идентифицированных нейронах с известной функцией может пролить свет на сложнейшие взаимодействия регуляторных систем, задействованных в формировании долговременной памяти.

## Acknowledgments

This work was supported by the Russian Foundation for Basic Research, project 14-04-01681, and the Basic Research of State Academies 2013–2014 Program (GP-14, section 63).

## Conflict of interest

The authors declare no conflict of interest.

## References

Abel T., Zukin R.S. Epigenetic targets of HDAC inhibition in neurodegenerative and psychiatric disorders. Curr. Opin. Pharmacol. 2008; 8(1):57-64. DOI 10.1016/j.coph.2007.12.002.

Balaban P.M. Cellular mechanisms of behavioral plasticity in terrestrial snail. Neurosci. Biobehav. Rev. 2002;26(5):597-630.

Bredy T.W., Wu H., Crego C., Zellhoefer J., Sun Y.E., Barad M. Histone modifications around individual BDNF gene promoters in prefrontal cortex are associated with extinction of conditioned fear. Learn. Mem. 2007;14(4):268-276. DOI 10.1101/lm.500907.

Danilova A.B., Grinkevich L.N. Inability of juvenile snails for long-term memory formation depends on acetylation status of histone H3 and can be improved by NaB treatment. PLoS One. 2012;7(7):1-8. e41828. DOI 10.1371/journal.pone.0041828.

Danilova A.B., Kharchenko O.A., Shevchenko K.G., Grinkevich L.N. Histone H3 acetylation is asymmetrically induced upon learning in identified neurons of the food aversion network in the mollusk *Helix lucorum.* Front. Behav. Neurosci. 2010;4(180):1-7.

Dyakonova T.L., Sh.-Rozha K. Effect of FMRFamide on electrical and plastic properties of identified neurons of grape snail. Zhurnal Vysshey Nervnoy Deyatelnosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity. 1986;36(4):751-759. (in Russian)

Dyatlov V.A. Role of calcium ions in processes of serotonin-induced modulation of neuronal response to acetylcholone application in *Helix pomatia*. Neurophysiology. 1988;5:489-492.

Elekes K., Ude J. An immunogold electron microscopic analysis of FMRFamide-like immunoreactive neurons in the CNS of *Helix pomatia*: ultrastructure and synaptic connections. J. Neurocytol. 1993; 22(1):1-13.

Gräff J., Tsai L.H. The potential of HDAC inhibitors as cognitive enhancers. Annu. Rev. Pharmacol. Toxicol. 2013;53:311-330. DOI 10.1146/annurev-pharmtox-011112-140216.

Grinkevich L.N. Epigenetics and long-term memory formation. Rossiyskiy Fiziologicheskiy Zhurnal im. I.M. Sechenova = I.M. Sechenov Physiological Journal. 2012;98(5)553-574. (in Russian)

Grinkevich L.N. p38 MAPK is involved in the regulation of epigenetic mechanisms of food aversion learning. Bulletin of Experimental Biology and Medicine. 2017;163(4):412-414. DOI 10.1007/s10517-017-3816-9.

Grinkevich L.N., Lisachev P.D., Kharchenko O.A., Vasil'ev G.V. Expression of MAP/ERK kinase cascade corresponds to the ability to develop food aversion in terrestrial snail at different stages of ontogenesis. Brain Res. 2008;1187:12-19. DOI 10.1016/j.brainres.2007.08.029.

Grinkevich L.N., Vorobiova O.V. Role of modulatory mediator serotonin in induction of epigenetic processes during long-term memory formation in *Helix.* Russian Journal of Genetics: Applied Research. 2014;4(6):526-532. DOI 10.1134/S2079059714060094.

Grinkevich L.N., Vorobiova O.V. Opposing roles of serotonin and neuropeptide FMRFamide in the regulation of epigenetic processes involved in the long-term memory. Russian Journal of Genetics: Applied Research. 2017;7(3):273-280. DOI 10.1134/S2079059717030054.

Guan Z., Giustetto M., Lomvardas S., Kim J.H., Miniaci M.C., Schwartz J.H., Thanos D., Kandel E.R. Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. Cell. 2002;111(4): 483-493.

Guan Z., Kim J.H., Lomvardas S., Holick K., Xu S., Kandel E.R., Schwartz J.H. p38 MAP kinase mediates both short-term and long-term synaptic depression in aplysia. J. Neurosci. 2003;23(19):7317-7325.

Hobert O., Johnston R.J., Chang S. Left-right asymmetry in the nervous system: the *Caenorhabditis elegans* model. Nat. Rev. Neurosci. 2002;3(8):629-640. DOI 10.1038/nrn897.

Kandel E. The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. Mol. Brain. 2012;5(14):1-12. DOI 10.1186/1756-6606-5-14.

Kharchenko O.A., Grinkevich V.V., Vorobiova O.V., Grinkevich L.N. Learning-induced lateralized activation of the MAPK/ERK cascade in identified neurons of the food aversion network in the mollusk *Helix lucorum*. Neurobiol. Learn. Mem. 2010;94:158-166. DOI 10.1016/j.nlm.2010.05.002.

Kim S., Kaang B.K. Epigenetic regulation and chromatin remodeling in learning and memory. Exp. Mol. Med. 2017;49(1):e281. DOI 10.1038/emm.2016.140.

Lenz O., Xiong J., Nelson M.D., Raizen D.M., Williams J.A. FMRFamide signaling promotes stress-induced sleep in Drosophila. Brain Behav. Immun. 2015;47:141-148. DOI 10.1016/j.bbi.2014.12.028.

Levenson J.M., O'Riordan K.J., Brown K.D., Trinh M.A., Molfese D.L., Sweatt J.D. Regulation of histone acetylation during memory formation in the hippocampus. J. Biol. Chem. 2004;279:40545-40559. DOI 10.1074/jbc.M402229200.

Levenson J.M., Sweatt J.D. Epigenetic mechanisms: a common theme in vertebrate and invertebrate memory formation. Cell Mol. Life Sci. 2006;63:1009-1016. DOI 10.1007/s00018-006-6026-6.

Monsey M.S., Ota K.T., Akingbade I.F., Hong E.S., Schafe G.E. Epigenetic alterations are critical for fear memory consolidation and synaptic plasticity in the lateral amygdala. PLoS One. 2011;6(5):e19958. DOI 10.1371/journal.pone.0019958.

Raffa R.B. The action of FMRFamide (Phe-Met-Arg-Phe-NH2) and related peptides on mammals. Peptides. 1988;9(4):915-922.

Rogers L.J., Vallortigara G. From antenna to antenna: lateral shift of olfactory memory recall by honeybees. PLoS One. 2008;3(6):1-5. DOI 10.1371/journal.pone.0002340.

Rőszer T., Bánfalvi G. FMRFamide-related peptides: anti-opiate transmitters acting in apoptosis. Peptides. 2012;34(1):177-185. DOI 10.1016/j.peptides.2011.04.011.

Takase K., Oda S., Kuroda M., Funato H. Monoaminergic and neuropeptidergic neurons have distinct expression profiles of histone deacetylases. PLoS One. 2013;8(3):e58473. DOI 10.1371/journal.pone.0058473.

Telegdy G., Bollók I. Amnesic action of FMRFamide in rats. Neuropeptides. 1987;10(2):157-163.

Zatylny-Gaudin C., Favrel P. Diversity of the RFamide peptide family in mollusks. Front. Endocrinol. (Lausanne). 2014;5(178):1-14. DOI 10.3389/fendo.2014.00178.

Zhen X., Du W., Romano A.G., Friedman E., Harvey J.A. The p38 mitogen-activated protein kinase is involved in associative learning in rabbits. J. Neurosci. 2001;21(15):5513-5529.

Zovkic I.B., Guzman-Karlsson M.C., Sweatt J.D. Epigenetic regulation of memory formation and maintenance. Learn. Mem. 2013;20:61-74. DOI 10.1038/npp.2012.79.

Вавиловский журнал генетики и селекции. 2018;22(5):611-619
DOI 10.18699/VJ18.402



# Estimating the impact of the Mongol expansion upon the gene pool of Tuvans

L.D. Damba[1, 2], E.V. Balanovskaya[2, 3], M.K. Zhabagin[4], Y.M. Yusupov[5], Y.V. Bogunov[2, 6], Z.M. Sabitov[7], A.T. Agdzhoyan[2, 6], N.A. Korotkova[2, 3], M.B. Lavryashina[8], B.B. Mongush[9], U.N. Kavai-ool[10], O.P. Balanovsky[3, 6] ✉

[1] Research Institute of Medical and Social Problems and Control of the Healthcare Department, Kyzyl, Russia
[2] Research Centre for Medical Genetics, Moscow, Russia
[3] Biobank of Northern Eurasia, Moscow, Russia
[4] National Center for Biotechnology, Astana, Kazakhstan
[5] Institute for Strategic Studies of the Republic of Bashkortostan, Ufa, Russia
[6] Vavilov Institute of General Genetics, RAS, Moscow, Russia
[7] L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
[8] Kemerovo State University, Kemerovo, Russia
[9] Tuvan Institute of Humanitarian and Applied Socio-Economic Studies, Kyzyl, Russia
[10] Tuvan State University, Kyzyl, Russia

With a view to trace the Mongol expansion in Tuvinian gene pool we studied two largest Tuvinian clans – those in which, according to data of humanities, one could expect the highest Central Asian ancestry, connected with the Mongol expansion. Thus, the results of Central Asian ancestry in these two clans component may be used as upper limit of the Mongol influence upon the Tuvinian gene pool in a whole. According to the data of 59 Y-chromosomal SNP markers, the haplogroup spectra in these Tuvinian tribal groups (Mongush, $N = 64$, and Oorzhak, $N = 27$) were similar. On average, two-thirds of their gene pools (63 %) are composed by North Eurasian haplogroups (N*, N1a2, N3a, Q) connected with autochtonous populations of modern area of Tuvans. The Central Asian haplogroups (C2, O2) composed less then fifth part (17 %) of gene pools of the clans studied. The opposite ratio was revealed in Mongols: there were 10 % North Eurasian haplogroups and 75 % Central Asian haplogroups in their gene pool. All the results derived – "genetic portraits", the matrix of genetic distances, the dendrogram and the multidimensional scaling plot, which mirror the genetic connections between Tuvinian clans and populations of South Siberia and East Asia, demonstrated the prominent similarity of the Tuvinian gene pools with populations from and Khakassia and Altai. It could be therefore assumed that Tuvinian clans *Mongush* and *Oorzhak* originated from autochtonous people (supposedly, from the local Samoyed and Kets substrata). The minor component of Central Asian haplogroups in the gene pool of these clans allowed to suppose that Mongol expansion did not have a significant influence upon the Tuvinan gene pool at a whole.

Key words: Y-chromosome; SNP-polymorphism; haplogroup; ethnogenesis; tribal group; Tuvans.

## Оценка вклада монгольской экспансии в генофонд тувинцев

Л.Д. Дамба[1, 2], Е.В. Балановская[2, 3], М.К. Жабагин[4], Ю.М. Юсупов[5], Ю.В. Богунов[2, 6], Ж.М. Сабитов[7], А.Т. Агджоян[2, 6], Н.А. Короткова[2, 3], М.Б. Лавряшина[8], Б.Б. Монгуш[9], У.Н. Кавай-оол[10], О.П. Балановский[3, 6] ✉

[1] НИИ медико-социальных проблем и управления Республики Тыва, Кызыл, Россия
[2] Медико-генетический научный центр, Москва, Россия
[3] АНО «Биобанк Северной Евразии», Москва, Россия
[4] Национальный центр биотехнологии, Астана, Казахстан
[5] Институт стратегических исследований Республики Башкортостан, Уфа, Россия
[6] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[7] Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан
[8] Кемеровский государственный университет, Кемерово, Россия
[9] Тувинский институт гуманитарных и прикладных социально-экономических исследований, Кызыл, Россия
[10] Тувинский государственный университет, Кызыл, Россия

Для выявления следа монгольской экспансии в генофонде тувинцев впервые изучены генофонды двух самых многочисленных тувинских родов, для которых по данным гуманитарных наук ожидается наибольший вклад центральноазиатского компонента, связываемого с монгольской экспансией. При таком подходе результаты исследования могут служить верхней оценкой «монгольского влияния» на генофонд тувинцев в целом. По данным о 59 SNP-маркерах Y-хромосомы спектры гаплогрупп генофондов этих тувинских родов (монгуш, $N = 64$; ооржак, $N = 27$) оказались сходными. В среднем две трети их генофондов (63 %) составляют «североевразийские» гаплогруппы (N*, N1a2, N3a, Q), связываемые с автохтонным населением современного ареала тувинцев, тогда как «центральноазиатские» гаплогруппы (C2, O2) составляют менее пятой части (17 %) генофонда изученных тувинских родов. Для монголов наблюдается прямо противоположное соотношение: 10 % «североевразийских» и 75 % «центральноазиатских» гаплогрупп. Все полученные результаты – «генетические портреты», матрица генетических расстояний, дендрограмма и график многомерного шкалирования, отражающие генетические связи тувинских родов с популяциями Южной Сибири и Центральной Азии, свидетельствуют о значительном сходстве генофондов тувинских родов с популяциями Хакасии и Алтая и позволяют сделать

Received 01.04.2018
Accepted for publication 18.06.2018


✉ e-mail: balanovsky@inbox.ru

вывод о формировании тувинских родов монгуш и ооржак на основе автохтонного населения (предположительно – местного самодийско-кетского субстрата). Малый вклад в генофонд этих родов «центральноазиатских» гаплогрупп позволяет считать, что и на генофонд тувинцев в целом монгольская экспансия не оказала значимого влияния.

