

Научный рецензируемый журнал

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.

Периодичность 8 выпусков в год

DOI 10.18699/VJGB-23-12

Учредители

Сибирское отделение Российской академии наук

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

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

Главный редактор

А.В. Кочетов – академик РАН, д-р биол. наук (Россия)

Заместители главного редактора

Н.А. Колчанов – академик РАН, д-р биол. наук, профессор (Россия)

И.Н. Леонова – д-р биол. наук (Россия)

Н.Б. Рубцов – д-р биол. наук, профессор (Россия)

В.К. Шумный – академик РАН, д-р биол. наук, профессор (Россия)

Ответственный секретарь

Г.В. Орлова – канд. биол. наук (Россия)

Редакционная коллегия

Е.Е. Андронов – канд. биол. наук (Россия)

Ю.С. Аульченко – д-р биол. наук (Россия)

О.С. Афанасенко – академик РАН, д-р биол. наук (Россия)

Д.А. Афонников – канд. биол. наук, доцент (Россия)

Л.И. Афтанас – академик РАН, д-р мед. наук (Россия)

Л.А. Беспалова – академик РАН, д-р с.-х. наук (Россия)

А. Бёрнер – д-р наук (Германия)

Н.П. Бондарь – канд. биол. наук (Россия)

С.А. Боринская – д-р биол. наук (Россия)

П.М. Бородин – д-р биол. наук, проф. (Россия)

А.В. Васильев – чл.-кор. РАН, д-р биол. наук (Россия)

М.И. Воевода – академик РАН, д-р мед. наук (Россия)

Т.А. Гавриленко – д-р биол. наук (Россия)

И. Гроссе – д-р наук, проф. (Германия)

Н.Е. Грунтенко – д-р биол. наук (Россия)

С.А. Демаков – д-р биол. наук (Россия)

И.К. Захаров – д-р биол. наук, проф. (Россия)

И.А. Захаров-Гезехус – чл.-кор. РАН, д-р биол. наук (Россия)

С.Г. Инге-Вечтомов – академик РАН, д-р биол. наук (Россия)

А.В. Кильчевский – чл.-кор. НАНБ, д-р биол. наук (Беларусь)

С.В. Костров – чл.-кор. РАН, д-р хим. наук (Россия)

А.М. Кудрявцев – чл.-кор. РАН, д-р биол. наук (Россия)

И.Н. Лаврик – д-р биол. наук (Германия)

Д.М. Ларкин – канд. биол. наук (Великобритания)

Ж. Ле Гуи – д-р наук (Франция)

И.Н. Лебедев – д-р биол. наук, проф. (Россия)

Л.А. Лутова – д-р биол. наук, проф. (Россия)

Б. Люгтенберг – д-р наук, проф. (Нидерланды)

В.Ю. Макеев – чл.-кор. РАН, д-р физ.-мат. наук (Россия)

В.И. Молодин – академик РАН, д-р ист. наук (Россия)

М.П. Мошкин – д-р биол. наук, проф. (Россия)

С.Р. Мурсалимов – канд. биол. наук (Россия)

Л.Ю. Новикова – д-р с.-х. наук (Россия)

Е.К. Потокина – д-р биол. наук (Россия)

В.П. Пузырев – академик РАН, д-р мед. наук (Россия)

Д.В. Пышный – чл.-кор. РАН, д-р хим. наук (Россия)

И.Б. Розозин – канд. биол. наук (США)

А.О. Рувинский – д-р биол. наук, проф. (Австралия)

Е.Ю. Рыкова – д-р биол. наук (Россия)

Е.А. Салина – д-р биол. наук, проф. (Россия)

В.А. Степанов – академик РАН, д-р биол. наук (Россия)

И.А. Тихонович – академик РАН, д-р биол. наук (Россия)

Е.К. Хлесткина – д-р биол. наук, проф. РАН (Россия)

Э.К. Хуснутдинова – д-р биол. наук, проф. (Россия)