Ключевые слова: Y-хромосома; SNP-полиморфизм; гаплогруппа; этногенез; род; тувинцы.

**HOW TO CITE THIS ARTICLE:**
Damba L.D., Balanovskaya E.V., Zhabagin M.K., Yusupov Y.M., Bogunov Y.V., Sabitov Z.M., Agdzhoyan A.T., Korotkova N.A., Lavryashina M.B., Mongush B.B., Kavai-ool U.N., Balanovsky O.P. Estimating the impact of Mongol expansion on gene pool of Tuvans. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):611-619. DOI 10.18699/VJ18.402

The Tuvan Republic (Tuva or Tyva) is situated in the very heart of Asia, lying at its geographical center among the mountains and foothills of the Altai-Sayan region. Both ancient and more recent ways of migration passed through Tyva, including the Silk Road that connected the East and the West of Eurasia. That is responsible for sophisticated culture-genesis of indigenous population, as the territory has long been exposed to various tribes mixing and formation, altering of languages (there is one of Turkic languages spoken today) and cultural paradigms (along with buddhism, traditional shamanism is still being practiced). Those processes might have considerably affected the gene pool of the Tuvans. However, one of the most acute problems in revealing the ontogeny is the extent of the Mongol expansion influence.

Tuvinian and Mongoloid tribe groups had long been historically united while being ruled by the Chinese Qing dynasty as a part of Outer Mongolia. After the Mongolian national revolution in 1911 and the events followed, Mongolia declared independence from China, while Tuvans got independent in 1926 to establish the Tuvan People's Republic. In 1944, the Republic became a part of the Soviet Union. The historical destinies interplay of Tuvinian and Mongoloid tribes may have left a considerable mark on the gene pool of modern Tuvans. The most correct way to reveal the impact of the Mongol expansion in the Tuvinian gene pool forming is to study the clans for which this influence is supposed to be most significant.

In 2010, Tuvans numbered 249 thousand people by population census (Results of the All-Russian Census of 2010) and were subdivided into 26 clans (Serdobov, 1971). Unfortunately, the ethnographers from the Tuvan Institute of Humanitarian and Applied Socio-Economic Studies report no integrated historical ethnographic and linguistic research on settling and populations of Tuvinian clans having been conducted since 1970. However, in 7 of 26 clans (namely, Mongush, Oorzhak, Ojun, Kuzhuget, Khertek, Salchak, Donghak), ethnographers suppose the Mongol ancestry to be considerable. Representatives of those seven clans make up to half of all the Tuvinian people, while two of them (Mongush and Oorzhak) comprising about 33 % of the whole population. Thus, to assess the Mongolian influence upon the gene pool of Tuvans, we have studied the gene pools of the two most numerous of those seven clans being presumingly most influenced by the Mongol expansion; it is difficult to make representative samples for the other minor clans. The evaluation obtained is considered as the upper assessment of the Mongolian ancestry; as for the rest of the Tuvans, ethnographers, historians and linguists conceive the influence to be insignificant.

For Mongush and Oorzhak, West and partially, Central Tyva are the autochtonous territories. However, the Mongush representatives (about 57.5 thousand people) inhabit all the koshuuns of Tyva, and the Oorzhak representatives (about 25 thousand people) live mainly in the western part of Tyva (Barun-Khemchinsky and Dzun-Khemchinsky koshuuns). There has not been consensus either on the scientific etymology of those clans' names, or their ethnogenesis. Two controversial hypotheses (Mongolian and Turkic ones) do interpret different combinations of ethnographic, historical and linguistic data. We have detailed these hypotheses at another paper (Damba et al., 2018), so only brief information is given below.

**The "Mongolian" hypothesis.** ***Mongush.*** The most popular is the version of Mongolian origin of the ethnonym directly linking it to the ethnonym "mongol". Ancient Mongols used to have the form "mongyus". A famous sinologist N.Ts. Munkuev considered it not only as one of transcriptions, but a special form of the ethnonym "mongol" (Tatarintsev, 1980).

***Oorzhak.*** This ethnonym is referred to medieval mongolisms related to the words "oghurchak" in old written Mongolian ("abandoned", "left", "lonely"), and "oortzog" in modern Mongolian ("distinct", "separate").

**The "Turkic" hypothesis.** According to B.I. Tatarintsev (1986), who was a turkologist specializing in Tyva, the ethnonyms "mongush" and "oorzhak" have Turkic origin, so does the ethnonym "tuva". Most historians agree that this ethnonym was first recorded in Chinese documents of VII century. The tribe "dubo" belonged to Tiele people, along with ancient Uyghurs (Bichurin, 1950). The tribe "tuba" ("tubas") was also mentioned in the Mongolian document of XIII century "The Secret History of the Mongols" as the one conquered in 1207 (The Secret History, 1941).

***Oorzhak.*** This clan is suggested to get its modern name "Oorzhak"–"Oghurtchak" from the Mongols not later than in XIII–XIV centuries during the Mongol military expansion, as in modern Mongolian language the word "ogurtchak" was changed by "oortzog" (Tatarintsev, 1986). The Oghur tribes are known from the written sources of V–VII centuries. Some of those tribes migrated westward, giving rise to the peoples of Khazar-Bulgarian subgroup of Turkic languages. On the other hand, up to most turkologists, the ethnonym "uyghur" originates from "oghur" as nomadic Uyghurs had been dominating the territories of modern Mongolia and nearby areas for more than a century (745–847 AD). Thus,

the ethnonym "oorzhak" may go back to the Turkic ethnonym "oghur" (Klyashtorny, Savinov, 2005). Later, the Mongols used the word "oghurtchak" to name the tribes migrating from their native lands to escape invaders. It is known that in XVII century, the Oorzhaks were migrating with other clans about Mountain Altai and along the banks of the upper Ob river outside of the present-day Tuva territory (Serdobov, 1971). B.O. Dolgikh (1960) was of the opinion that the Oorzhaks were Mongolian-speaking in XVII–XVIII centuries, nevertheless he cited the document dated to 1652 and referring the Oorzhaks to Turkic-speaking clans.

***Mongush.*** B.I. Tatarintsev gave preference to searching for ethymology of the name "*mongush*" in South Siberian and Old Turkic languages. In Turkic ethnonymy, "mongush" correlates to the Kyrgyz "munghush", the Bashkir "munash", the Chuvash "mon", the Yakut "mun", and the Old Turkic "mungas" (Old Turkic Dictionary, 1969). B.I. Tatarintsev admitted the presence of Mongolian language component in the Tuvinian ethnos, but he supposed the result to be minimal; thus he did not manage to find its direct reflection on the Tuvinian ethnomy and the ethnogenesis of Tuvinian clans (Tatarintsev, 1986).

Hence there are two alternative versions of the origin of the both most numerous Tuvinian clans. The first one connects their ethnogenesis to Mongols and the entire invasive population from Central Asia, while the other associates their origin with local Turkic-speaking population, the autochtonous people later adopted Turkic.

Purpose of the study: the work targets to analyze the gene pool of most numerous Tuvinian clans Mongush and Oorzhak by genotyping of a wide range of the Y-chromosome SNP markers. In those clans a significant impact of Mongolian factor is supposed. The aim is to detect their genetic linkage with peoples of Siberia and Central Asia and establish "the upper assessment of Mongolian influence" upon the entire gene pool of Tuvans.

## Materials and methods

As the material for the study, genomic DNA was obtained by phenol-chloroform extraction of venous blood.

The total number of population samples $N = 91$; the samples from the Mongush clan were collected in Chaa-Kholsky koshuun ($N = 64$) and of the Oorzhak clan – in Barun-Khemchiksky koshuun ($N = 27$) of the Tyva Republic (Fig. 1). Blood samples were collected only from male indigenous Tuvan representatives, degree of relatedness between any of them being more than 3, and all the ancestors belonging to the same clan and living at the same territory for more than 3 generations. The informed consent to participate in the investigation controlled by the Comission of Ethics of the Research Centre for Medical Genetics (Moscow) was received from all the individuals under the study.

SNP-markers genotyping was carried out by real-time PCR method at StepOnePlus and 7900HT systems (Applied Biosystems) using Taqman assays (Applied Biosystems). 59 SNP-markers were genotyped: M130, M217, F2613, F2386, F1788, F3918, F3830, M86, F5485, SK1066, F3791, F11899, F5481, F11791, F14768, F3960, P53.1, CTS4021, M407, Z12266, M174, M69, M170, M253, M223, P37, M304, M267, M172, M47, M67, M92, M12, M9, M20, M231, LLY22g, M178, L708, L666, B211, M2118, VL29, Z236, F4205, P31, M122, M242, M120, M378, M207, M198, M458, M343, M73, M269, M124, M70, and chrY:15310670 T>C. The nomenclature of haplogroups (hg) in the table of hg frequencies (see the Table) is given as defined by ISOGG Y-DNA Haplogroup Tree 2018 (https://isogg.org/tree/ISOGG_YDNATreeTrunk.html), names of new-discovered markers subdividing hg N3 into sub-clades N3a1, N3a2, N3a3, N3a4, and N3a5 being in accordance to proposed by Ilumäe and his colleagues (Ilumäe et al., 2016).

Nei's genetic distances were calculated by hg frequences (Nei, 1975). Software Djgenetic (www.genofond.ru) elaborated by Y.A. Seryogin and E.V. Balanovskaya was used for calculations. Multivariate analysis methods were conducted using a STATISTICA 6.0 software package (StatSoft. Inc., 2001) applying Ward's method. The results were compared with the unpublished data on Y-chromosome hg frequencies in populations of Siberia and Central Asia from "Y-base" database elaborated under the supervision of O.P. Balanovsky (www.genofond.ru).

## Results and discussion

### "Genetic portrays" of Mongush and Oorzhak clans

In the Mongush clan, there have been revealed 10 haplogroups within the gene pool (see Fig. 1 and Table). Three most frequent hg C2, hg Q, and hg R1a1a comprising more than a half of the entire gene pool were detected with equal frequencies (19 %). More than one third of the gene pool (39 %) was constituted by sub-clades of hg N. Other haplogroups (total 6 %), namely O2, R2a, R1b occurred in single Mongush representatives.

In the gene pool of the Oorzhak clan, there have been detected 9 haplogroups (see Fig. 1 and Table). Hg N3a turned to be the most frequent (30 %), while N1a2, N3a5 and R1a1a haplogroups displayed equal frequencies of 15 %. Lower frequency was observed for hg Q and hg C2b1a3. Others (total 12 %) – C2b1a2a, O2, R1b – were seen in single individuals from the Oorzhak clan.

As a whole, gene pools of the two clans under the study were characterized by similar hg spectra, though their frequencies differed; this can be explained by the fact that one sampling value was twice as large as another. In both clans, two thirds of Y gene pools (in total sample, average 63 %) were represented by North Eurasian hg N and hg Q (see Table), while on the contrary, Central Asian hg C2 and hg O had an insignificant place in the gene pools of Tuvinian clans, average frequency being about 17 %.

Let us consider the hg distribution detected in gene pools of Tuvinian clans in detail.

**Haplogroup C2** peaks in Central Asia (Wells et al., 2001; Zerial et al., 2003), though its variants are abundant in other peoples of Siberia and Far East. For instance, in one of Buryat clans, namely Ekhirids, hg C2 frequency is 88 % (Y-base); in Kazakhs from different regions of Kazakhstan, total occurrence of hg C2 variants averages between 17 and 81 % (Abilev et al., 2012; Zhabagin et al., 2013, 2014, 2017), in populations of the Amur River (such as Nanais, Negidals, Nivkhs, Ulchs) – between 40 and 65 %, in Evenks – up to 68 % (Y-base), in Kyrgyz people of Pamir-Alay – up to 22 %, correspondingly; of all Turkic peoples of Altai, relatively

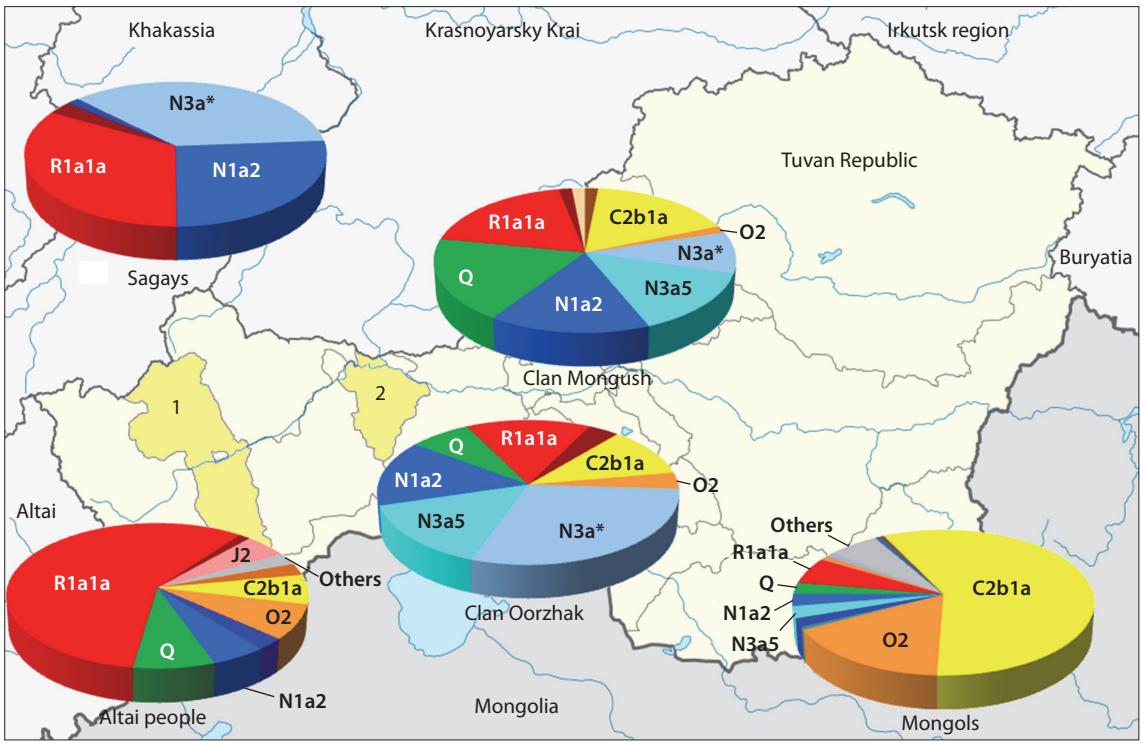



**Fig. 1.** Y-chromosomal haplogroup spectra in gene pools of Tuvinian Oorzhak and Mongush clans and of the neighboring populations of South Siberia and Central Asia.

The sectors of the diagrams display proportions of haplogroups in gene pools. 1 – Barun-Khemchiksky koshuun (site of collecting sample material from Oorzhak representatives); 2 – Chaa-Kholsky koshuun (site of collecting sample material from Mongush representatives).

Y-chromosome frequencies in gene pools of Tuvinian clans Oorzhak and Mogush and neighboring populations of South Siberia and Central Asia (%)

| Y-chromosomal haplogroup | SNP marker | Clan Oorzhak (*N* = 27) | Clan Mongush (*N* = 64) | Mongols (*N* = 852) | Sagays (*N* = 69) | Altai people (*N* = 76) |
|---|---|---|---|---|---|---|
| D | M174 | 0 | 0 | 0.8 | 0 | 2.6 |
| C2* | M217 | 0 | 1.6 | 0.4 | 0 | 0 |
| C2b1a1 | F3918 | 0 | 0 | 4.0 | 0 | 0 |
| C2b1a3 | F3791 | 7.4 | 0 | 17.5 | 0 | 2.6 |
| C2b1a2a | M86 | 3.7 | 17.2 | 26.6 | 0 | 3.9 |
| C2c1 | F2613 | 0 | 0 | 9.6 | 0 | 0 |
| O2 | M122 | 3.7 | 1.6 | 16.7 | 0 | 7.9 |
| N* | M231 | 0 | 0 | 1.8 | 1.4 | 2.6 |
| N3a | L708 | 29.6 | 9.4 | 0 | 36.2 | 0 |
| N3a5 | F4205 | 14.8 | 14.1 | 2.7 | 0 | 0 |
| N1a2 | L666 | 14.8 | 15.6 | 2.7 | 26.1 | 5.3 |
| Q | M242 | 7.4 | 18.8 | 2.6 | 0 | 7.9 |
| R1a1a | M198(xM458) | 14.8 | 18.8 | 5.9 | 33.3 | 57.9 |
| R1b | M343(xM269) | 3.7 | 1.6 | 1.1 | 2.9 | 1.3 |
| R2a | M124 | 0 | 1.6 | 0.9 | 0 | 0 |
| J2 | M172 | 0 | 0 | 0.7 | 0 | 5.3 |
| Others | | 0 | 0 | 6.1 | 0 | 2.6 |
| | | | Summary values (%) | | | |
| "North Eurasia" | N и Q | 67 | 58 | 10 | 64 | 16 |
| "Central-Asian" | C и O | 15 | 20 | 75 | – | 14 |

Note: Symbol "*" in hg names highlights the ancestral marker presence and absence of mutations of all the daughter haplogroups. "Others" combine the haplogroups with frequency lower than 1 %.

high hg C2 frequency (16 %) is detected only in Telengits (Balanovskaya et al., 2014; Balaganskaya et al., 2011a, 2016). In Tuvinian clans under the study, hg C2 frequency is rather low – 19 % in Mongush and 11 % in Oorzhak, while in Mongols it makes up almost two thirds of the entire gene pool an comprises different genetic lines (subhaplogroups).

Despite being considered as typical for population of Central Asia, **haplogroup O** is rare in Tuvinian clans, that is one carrier of this hg has been detected in each of Mongush and Oorzhak clans, while in Mongols, its frequency is the second and reaches up to 17 % of the entire gene pool (see Fig. 1 and Table).

Hence, in the gene pools of even considered to be the most affected by the Mongol expansion clans, the arbitrarily Central Asian hg C and hg O, which compose three quarters of the Mongolian gene pool (75 %, see Table), are minor. Thus, in the entire Tuvinian gene pool, one can infer that frequency of those haplogroups is considerably lower (judging from the ratio between the populations of clans presumably affected by the Mongol expansion and of clans for which such influence is precluded). Further analysis of distribution throughout Eurasia of hg C and hg O new sub-clades, which can be traced by genome-wide studies of Y-chromosome and mass screening in indigenous peoples of Eurasia by new sub-clades, will permit to clarify the ethnogenetic links between Tuvans and other populations. Yet, the obtained data make us consider the low percentage of "Central Asian" component in gene pools of Tuvinian clans as reflection of a small genetic layer that can probably mark the trace of the Mongol ancestry. However, the other hypotheses need checking. One branch of the Silk Road (an ancient trade route that had connected East Asia and the Mediterranean 15 centuries before the Mongol expansion) crossed Tyva and could bring Central Asian gene pools as well. The territory of present-day Tyva used to be under cultural influence of Hun (Xiongnu or Chanyu) Empire, which could be attended with one of the very first waves of genetic interaction between Central Asian and South Siberian populations.

**Haplogroup N** is abundant all over North Eurasia from Scandinavia to Far East (Rootsi et al., 2007). The study on whole Y-chromosome sequencing conducted with participation of our group (Ilumäe et al., 2016) subdivided this haplogroup into several branches with their regional distribution. In gene pools of the Tuvans involved, hg N was represented by two sub-clades, namely N1a2 and N3a.

**Sub-clade N1a2** peaks in populations of West Siberia (in Nganasans, frequency is 92 %) and South Siberia (in Khakas 34 %, in Tofalars 25 %) (Y-base). In Tuvans, N1a2 occurrence is nearly16 % in Mongush and 15 % in Oorzhak clans, respectively, while in Mongols, the frequency is three times less (5 %). Hg N1a2 is supposed to display the impact of the Samoyedic component to the gene pool of Tuvinian clans (Kharkov et al., 2013).

**Sub-clade N3a** is major in the Oorzhak clan comprising almost half of the gene pool (45 %); it is represented by two sub-clades, namely **N3a*** and **N3a5**. The same sub-branches are specific to the Mongush clan as well, though with lower frequencies: N3a* – 9 % and N3a5 – 14 % (see Table). In Khori-Buryats from the Transbaikal region, a high frequency is observed – 82 % (Kharkov et al., 2014), while in Mongols, N3a5 occurs rather rarely (6 %). Hg N3a* was detected in populations of South Siberia only, and was widely spread in Khakas-Sagays and Shors (up to 40 %) (Ilumäe et al., 2016) (Y-base).

**Haplogroup Q** is most spread in West and South Siberia. It reaches its maximum frequencies (Y-base) in gene pools of Kettic people (84 %), Selkups (66 %), populations of North Altai (more than a half of the gene pool in Chelkans), Koibals and Kyzyl Khakas (44 %) (Balaganskaya et al., 2011b). Hg Q makes up one fifth of the Mongush gene pool, with frequency only 3 % in Mongols. Presence of this haplogroup in Tuvinian clans can reflect the genetic contribution of Samoyedic and Paleosiberian tribes.

Thus, in the gene pools of Oorzhak and Mongush Tuvinian clans, which are speculated to be mostly affected by Mongolian ancestry, two thirds of the entire gene pools are represented by typical "North Eurasian" haplogroups (58 % in Mongush and 67 % in Oorzhak). This proves the hypothesis of significant autochtonous component presence within the gene pools of those clans; moreover, we have not detected hg N3a* anywhere else, but South Siberian populations.

Within the pan-Eurasian **haplogroup R1a1a**, two large genetic lines (sub-haplogroups) are identified: "European" (marker M458) and "Asian" (marker Z93) the latter almost never occurring in Europe (Balanovsky, 2015) but abundant in South Siberia and northern Hindustan. In the Altai-Sayan region, high frequencies of the "Asian" branch are spread in many peoples – Shors, Tubalars, Altai-Kizhi people, Telengits, Sagays, Kyzyl Khakas, Koibals, Teleuts (Y-base) (Kharkov et al., 2009). Hg R1a1a comprises perceptible parts of gene pools of Tuvinian clans (19 % in Mongush, and 15 % in Oorzhak), though its occurrence in Mongols is much lower (6 %). Those results also count in favor of the hypothesis of autochtonous component dominance even in the gene pools of clans potentially most influenced by Mongolian ancestry. If we add R1a1a variants to the "North Eurasian" haplogroups, the "not-Central Asian" component will compose average four fifth of the entire gene pools for Tuvinian clans (in Mongush 77 %, and in Oorzhak 81 %), being only 16 % in Mongols. Such data are definitely contrary to the hypothesis of a crucial influence of the Mongol expansion upon the development of Tuvinian gene pool.

### Tuvinian clans position within the genetic space of Siberia and Central Asia

Position of the analyzed Mongush and Oorzhak clans within the genetic space of Siberian and Central Asian peoples by all their haplogroups cumulatively is assessed by the matrix of genetic distances (d) that displays the populations studied by genotyping of the same wide range of the Y chromosomal SNP markers as Tuvans (Fig. 2). The first column presents average genetic distances ($\overline{d}$) for both clans, then all the populations are ranged by their average genetic affinity to the Tuvinian clans and divided into two classes – of close to the clans' populations ($\overline{d} < 1$) and of distant ones ($\overline{d} > 1$, highlighted on the grey background, Fig. 2, *a*).

To begin with, let us note that Khakas-Sagays are almost two times closer genetically to the Oorzhak clan (d = 0.16) than both clans to each other (d = 0.28), while the other Khakas groups of Koibals and Kyzyl people are the same close to the Mongush clan (d = 0.26) as the clans to each other (d = 0.28).

*a*

| Populations | $\bar{d}$ genetic distances | Clan Oorzhak | Clan Mongush |
|---|---|---|---|
| Clan Oorzhak | 0.28 | 0 | 0.28 |
| Clan Mongush | 0.28 | 0.28 | 0 |
| Sagays | 0.31 | 0.16 | 0.45 |
| Koibals | 0.47 | 0.68 | 0.26 |
| Shors | 0.53 | 0.48 | 0.57 |
| Tubalars | 0.69 | 0.92 | 0.46 |
| Altai-Kizhi | 0.70 | 0.89 | 0.51 |
| Chelkans | 0.72 | 1.05 | 0.40 |
| Kachins | 0.79 | 0.85 | 0.73 |
| Telengits | 0.80 | 0.95 | 0.65 |
| Buryats from Transbaikal Region | 0.98 | 0.99 | 0.97 |
| Mongols | 1.01 | 1.36 | 0.65 |
| Khamnigans | 1.02 | 1.03 | 1.01 |
| Kumandins | 1.44 | 1.56 | 1.32 |
| Kara-Kalpaks | 1.49 | 1.75 | 1.22 |
| Kirghiz | 1.57 | 1.75 | 1.39 |
| Kazakh | 1.70 | 1.71 | 1.70 |
| Buryats from Buryatia | 1.86 | 1.88 | 1.84 |
| Buryats from Irkutsk Region | 3.32 | 3.37 | 3.26 |

*b*

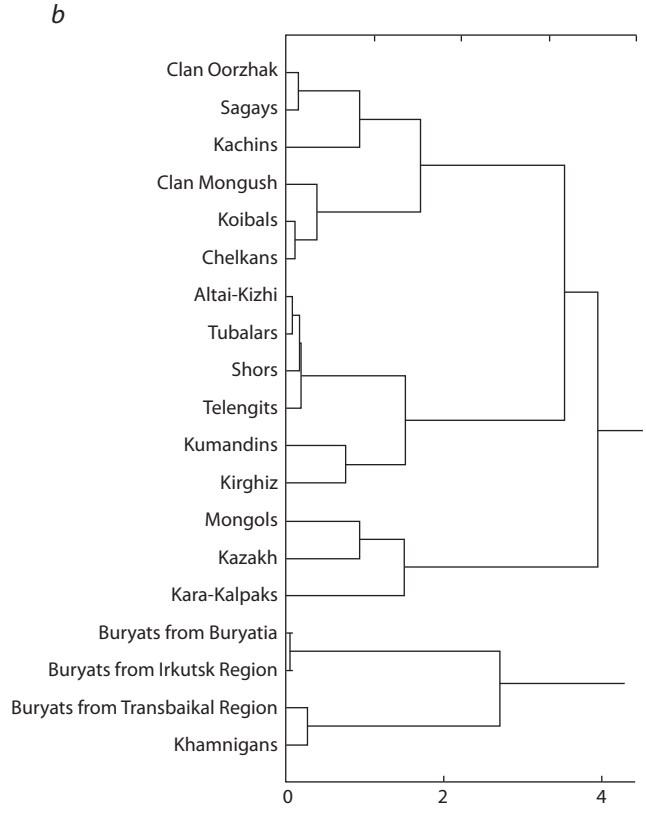



**Fig. 2.** Matrix of genetical distances from the gene pools of Tuvinian clans Oorzhak and Mongush to the populations of Siberia and Central Asia (*a*) and dendrogram of their genetic relationship (*b*).

By and large Khakas, Shors and populations of Northern Altai (but for very peculiar Kumandins; in the dendrogram, they associate with Kyrgyz people) and peoples of Southern Altai. In fact, Buryats of the Transbaikal region are as distant from Tuvans ($\bar{d}$ = 0.98) as Mongols are ($\bar{d}$ = 1.01). Greater similarity of Transbaikal Buryats to Tuvans, as compared with their geographically neighboring Buryats of Irkutsk region (d = 3.37) and of the Buryat Republic (d = 1.88), results from a very high frequency of "North Eurasian" hg N3a5 in the Transbaikal region (82 %), which composes about 15 % in Tuvinian clans.

Mongols are the first in the list of populations genetically distant from Tuvinian clans ($\bar{d}$ > 1), however the Mongush clan is twice closer to them genetically (d = 0.65) than to the Oorzhak clan (d = 1.36). The same is observed for the distance to Karakalpaks (d = 1.22 from the Mongush clan, and d = 1.75 from the Oorzhak one) and to Kyrgyz people (d = 1.39 from the Mongush, and d = 1.75 from the Oorzhak clans). This fact highlights relatively more significant contribution of the "Central Asian" component to the gene pool of the Mongush clan, being in conformity with ethnography data. Yet, the absolute value of this contribution is less in comparison with the genetic impact of the autochtonous population of the regions, as it is shown in dendrogramme (see Fig. 2, *b*). The Oorzhak clan is in one cluster with Sagays and Katschins of Khakassia, while the Mongush clan got to the same cluster with Koibals of Khakasia and Chelkans of Northern Altai. The neighboring cluster comprises Shors and other populations of Altai along with Kyrgyz people. Mongols together with Kazakhs and Karakalpaks are joined to a distant from Tuvans cluster, Buryats and Hamnigans to form an even more genetically remote one. Hence the dendrogram exhibits the conclusion that in Tuvinian gene pool, there prevails the ancestry component from autochtonous population who are genetically similar to their western neighbors (present-day population of Khakhassia and Altai), but distant from Mongols and Buryats, who are their neighbors in the west and the south.

Multidimensional scaling plot appears to be more correct and informative, five clusters being presented (Fig. 3). Great variety of Buryat populations is reflected by two well-distinctive clusters of "west Buryats" and "east Buryats". Southern neighbors Mongols are closer to Tuvans than their eastern neighbors (Buryats of Buryatia and Irkutsk region); but along with other populations of Central Asia (Karakalpaks, Kazakhs and Kyrgyz people), Mongols compose their own "Central Asian" cluster. Within the genetic space of Siberia and Central Asia, a compact cluster of Altai and the Tuvinian-Khakassian one are the closest. In its entirety, the plot does confirm once again the results obtained: both Tuvinian clans have most close relationship to populations of Khakassia and Altai; a weak trace of "Central Asian" component is more distinctive in the Mongush than in the Oorzhak clan.

## Conclusion

Gene pools of two largest Tuvinian clans appeared to be similar by Y-chromosomal spectra of haplogroups, suggesting the affinity of their origin. Different hg frequencies might mainly result from small samples. However, genetic relation of the Mongush clan to Mongolian populations is slightly closer revealing somewhat stronger influence of their southern neighbors. Though the case in point is rather a weak trace of Central Asian genetic ancestry. It is under the question, whether it can be associated with the late Mongol expansion only (manifested by the ethnonym *mongush*) or to earlier constant migrations from Central Asia. Further detection of noval Y-chromosomal sub-clades and mass screening in population of North, Central and East Asia by SNP markers of those new Y-chromosomal sub-clades might provide the answer.

Nevertheless the source of a weak "Central Asian" genetic ancestry is, the performed study unambiguously indicates that the main part of gene pools of Tuvinian clans confidently being associated by ethnographers with the Mongol expansion is inherited from the autochtonous population of the Altai-Sayan highlands. Predominance of "North Eurasian" Q, N1a2 and N3a haplogroups allows to assume that Tuvinian gene pool was formed on the basis of Samoyed and Kets tribes having inhabited the territory of present-day Tyva since the Neolithic (6000–3000 BC). The conclusion about rather low proportion of East Eurasian haplogroups within the gene pools of the Mongush and Oorzhak Tuvinian clans, referring to a later and less significant genetic layer, correlates well with the anthropological data, which consider the South Siberian component of Tuvinian ethnogenesis as an earlier and more relevant in comparison with the South Central Asian one (Aksyanova, 2009). It must be emphasized that genetic and anthropological data betoken only the biological trace of migrations, their demographic intensity and contribution to the gene pool. They say nothing about the intensity of cultural influence,which is the field of ethnographic and historical estimation.

The present study was intentionally concentrated on Y-chromosome paternal (male-specific) lines as they are inherited in the same manner as enthonyms of clans do. However, gene pools of South Siberian and Central Asian ethnic groups are being analyzed by genotyping of the wide range of the autosomal markers. Hence, the work is just a first step towards reconstruction of interaction between indigenous population of South Siberia and Central Asia. Only integral population-genetic and historical-ethnographic investigations into clan structure of Tuvans, Mongols (as present-day Mongols comprise tribes of various ethnogenesis), and other ethnic groups of Central Asia and Siberia might open some pages of history of the very center of Asia, the largest in both territory and population part of the world.

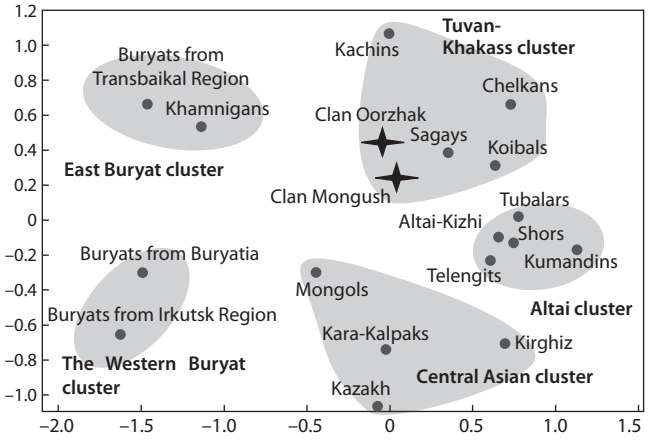



**Fig. 3.** Gene pools of Tuvinian clans Mongush and Oorzhak in the context of the population of South Siberia and Central Asia.

Tuvinian populations: Oorzhak clan, Mongush clan; Khakassia ethnic groups: Katschns, Sagays, Koibals; populations of North Altai: Chelkans, Tubalars, Kumandins; populations of South Altai: Altai-Kizhi, Telengits. The multidimensional scaling plot by frequencies of 75 Y-chromosomal haplogroups in 19 populations of South Siberia and Central Asia. Stress value = 0.13; alienation = 0.15.

## Acknowledgements

This work was supported by the Russian Foundation for Basic Research (grant No. 17-06-00472_a) within the framework of the State Target of FASO Russia for the Research Center for Medical Genetics and of the State Target of FASO Russia for Vavilov Institute of General Genetics (project no. AAAA-A16-116111610171-1).

We are especially grateful to the staff of central kozhuun hospitals of Barun-Khemchinsk and Chaa-Kholsk for invaluable help in collecting the biological material and also to our informant Oleg Konstantinovich Tamdyn for providing the data on population of Tuvinian clans.

## Conflict of interest

The authors declare to have no conflict of interest.

## References

Abilev S., Malyarchuk B., Derenko M., Wozniak M., Grzybowski T., Zakharov I. The Y-chromosome C3* star-cluster attributed to Genghis Khan's descendants is present at high frequency in the Kerey clan from Kazakhstan. Hum. Biol. 2012;84(1):79-89. DOI 10.3378/ 027.084.0106.

Balaganskaya O.A., Damba L.D., Zhabagin M.K., Agdzhoyan A.T., Yusupov Y.M., Sabitov Z.M., Bogunov Y.V., Balaganskiy A.G., Sultanova G.D., Dolinina D.O., Padyukova A.D., Shalyaho R.A., Markina N.V., Bukin A.G., Lavryashina M.B., Balanovskaya E.V., Baranova E.G., Balanovsky O.P. Mongolian trace in gene pool of populations along the steppe of Eurasia. Sovremennye Problemy Nauki i Obrazovaniya = Modern Problems of Science and Education. 2016;4:211. (in Russian)

Balaganskaya O.A., Lavryashina M.B., Kuznetsova M.A., Romanov A.G., Dibirova Kh.D., Frolova S.A., Kuznetsova A.A., Zakharova T.A., Baranova E.E., Teutchezh I.E., Romashkina M.V., Sabitov Z., Tazhigulova I., Nimadava P., Balanovskaya E.V., Balanovsky O.P. Genetic structure of Altai people (Russia, Kazarhstan, Mongolia). Vestnik Moskovskogo Universiteta. Seria XXIII. Antropologia = Vestnik of Moscow University. Ser. XXIII. Antropology. 2011b;2: 25-39. (in Russian)

Balanovskaya E.V., Balaganskaya O.A., Damba L.D., Dibirova Kh.D., Agdzhoyan A.T., Bogunov Y.V., Zha-

bagin M.K., Isakova Z.T., Lavryashina M.B., Balanovsky O.P. Impact of the environment on the gene pool of Turkic-speaking populations of mountains and steppes in Altay, Sayan, Tien Shan and Pamir. Vestnik Moskovskogo Universiteta. Seria XXIII. Antropologia = Vestnik of Moscow University. Ser. XXIII. Antropology. 2014;2:46-55. (in Russian)

Balanovsky O.P. The Gene Pool of Europe. Moscow: KMK Publ., 2015. (in Russian)

Bitchurin N.Y. [Iakinf]. The Collection of Information About the Peoples Inhabiting Central Asia in Ancient Times. Moscow; Leningrad: AN SSSR; Mikluho-Maklay Institute of Ethnography, 1950;302-317. (in Russian)

Damba L.D., Aiyzhy E.V., Mongush B.B., Zhabagin M.K., Yusupov Y.M., Bogunov Y.V., Sabitov Z.M., Agdzhoyan A.T., Markina N.V., Dorzhu C.M., Balanovskaya E.V., Balanovsky O.P. Complex approach in tribal structure of Tuvans by the example of tribal groups *Oorzhak* and *Mongush*. Vestnik Tuvinskogo Gosudarstvennogo Universiteta = Bulletin of Tuvan State University. 2018;2. (in Russian)

Dolgikh B.O. The Clan and Tribal Composition of Siberian Ethnic Groups in 17th Century. Moscow: AN SSSR Publ., 1960. (in Russian)

Ilumäe A.M., Reidla M., Chukhryaeva M., Järve M., Post H., Karmin M., Saag L., Agdzhoyan A., Kushniarevich A., Litvinov S., Ekomasova N., Tambets K., Metspalu E., Khusainova R., Yunusbayev B., Khusnutdinova E.K., Osipova L.P., Fedorova S., Utevska O., Koshel S., Balanovska E., Behar D.M., Balanovsky O., Kivisild T., Underhill P.A., Villems R., Rootsi S. Human Y chromosome haplogroup N: A non-trivial time-resolved phylogeography that cuts across language families. Am. J. Hum. Genet. 2016;99:163-173. DOI 10.1016/j.ajhg.2016.05.025.

Kharkov V.N., Khamina K.V., Medvedeva O.F., Simonova K.V., Eremina E.R., Stepanov V.A. Gene pool of Buryats: Clinal variability and territorial subdivision based on data of Y-chromosome markers. Russ. J. Genet. 2014;50(2):180-190. DOI 10.1134/S1022795413110082.

Kharkov V.N., Khamina K.V., Medvedeva O.F., Simonova K.V., Khitrinskaya I.Yu., Stepanov V.A. Genepool structure of Tuvans inferred from Y-chromosome marker data. Russ. J. Genet. 2013;49(12):1236-1244. DOI 10.1134/S102279541312003X.

Kharkov V.N., Medvedeva O.F., Luzina F.A., Kolbasko A.V., Gafarov N.I., Puzyrev V.P., Stepanov V.A. Comparative characteristics of the gene pool of Teleuts inferred from Y-chromosomal marker data. Russ. J. Genet. 2009; 45(8):994-1003. DOI 10.1134/S1022795409080158.

Klyashtorny S.G., Savinov D.G. Steppe Empires of Ancient Eurasia. St. Petersburg: Faculty of Philology of St. Petersburg State University, 2005. (in Russian)

Mannai-ool M.Kh. Tuvan People. The Origin and Formation of the Ethnos. Novosibirsk: Nauka Publ., 2004;99-166. (in Russan)

Nei M. Molecular Population Genetics and Evolution. Amsterdam: North-Holland Publ. Co., 1975.

Old Turkic Vocabulary. Leningrad: Nauka Publ., 1969;120-364. (in Russian)

Rootsi S., Zhivotovsky L.A., Baldovic M., Kayser M., Kutuev I.A., Khusainova R., Bermisheva M.A., Gubina M., Fedorova S.A., Ilumäe A.M., Khusnutdinova E.K., Voevoda M.I., Osipova L.P., Stoneking M., Lin A.A., Ferak V., Parik J., Kivisild T., Underhill P.A., Villems R. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. Eur. J. Hum. Genet. 2007;15(2):204-211. DOI 10.1038/si.eihg.5201748.

Serdobov N.A. The History of Formation of Tuvan People. Kyzyl: Tuvinskoye Knizhnoye Izdatelstvo, 1971. (in Russian)

Tatarintsev B.I. About the Some Tuvan Ethnonyms. The New Research of Tuva Archeology and Ethnogenesis of Tuvan People. Kyzyl: TNIIYALI Publ., 1980;144-145. (in Russan)

Tatarintsev B.I. The Problems of Investigation of Tuvan Ethnonymics (The Case of Some Ethnonyms of Presumed Mongol Origin). The Research of Tuvan Philology. Kyzyl, 1986;64-86. (in Russan)

The Results of population census in Russia in 2010. Avaible at http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm (in Russian)

The Secret Story. The Mongolian Chronicle of 1240 Under the Name Mongrol-un Nirucatobciyan. Yuan Chao Bi Shi (translated by S.A. Kozin). Mongolian Mundane Collection. Moscow; Leningrad, 1941. (in Russian)

Wells R.S., Yuldasheva N., Ruzibakiev R., Underhill P.A., Evseeva I., Blue-Smith J., Jin L., Su B., Pitchappan R., Shanmugalakshmi S., Balakrishnan K., Read M., Pearson N.M., Zerjal T., Webster M.T., Zholoshvili I., Jamarjashvili E., Gambarov S., Nikbin B., Dostiev A., Aknazarov O., Zalloua P., Tsoy I., Kitaev M., Mirrakhimov M., Chariev A., Bodmer W.F. The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc. Natl. Acad. Sci. USA. 2001;98(18):10244-10249. DOI 10.1073/pnas.171305098.

Zerjal T., Xue Y., Bertorelle G., Wells R.S., Bao W., Zhu S., Qamar R., Ayub Q., Mohyuddin A., Fu S., Li P., Yuldasheva N., Ruzibakiev R., Xu J., Shu Q., Du R., Yang H., Hurles M.E., Robinson E., Gerelsaikhan T., Dashnyam B., Mehdi S.Q., Tyler-Smith C. The genetic legacy of the Mongols. Am. J. Hum. Genet. 2003;72(3):717-721. DOI 10.1086/367774.

Zhabagin M., Balanovska E., Sabitov Zh., Kuznetsova M., Agdzhoyan A., Balaganskaya O., Chukhryaeva M., Markina N., Romanov A., Skhalyakho R., Zaporozhchenko V., Saroyants L., Dalimova D., Davletchurin D., Turdikulova Sh., Yusupov U., Tazhigulova I., Akilzhanova A., Tyler-Smith C., Balanovsky O. The connection of the genetic, cultural and geographic landscapes of Transoxiana. Sci. Rep. 2017;7(1):3085. DOI 10.1038/s41598-017-03176-z.

Zhabagin M.K., Dibirova Kh.D., Frolova S.A., Sabitov Z.M., Yusupov Y.M., Utevskaya O.M., Tarlikov P.V., Tazhigulova I.M., Balaganskaya O.A., Nimadava P., Zakharov I.A., Balanovsky O.P. The relation between the Y-chromosomal variation and the clan structure: the gene pool of the steppe

aristocracy and the steppe clergy of the Kazakhs. Vestnik Moskovskogo Universiteta. Seria XXIII. Antropologia = Vestnik of Moscow University. Ser. XXIII. Antropology. 2014;1:96-101. (in Russian)

Zhabagin M.K., Sabitov Z., Balaganskaya O., Bogunova A., Frolova S., Tazhigulova I., Nimadava P., Zakharov-Gezehus I.A., Balanovsky O.P. The possible role of social selection in the distribution of major haplotypes of Y-haplogroup C* in the people of Central Asia. 5th Int. conf. "Alexeev reading" on the memory of T.I. Alexeeva and V.P. Alexeev. Moscow, 2013. (in Russian)

Вавиловский журнал генетики и селекции. 2018;22(5):620-626
DOI 10.18699/VJ18.403

# Whole exome analysis of primary immunodeficiency

E.S. Rahmani[1], H. Azarpara[2], M. Karimipoor[3], H. Rahimi[3]

[1] Department of Medical Biotechnology, College of Allied Medicine, Iran University of Medical Sciences, Tehran, Iran
[2] Department of Medicine, Iran University of Medical Sciences, Tehran, Iran
[3] Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

The human primary immunodeficiency diseases (PIDs) refer to a rare heterogeneous group of single-gene inherited disorders causing malfunctions in the immune system, and thus the affected patients have a predisposition to severe life-threatening infections. The heterogeneous nature of PIDs, which involves at list 300 different genes, makes diagnosis of the disease a complex issue. Although studies revealed that six million people have a kind of PID, but due to a complex diagnosis procedure many affected individuals have not gotten a correct diagnosis. However, thanks to advancing in the DNA sequencing method and availability of sophisticated sequencers molecular characterization of genetic disorders have been revolutionized. The whole exome sequencing (WES) method can help clinicians detect Mendelian disease and other complex genetic disorders. The presented study used WES to investigate two infants with symptoms of primary immunodeficiency including hemophagocytic lymphohistiocytosis (HLH) and severe combined immunodeficiency (SCID). It has been shown that the HLH patient had a mutation in the UNC13D gene (NM_199242.2:c.627delT), and the SCID patient had a mutation in the RAG1 gene (NM_000448.2:c.322C>G). It has been demonstrated that WES is a fast and cost-effective method facilitating genetic diagnosis in PID sufferers.

Key words: primary immunodeficiency; next-generation sequencing; UNC13D; RAG1.

## Полноэкзомный анализ первичного иммунодефицита

Э.Ш. Рахмани[1], Х. Азарпара[2], М. Каримипур[3], Х. Рахими[3]

[1] Факультет медицинских биотехнологий, Школа комплементарной медицины, Иранский университет медицинских наук, Тегеран, Иран
[2] Факультет медицины, Иранский университет медицинских наук, Тегеран, Иран
[3] Факультет молекулярной медицины, Исследовательский центр биотехнологий, Институт Пастера Ирана, Тегеран, Иран

Первичный иммунодефицит представляет собой гетерогенную группу редких наследственных мутаций единичного гена, вызывающих сбой в работе иммунной системы человека и проявляющихся в предрасположенности пациентов к тяжелым жизнеугрожающим инфекциям. Гетерогенная природа первичного иммунодефицита, при которой мутации могут быть подвержены по меньшей мере 300 различных генов, серьезно осложняет его диагностику. И хотя было подсчитано, что от этого заболевания могут страдать около шести миллионов человек, только немногие из них могут рассчитывать на постановку правильного диагноза. Однако развитие методов секвенирования ДНК и доступность высокотехнологичного оборудования позволили сделать значительный шаг вперед в области молекулярных исследований генетических заболеваний. Технология полноэкзомного анализа ДНК может оказать существенную помощь врачам при диагностировании менделевских предрасположенностей к микробактериальным инфекциям и других форм редких генетических заболеваний. В представленном исследовании мы использовали метод полноэкзомного анализа ДНК для обследования двух младенцев с симптомами первичного иммунодефицита, такими как гемофагоцитарный лимфогистиоцитоз (ГЛГ) и тяжелый комбинированный иммунодефицит (ТКИД). Полноэкзомный анализ выявил мутацию UNC13D гена (NM_199242.2:c.627delT) у пациента с ГЛГ и мутацию RAG1 гена (NM_000448.2:c.322C>G) – у пациента с ТКИД. Исследование показало, что полноэкзомный анализ – это быстрый и экономичный метод, помогающий поставить правильный диагноз пациентам с первичным иммунодефицитом.

Ключевые слова: первичный иммунодефицит; секвенирование нового поколения; UNC13D; RAG1.

**HOW TO CITE THIS ARTICLE:**
Rahmani E.S., Azarpara H., Karimipoor M., Rahimi H. Whole exome analysis of primary immunodeficiency. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(5):620-626. DOI 10.18699/VJ18.403

Received 09.04.2018
Accepted for publication 28.06.2018

These authors contributed equally to this work.

e-mail: rahimi.h1981@gmail.com

Human primary immunodeficiencies (PIDs) include at least 300 genetically-defined single-gene inherited disorders causing malfunctions in the immune system (Bousfiha et al., 2015). Recent studies have revealed that six million people worldwide have a kind of PID, whereas only 27000–60000 individual have been diagnosed. Despite the fact that the prevalence of PIDs is highest among children, there are also many adult patients with PIDs (Bousfiha et al., 2013). Patients with PIDs are predisposed to severe infections such as the EBV, Neisseria, Papillomavirus, Streptococcus pneumoniae, weakly virulent mycobacteria, Herpes simplex virus, and Candida Albicans. Furthermore, the immune dysregulation and aberrant inflammatory responses (Abolhassani et al., 2014) such as Allergy, Angioedema, Hemophagocytosis, Autoinflammation, and Autoimmunity have also been diagnosed in such patients (Bousfiha et al., 2013). PIDs are classified into nine groups based on the clinical and laboratory parameters (Picard et al., 2015). The first group includes the most common combined T- and B-cell immunodeficiency (CID) and a severe PID form known as severe combined immunodeficiency (SCID) (Bonilla et al., 2005). SCID is characterized by profound defects of T-cell development and it affects some of the B and NK cells (Kwan, Puck, 2015). Hemophagocytic lymphohistiocytosis (HLH) is a heterogeneous group of disorders related to dysregulation of the immune system (PID class 4) that are classified into two groups, namely genetic (familial HLH (FHL)) and acquired forms based on the etiology. FHL includes five forms of loss-of-function mutation (FHL1-5) leading to defects in the cytotoxic granule secretion pathway in NK and $CD8^+$ T-cells, and they consequently lead to failure in exocytosis of granules in immunologic synopsis and completely eradicate target cells, in situation of immune response (Sifers et al., 2016). HLH can also occur due to infections and autoinflammatory/autoimmune and malignant diseases which are known as the acquired forms (Janka, 2012).

Current procedures for PID diagnosis are very complex and involve using specialized immunologic tests, including lymphocyte proliferation and cytotoxic assay, evaluation of serum immunoglobulin level, flow cytometry, neutrophil function assays and complementary analysis (McCusker, Warrington, 2011). However, immunological evaluation is performed to assess a patient's immune status for primary PID diagnosis. In addition, phenotype – based PID diagnosis is often complex, expensive and not always successful. Genetic investigation of PIDs is also very complex, and covers more than 300 genes that may be involved. Allelic heterogeneity and locus heterogeneity also increase the complexity of genetic analysis (Moens et al., 2014; Stoddard et al., 2014).

Advances in next-generation sequencing (NGS), particularly in whole exome sequencing (WES) have revolutionized molecular diagnosis of Mendelian disorders (Gilissen et al., 2012), and thus the traditional methods can be replaced by interrogation of a large set of genes in the single test in a timely and cost-effective manner instead of gene-by-gene approaches (Stoddard et al., 2014; Vrijenhoek et al., 2015). Thirty four new gene defects in PIDs have been diagnosed using NGS technology, so far (Conley, Casanova, 2014). The most recent research by S. Tamura et al. (2015) described the identification of novel compound heterozygous mutation in the DNA ligase IV (LIG4) gene through WES.

The present study was aimed at using WES for molecular characterization of two families with PID-affected children to be confirmed by the Sanger sequencing. We found a mutation in UNC13D and RAG1 genes in HLH and SCID families respectively. Our results have proved WES to be a useful method for detecting pathogenic variants in PID sufferers.

## Patients, materials, and methods

**Patients.** The patients selected for molecular characterization using the whole exome sequencing were two children from two Iranian families in consanguineous marriage hospitalized in Children's medical center (Tehran) and their PIDs were confirmed with specialized immunology methods. The first patient was a six-month female infant with clinical symptoms of HLH, and the second - a two-month male infant diagnosed with SCID. The HLH patient had prolonged fever, hepatosplenomegaly, an infection caused by Epstein–Barr virus (EBV) and increased levels of ferritin, while the SCID patient suffered of recurrent diarrhea, respiratory infections without any circulating T and B cells ($T^-$ $B^-$ SCID). For the children could participate in the study, the parents had signed an ethical consent form.

**Blood sampling.** Peripheral blood samples were taken from the children and their family members. The clinical information obtained from their medical records included date and year of diagnosis, disease class and severity, surgical history, medication and family history. Genomic DNA was extracted using salting-out protocol (Miller et al., 1988).

**Whole exome sequencing and data analysis.** One hundred ng/μl of high-quality genomic DNA was first used for whole exome enrichment in the IonAmpliSeq Exome RDY plates and Ion AmpliSeq HiFi Mix. After ligation of Proton adapters and quantification by qPCR, the final library was sequenced using the Ion Proton platform which produced raw FASTQ files at an average coverage depth of 50X.

The following general workflow was used for performing bioinformatics analysis of the FASTQ raw data to prioritize causative variants (Fig. 1).

The raw FASTQ files were processed using the NGS QC toolkit (Patel, Jain, 2012) to estimate the quality and states of sequence reads. Sequencing reads, which have some errors such as adaptor and primer contamination, low quality 5′ and 3′ end bases, short reads and those with quality scores (Phered score) bellow 20, were trimmed from the FASTQ files using the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). BWA-MEM algorithm (Caboche et al., 2014) was used to align reads against the reference genome sequence (GRCh37), and then the results were stored in the SAM (Sequence Alignment/Map) file format. BWA-MEM is fast and accurate alignment software for nucleotide sequences of about 70 bp–1 Mbp. Duplicate reads were marked by Picard tools (https://broadinstitute.github.io/picard/). Afterwards, the SNP and InDel calling was performed with the Genome Analysis Toolkit (GATK; v 3.6) (Van der Auwera et al., 2013). The functional variant annotation was performed using the following software and databases: ANNOVAR (Wang et al., 2010), KGGSeq (v1.0) (Li et al., 2012), the 1000-Genome Project (www.1000genomes.org/), NHLBI GO Exome Sequencing Project (ESP), Exome Variant Server (EVS) (http://evs.gs.washington.edu/EVS/), Exome Aggregation Consortium

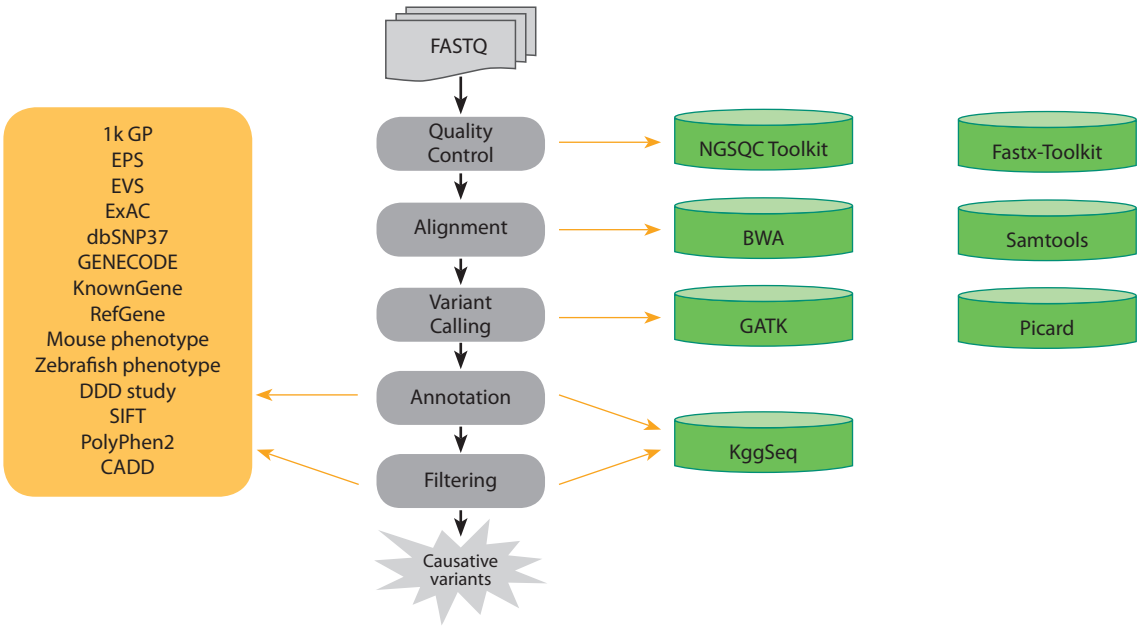



**Fig. 1.** WES data analysis flowchart.

The green circles denote software, while the orange box shows all the databases and algorithms provided by KGGSeq that were used for annotation and filtering.

**Table 1.** PCR primers

| Sample | Chr | Pos | Ref | Alt | Primer-Forward | Tm, °C | Primer-Reverse | Tm, °C | Product length, bp |
|---|---|---|---|---|---|---|---|---|---|
| HLH | 17 | 73836899 | A | – | AGGTATGGGAAGGGAAGGGATC | 60.3 | GCACCCCAGCATCCAGTGTG | 63.1 | 257 |
| SCID | 11 | 36595176 | C/G | T | GGACTTGTTTTCATTGTTCTCAG | 54.9 | CGAGTCAACATCTGCCTTCAC | 58.1 | 544 |

Chr – chromosome; Pos – position; Ref – reference sequence; Alt – alternative sequence.

(ExAC) (http://exac.broadinstitute.org; release 0.3), dbSNP, GENECODE (Harrow et al., 2012), knownGene, RefGene (UCSC), mouse phenotype (Eppig et al., 2015) and DDD study (Deciphering Developmental Disorders) (Firth et al., 2009). The KGGSeq filtering strategy was executed to filter out both common benign variants and recurrent artifact as well as to find causal variants. At the first level, variant quality control was checked using the KGGSeq software, which includes various filters such as genotype QC, variants QC and sample QC to filter out errors and low-quality variants. A cut-off of 20, 50 and 20 was done for variants' Phred quality score, mapping quality score and depth coverage respectively. SIFT (Ng, Henikoff, 2003), PolyPhen-2 (Adzhubei et al., 2010) and CADD (Kircher et al., 2014) tools were used for predicting and scoring the possible impact of amino acid substitution on the structure and function of human proteins. Prioritization was performed with the focus on a primary immunodeficiency panel, which was gathered from the NCBI gene database (http://www.mcbi.nlm.nih.gov/gene/) and recommended genes from National Immunology Society resulting in 400 genes and genomic regions. The causative variants were limited to following criteria: 1) a minor allele frequency (MAF) of less than 0.1 % in data from the 1000-Genome Project, EVS, and ExAC, 2) minimum CADD score of 20, and 3) high or moderate effect determined by SNPEff (v 4.2) (Cingolani et al., 2012).

**Sanger sequencing validation.** The chromosomal regions containing the candidate causative variants in patients were amplified by polymerase chain reaction (Eppendorf, Germany) with Taq DNA polymerase (Amplicon, Germany) and designed primer for each of them at PCR conditions as summarized in Table 1. The sequencing diagrams were obtained using the Chromas software (v 2.6) (http://technelysium.com.au).

## Results

Molecular diagnosis of PIDs by means of whole exome sequencing is a cost-effective and valuable approach for patients. In the present project, WES was used to study two families with HLH and SCID symptoms respectively. After performing the sequencing, the total reads were about 44 and 55 million for HHL and SCID patients respectively (Table 2). QC analysis of the FASTQ files showed that the average length of reads consisted of 190 nucleotides and 51 % of GC content, and more than 86 present of reads had quality scores equal or more than the cut-off (Q20) (see Table 2). After QC checking and trimming, the trimmed files were mapped to human reference genome version of GRCh37. Alignment statistics report using Samtools (Li et al., 2009) indicated that 95 %

of the reads were mapped to reference genome in the both files (see Table 2). 49426 variants were observed in the HLH patient, and 49629 variants – in the SCID patient after variant calling (see Table 2). After using the KGGSeq filters, the observed variations for the two patients were 48838 (HLH) and 49181 (SCID) respectively. Allele frequency trimming performed using 1000G, ESP6500 and ExAc indicated that 30569 and 29968 variants had greater than 1 % of MAF in the HLH and SCID patients respectively. Functional effect prediction of variants was performed using the SIFT, Polyphen-2 and CADD software resulting in 80 pathogenic and deleterious variants in the both patients. Top 10 variants with higher scores are presented in Table 3. Finally, two variants were considered as potential causative variants for the abovementioned phenotype of HLH and SCID. Selected variants included a frameshift variant in Exon 8 of the UNC13D gene (NM_199242.2:c.627delT) and a premature stop codon variant in Exon 2 of the RAG1 gene (NM_000448.2:c.322C>G) in the HLH and SCID patient respectively (Table 4). Sanger sequencing confirmed the both variants (Fig. 2).

**Table 2.** Quality control and alignment statistics

| Parameter | HLH | SCID |
|---|---|---|
| Total reads (Mb) | 44.509 | 55.250 |
| Total bases (Gb) | 8.457 | 10.446 |
| Average length of read (bp) | 190 | 189 |
| Q20 | 86.57 % | 86.15 % |
| % GC | 51.51 % | 51.92 % |
| Mapped reads | 44354714 | 55004005 |
| Total aligned base reads | 8307987731 | 10246246706 |
| Percent reads on target | 95.12 % | 94.72 % |
| Percent base reads on target | 94.22 % | 93.94 % |
| Total variant count | 49248 | 49456 |
| SNV | 46455 | 46670 |
| INDEL | 2793 | 2786 |

Q20 – percent of base number calls with quality value of 20 or higher; % GC – percentage of GC content; SNV – single nucleotide variant; INDEL – insertion and deletion; CNV – copy number variation.

**Table 3.** Top 10 variants with high scores for both patients

| Chr | Pos | Allele | Gene | Type | Score | Causality |
|---|---|---|---|---|---|---|
| | | | HLH | | | |
| 19 | 57326850 | T/C | PEG3 | Missense | 0.9967 | Damaging |
| 17 | 61601686 | G/T | KCNH6 | » | 0.9518 | » |
| 16 | 2577847 | C/G | AMDHD2 | » | 0.9176 | » |
| 12 | 121176083 | G/A | ACADS | » | 0.9158 | » |
| 2 | 189918622 | G/A | COL5A2 | » | 0.8126 | » |
| 1 | 2938989 | G/A | ACTRT2 | » | 0.7911 | » |
| 17 | 21319121 | C/T | KCNJ12 | » | 0.7460 | » |
| 10 | 14862082 | C/G | CDNF | » | 0.7220 | » |
| 11 | 5758062 | T/C | OR56B1 | » | 0.6986 | » |
| 14 | 20692643 | T/C | OR11H6 | » | 0.6830 | » |
| | | | SCID | | | |
| 17 | 61601686 | G/T | KCNH6 | » | 0.9518 | » |
| 12 | 121176083 | G/A | ACADS | » | 0.9158 | » |
| 18 | 61387333 | T/A | SERPINB11 | » | 0.9145 | » |
| 1 | 94466628 | G/A | ABCA4 | » | 0.8452 | » |
| 10 | 54531235 | C/T | MBL2 | » | 0.7919 | » |
| 16 | 84456014 | C/T | ATP2C2 | » | 0.7698 | » |
| 7 | 158591753 | G/A | ESYT2 | » | 0.7685 | » |
| 1 | 11856378 | G/A | MTHFR | » | 0.7582 | » |
| 14 | 24707479 | G/A | GMPR2 | » | 0.7563 | » |
| 17 | 21319121 | C/T | KCNJ12 | » | 0.7460 | » |

**Table 4.** Potential causative variants

| Sample | Chromosome | Position | Damaging variant | Reference allele | Alternative allele | Type | rsID |
|---|---|---|---|---|---|---|---|
| HLH | 17 | 73836899 | Exon 8:c.627 del T | A | – | Homozygous | rs755619812 |
| SCID | 11 | 36595176 | Exon 2: c.322 C>T | C/G | T | Homozygous | rs193922464 |

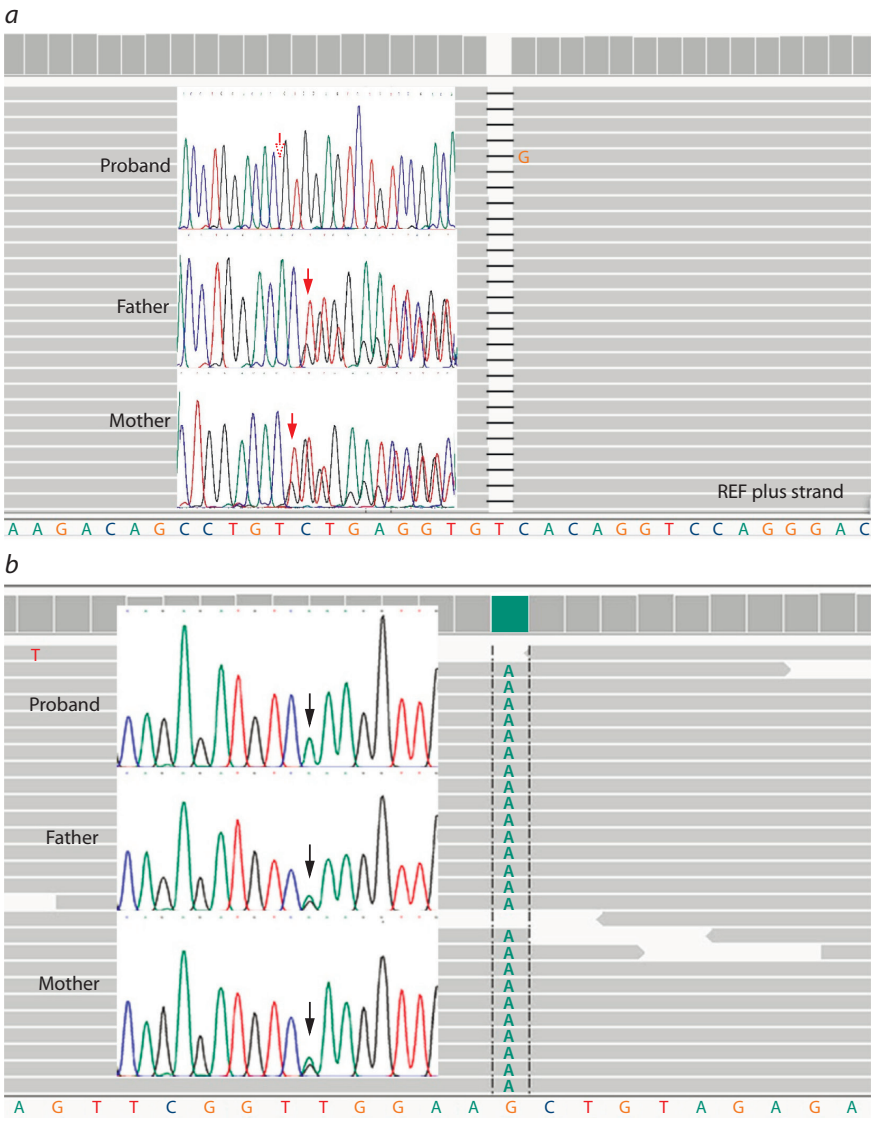



**Fig. 2.** Sanger sequencing confirmation of considered variants in HLH and SCID patients:

*a* – in the HLH patient, IGV screenshot shows a homozygous deletion in the UNC13D gene confirmed both in the proband and the parents; *b* – IGV screenshot from the SCID patient also shows a homozygous single nucleotide substitution in the RAG1 gene confirmed both in the proband and the parents.

## Discussion

Molecular diagnosis of PIDs with at least 300 genes, allelic and locus heterogeneity is a challenging issue. Recent studies indicate that 121 gene defects have been identified in these diseases in addition to the list of genes involved in PID (179 genes) since 2011 (Al-Herz et al., 2011; Picard et al., 2015). Utilization of the Sanger sequencing as a direct sequencing method for finding mutations in genes has become a de facto issue in genetic diagnosis (Sanger et al., 1977). Despite the fact that the gene-by-gene approach is a conventional diagnostic test for monogenic diseases, such methods are too expensive and time-consuming in multigenic diseases such as the PIDs (Chou et al., 2012; Sikkema-Raddatz et al., 2013). In contrast, WES method has become a favorable test for genetic diagnosis of multigenic diseases and it sequences all coding regions using amplification and parallel sequencing in a single test (Sikkema-Raddatz et al., 2013; Stoddard et al., 2014). For example, E. Mukda et al. (2017) evaluated 25 HLH patients and diagnosed pathogenic mutations in PFR1, UNC13D, STXBP2, LYST and XIAP genes. G.S. Schulert et al. (2015) also conducted WES in 16 HLH patients and diagnosed the disease-causing mutations in PFR1 and LYST gene. CARD11 gene inactivation due to a premature stop codon was diagnosed in a SCID patient by means of WES (Greil et al., 2013). Accordingly, the present study sought to obtain a molecular diagnosis of two patients with HLH and SCID by means of WES. A variant (rs755619812) in UNC13D gene and a variant (rs193922464) in RAG1 gene were detected in HLH and SCID patients respectively after exome sequencing and processing the raw data. In HLH case, the UNC13D gene (17q25.1) comprised 32 exons and 4.5-kb transcript with the highest expression in spleen, thymus, and peripheral blood leukocytes. In lymphocytes, the encoded protein (Munc13-4, 1090 amino acids, UniProt ID: Q70J99) played an important role in cytotoxic granule exocytosis. Munc13-4 is an essential protein for maturation, docking, and priming of cytotoxic granules in cytotoxic cells (CTLs and NK) (Feldmann et al., 2003). Pathogenic mutations in the UNC13D gene led to an ineffective protein create type 3 of familial HLH (FHL-3). We found a homozygous nucleotide deletion (c.627delT:p.V210Wfs*39) in exon 8 of UNC13D, and it caused a frameshift mutation (see Fig. 2, *a*). This damaging mutation led to the substitution of valine with tryptophan at position 210 resulting in a premature stop codon at 39 codons after this substitution which produced a truncated protein. This mutation results in defects in the killing ability of NK and $CD8^+$ T-cells and uncontrolled hyperinflammation in HLH patients. In 2006, Stasdt et al. studied 63 HLH patients and described the c.627delT mutation in two infants in order to find the mutations spectra of the PFR1, UNC13D, STX11 and RAB27A genes (Stadt et al., 2006).

In the second patient with clinical symptoms of SCID, we found a homozygous single nucleotide substitution in the first position of Arginine codon 108 in Exon 2 of the RAG1 gene (c.322C>G: R108*) leading to a premature stop codon and consequently a truncated protein (see Fig. 2, b). The human RAG1 gene (11p12) consists of 2 exons and encodes a protein with 1043 amino acids (UniProt ID: P15918). Recombination-activating protein 1 (RAG1) is a catalytic component of the RAG complex as a multi-protein complex that mediates the DNA cleavage phase during V(D)J recombination. In the RAG complex (RAG1/2), RAG1 mediates the DNA-binding to the conserved recombination

signal sequences (RSS) and catalyzes DNA cleavage activities by introducing a double-strand break between RSS and the adjacency to each coding V, D, and J DNA segment (Melek, Gellert, 2000). This somatic recombination leads to the diversity of immunoglobulins and T-cell receptors (TCRs). The pathogenic mutations affecting the active core (amino acids 384-1009) of RAG1 produce the clinical symptoms of SCID disease such as absence of lymphocyte (B- and T-cells) circulation (Corneo et al., 2001) observed in our patient.

## Conclusion

Whole exome sequencing has proved itself as a fast and cost-effective method for detection of causative rare mutations in PID patients.

## Acknowledgements

We would like to thank the families who participated in this project as well as Pasteur Institute of Iran for supporting this project (grant No. 756).

## Conflict of interest

The authors declare they have no conflict of interest.

## References

Abolhassani H., Wang N., Aghamohammadi A., Rezaei N., Lee Y.N., Frugoni F., Notarangelo L.D., Pan-Hammarström Q., Hammarström L. A hypomorphic recombination-activating gene 1 (RAG1) mutation resulting in a phenotype resembling common variable immunodeficiency. J. Allergy Clin. Immunol. 2014;134(6):1375-1380. DOI 10.1016/j.jaci.2014.04.042.

Adzhubei I.A., Schmidt S., Peshkin L., Ramensky V.E., Gerasimova A., Bork P., Kondrashov A.S., Sunyaev S.R. A method and server for predicting damaging missense mutations. Nat. Methods. 2010; 7(4):248-249. DOI 10.1038/nmeth0410-248.

Al-Herz W., Bousfiha A., Casanova J.-L., Chapel H., Conley M.E., Cunningham-Rundles C., Etzioni A., Fischer A., Franco J.L., Geha R.S., Hammarström L., Nonoyama S., Notarangelo L.D., Ochs H.D., Puck J.M., Roifman C.M., Seger R., Tang M.L. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. Front. Immunol. 2011;2:54. DOI 10.1007/s10875-015-0201-1.

Bonilla F.A., Bernstein I.L., Khan D.A., Ballas Z.K., Chinen J., Frank M.M., Kobrynski L.J., Levinson A.I., Mazer B., Nelson R.P., Orange J.S., Routes J.M., Shearer W.T., Sorensen R.U. Practice parameter for the diagnosis and management of primary immunodeficiency. Ann. Allergy Asthma Immunol. 2005;94(5):S1-S63. DOI 10.1016/S1081-1206(10)61142-8.

Bousfiha A.A., Jeddane L., Ailal F., Benhsaien I., Mahlaoui N., Casanova J.L., Abel L. Primary immunodeficiency diseases worldwide: more common than generally thought. J. Clin. Immunol. 2013;33(1): 1-7. DOI 10.1007/s10875-012-9751-7.

Bousfiha A., Jeddane L., Al-Herz W., Ailal F., Casanova J.L., Chatila T., Conley M.E., Cunningham-Rundles C., Etzioni A., Franco J.L., Gaspar H.B., Holland S.M., Klein C., Nonoyama S., Ochs H.D., Oksenhendler E., Picard C., Puck J.M., Sullivan K.E., Tang M. The 2015 IUIS phenotypic classification for primary immunodeficiencies. J. Clin. Immunol. 2015;35(8):727-738. DOI 10.1007/s10875-015-0198-5.

Caboche S., Audebert C., Lemoine Y., Hot D. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. BMC Genomics. 2014;15(1):1. DOI 10.1186/1471-2164-15-264.

Chou J., Ohsumi T.K., Geha R.S. Use of whole exome and genome sequencing in the identification of genetic causes of primary immunodeficiencies. Curr. Opin. Allergy Clin. Immunol. 2012;12(6):623-628. DOI 10.1097/ACI.0b013e3283588ca6.

Cingolani P., Platts A., Wang L.L., Coon M., Nguyen T., Wang L., Land S.J., Lu X., Ruden D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain $w^{1118}$; iso-2; iso-3. Fly. 2012;6(2):80-92. DOI 10.4161/fly.19695.

Conley M.E., Casanova J.-L. Discovery of single-gene inborn errors of immunity by next generation sequencing. Curr. Opin. Immunol. 2014;30:17-23. DOI 10.1016/j.coi.2014.05.004.

Corneo B., Moshous D., Güngör T., Wulffraat N., Philippet P., Le Deist F.L., Fischer A., de Villartay J.P. Identical mutations in RAG1 or RAG2 genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome. Blood. 2001;97(9):2772-2776. DOI 10.1182/blood.V97.9.2772.

Eppig J.T., Blake J.A., Bult C.J., Kadin J.A., Richardson J.E. Mouse Genome Database Group. The Mouse Genome Database (MGD): Facilitating mouse as a model for human biology and disease. Nucleic Acids Res. 2015;43:D726-D736. DOI 10.1093/nar/gku967.

Feldmann J., Callebaut I., Raposo G., Certain S., Bacq D., Dumont C., Lambert N., Ouachée-Chardin M., Chedeville G., Tamary H., Minard-Colin V. Munc13-4 is essential for cytolytic granules fusion and is mutated in a form of familial hemophagocytic lymphohistiocytosis (FHL3). Cell. 2003;115(4):461-473. DOI 10.1016/S0092-8674(03)00855-9.

Firth H.V., Richards S.M., Bevan A.P., Clayton S., Corpas M., Rajan D., Van Vooren S., Moreau Y., Pettett R.M., Carter N.P. DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. Am. J. Hum. Genet. 2009;84(4):524-533. DOI 10.1016/j.ajhg.2009.03.010.

Gilissen C., Hoischen A., Brunner H.G., Veltman J.A. Disease gene identification strategies for exome sequencing. Eur. J. Hum. Genet. 2012;20(5):490-497. DOI 10.1038/ejhg.2011.258.

Greil J., Rausch T., Giese T., Bandapalli O.R., Daniel V., Bekeredjian-Ding I., Stütz A.M., Drees C., Roth S., Ruland J., Korbel J.O., Kulozik A.E. Whole-exome sequencing links caspase recruitment domain 11 (CARD11) inactivation to severe combined immunodeficiency. J. Allergy Clin. Immunol. 2013;131(5):1376-1383. e1373. DOI 10.1016/j.jaci.2013.02.012.

Harrow J., Frankish A., Gonzalez J.M., Tapanari E., Diekhans M., Kokocinski F., Aken B.L., Barrell D., Zadissa A., Searle S., Barnes I., Bignell A., Boychenko V., Hunt T., Kay M., Mukherjee G., Rajan J., Despacio-Reyes G., Saunders G., Steward C., Harte R., Lin M., Howald C., Tanzer A., Derrien T., Chrast J., Walters N., Balasubramanian S., Pei B., Tress M., Rodriguez J.M., Ezkurdia I., van Baren J., Brent M., Haussler D., Kellis M., Valencia A., Reymond A., Gerstein M., Guigó R., Hubbard T.J. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012;22(9):1760-1774. DOI 10.1101/gr.135350.111.

Janka G. Familial and acquired hemophagocytic lymphohistiocytosis. Annu. Rev. Med. 2012;63:233-246. DOI 10.1146/annurev-med-041610-134208.

Kircher M., Witten D.M., Jain P., O'Roak B.J., Cooper G.M., Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 2014;46(3):310. DOI 10.1038/ng.2892.

Kwan A., Puck J.M. Newborn screening for severe combined immunodeficiency. Curr. Pediatr. Rep. 2015;3(1):34-42. DOI 10.1007/s40124-014-0068-2.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079. DOI 10.1093/bioinformatics/btp352.

Li M.-X., Gui H.-S., Kwan J.S., Bao S.-Y., Sham P.C. A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. Nucleic Acids Res. 2012;gkr1257.

McCusker C., Warrington R. Primary immunodeficiency. Allergy Asthma Clin. Immunol. 2011;7(1):1. DOI 10.1186/1710-1492-7-S1-S11.

Melek M., Gellert M. RAG1/2-mediated resolution of transposition intermediates: two pathways and possible consequences. Cell. 2000; 101(6):625-633. DOI 10.1016/S0092-8674(00)80874-0.

Miller S.A., Dykes D.D., Polesky H.F. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215. DOI 10.1093/nar/16.3.1215.

Moens L.N., Falk-Sörqvist E., Asplund A.C., Bernatowska E., Smith C.E., Nilsson M. Diagnostics of primary immunodeficiency diseases: a sequencing capture approach. PloS One. 2014;9(12): e114901. DOI 10.1371/journal.pone.0114901.

Mukda E., Trachoo O., Pasomsub E., Tiyasirichokchai R., Iemwimangsa N., Sosothikul D., Chantratita W., Pakakasama S. Exome sequencing for simultaneous mutation screening in children with hemophagocytic lymphohistiocytosis. Int. J. Hematol. 2017;1-9. DOI 10.1007/s12185-017-2223-3.

Ng P.C., Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31(13):3812-3814. DOI 10.1093/nar/gkg509.

Patel R.K., Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PloS One. 2012;7(2):e30619. DOI 10.1371/journal.pone.0030619.

Picard C., Al-Herz W., Bousfiha A., Casanova J.-L., Chatila T., Conley M.E., Cunningham-Rundles C., Etzioni A., Franco J.L., Gaspar H.B., Holland S.M., Klein C., Nonoyama S., Ochs H.D., Oksenhendler E., Picard C., Puck J.M., Sullivan K., Tang M.L. Primary immunodeficiency diseases: an update on the classification from the International Union of Immunological Societies Expert Committee for primary immunodeficiency 2015. J. Clin. Immunol. 2015;35(8): 696-726. DOI 10.1007/s10875-015-0201-1.

Sanger F., Nicklen S., Coulson A.R. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci USA. 1977;74(12):5463-5467. DOI 10.1073/pnas.74.12.5463.

Schulert G.S., Zhang M., Fall N., Husami A., Kissell D., Hanosh A., Zhang K., Davis K., Jentzen J.M., Napolitano L., Siddiqui J., Smith L.B., Harms P.W., Grom A.A., Cron R.Q. Whole-exome sequencing reveals mutations in genes linked to hemophagocytic lymphohistiocytosis and macrophage activation syndrome in fatal cases of H1N1 influenza. J. Infect. Dis. 2015;213(7):1180-1188. DOI 10.1093/infdis/jiv550.

Sifers T.M., Raje N., Dinakar C. Hemophagocytic lymphohistiocytosis: A concise review for the practicing physician. Allergy Asthma Proc. 2016;37(3):256-258. DOI 10.2500/aap.2016.37.3948.

Sikkema-Raddatz B., Johansson L.F., Boer E.N., Almomani R., Boven L.G., van den Berg M.P., van Spaendonck-Zwarts K.Y., van Tintelen J.P., Sijmons R.H., Jongbloed J.D., Sinke R.J. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. Hum. Mutat. 2013;34(7):1035-1042. DOI 10.1002/humu.22332.

Stadt U.Z., Beutel K., Kolberg S., Schneppenheim R., Kabisch H., Janka G., Hennies H.C. Mutation spectrum in children with primary hemophagocytic lymphohistiocytosis: molecular and functional analyses of PRF1, UNC13D, STX11, and RAB27A. Hum. Mutat. 2006;27(1):62-68. DOI 10.1002/humu.20274.

Stoddard J.L., Niemela J.E., Fleisher T.A., Rosenzweig S.D. Targeted NGS: a cost-effective approach to molecular diagnosis of PIDs. Front. Immunol. 2014;5:531. DOI 10.3389/fimmu.2014.00531.

Tamura S., Higuchi K., Tamaki M., Inoue C., Awazawa R. Mitsuki N., Nakazawa Y., Mishima H., Takahashi K., Kondo O., Imai K., Morio T., Ohara O., Ogi T., Furukawa F., Inoue M., Yoshiura K., Kanazawa N. Novel compound heterozygous DNA ligase IV mutations in an adolescent with a slowly-progressing radiosensitive-severe combined immunodeficiency. Clin. Immunol. 2015;160(2):255-260. DOI 10.1016/j.clim.2015.07.004.

Van der Auwera G.A., Carneiro M.O., Hartl C., Poplin R., del Angel G., Levy-Moonshine A., Jordan T., Shakir K., Roazen D., Thibault J., Banks E., Garimella K.V., Altshuler D., Gabriel S., DePristo M.A. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr. Protoc. Bioinformatics. 2013;11:11.10.11-11.10.33. DOI 10.1002/0471250953.bi1110s43.

Vrijenhoek T., Kraaijeveld K., Elferink M., De Ligt J., Kranendonk E., Santen G., Nijman I.J., Butler D., Claes G., Costessi A., Dorlijn W., van Eyndhoven W., Halley D.J., van den Hout M.C., van Hove S., Johansson L.F., Jongbloed J.D., Kamps R., Kockx C.E., de Koning B., Kriek M., Lekanne Dit Deprez R., Lunstroo H., Mannens M., Mook O.R., Nelen M., Ploem C., Rijnen M., Saris J.J., Sinke R., Sistermans E., van Slegtenhorst M., Sleutels F., van der Stoep N., van Tienhoven M., Vermaat M., Vogel M., Waisfisz Q., Marjan Weiss J., van den Wijngaard A., van Workum W., Ijntema H., van der Zwaag B., van IJcken W.F., den Dunnen J., Veltman J.A., Hennekam R., Cuppen E. Next-generation sequencing-based genome diagnostics across clinical genetics centers: implementation choices and their effects. Eur. J. Hum. Genet. 2015;23(9):1142-1150. DOI 10.1038/ejhg.2014.279.

Wang K., Li M., Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164. DOI 10.1093/nar/gkq603.