М. Чен – д-р биол. наук (Китайская Народная Республика)

Ю.Н. Шавруков – д-р биол. наук (Австралия)

Р.И. Шейко – чл.-кор. НАНБ, д-р с.-х. наук (Беларусь)

С.В. Шестаков – академик РАН, д-р биол. наук (Россия)

Н.К. Янковский – академик РАН, д-р биол. наук (Россия)

Scientific Peer Reviewed Journal

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

*Founded in 1997**Published 8 times annually*

DOI 10.18699/VJGB-23-12

Founders

Siberian Branch of the Russian Academy of Sciences

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

The Vavilov Society of Geneticists and Breeders

Editor-in-Chief

A.V. Kochetov, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia

Deputy Editor-in-Chief

N.A. Kolchanov, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia

I.N. Leonova, Dr. Sci. (Biology), Russia

N.B. Rubtsov, Professor, Dr. Sci. (Biology), Russia

V.K. Shumny, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia

Executive Secretary

G.V. Orlova, Cand. Sci. (Biology), Russia

Editorial board

O.S. Afanasenko, Full Member of the RAS, Dr. Sci. (Biology), Russia

D.A. Afonnikov, Associate Professor, Cand. Sci. (Biology), Russia

L.I. Aftanas, Full Member of the RAS, Dr. Sci. (Medicine), Russia

E.E. Andronov, Cand. Sci. (Biology), Russia

Yu.S. Aulchenko, Dr. Sci. (Biology), Russia

L.A. Bespalova, Full Member of the RAS, Dr. Sci. (Agricul.), Russia

N.P. Bondar, Cand. Sci. (Biology), Russia

S.A. Borinskaya, Dr. Sci. (Biology), Russia

P.M. Borodin, Professor, Dr. Sci. (Biology), Russia

A. Börner, Dr. Sci., Germany

M. Chen, Dr. Sci. (Biology), People's Republic of China

S.A. Demakov, Dr. Sci. (Biology), Russia

T.A. Gavrilenko, Dr. Sci. (Biology), Russia

I. Grosse, Professor, Dr. Sci., Germany

N.E. Gruntenko, Dr. Sci. (Biology), Russia

S.G. Inge-Vechtomov, Full Member of the RAS, Dr. Sci. (Biology), Russia

E.K. Khlestkina, Professor of the RAS, Dr. Sci. (Biology), Russia

E.K. Khusnutdinova, Professor, Dr. Sci. (Biology), Russia

A.V. Kilchevsky, Corr. Member of the NAS of Belarus, Dr. Sci. (Biology),
Belarus

S.V. Kostrov, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia

A.M. Kudryavtsev, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

D.M. Larkin, Cand. Sci. (Biology), Great Britain

I.N. Lavrik, Dr. Sci. (Biology), Germany

J. Le Gouis, Dr. Sci., France

I.N. Lebedev, Professor, Dr. Sci. (Biology), Russia

B. Lugtenberg, Professor, Dr. Sci., Netherlands

L.A. Lutova, Professor, Dr. Sci. (Biology), Russia

V.Yu. Makeev, Corr. Member of the RAS, Dr. Sci. (Physics and Mathem.),
Russia

V.I. Molodin, Full Member of the RAS, Dr. Sci. (History), Russia

M.P. Moshkin, Professor, Dr. Sci. (Biology), Russia

S.R. Mursalimov, Cand. Sci. (Biology), Russia

L.Yu. Novikova, Dr. Sci. (Agricul.), Russia

E.K. Potokina, Dr. Sci. (Biology), Russia

V.P. Puzyrev, Full Member of the RAS, Dr. Sci. (Medicine),
RussiaD.V. Pyshnyi, Corr. Member of the RAS, Dr. Sci. (Chemistry),
Russia

I.B. Rogozin, Cand. Sci. (Biology), United States

A.O. Ruvinsky, Professor, Dr. Sci. (Biology), Australia

E.Y. Rykova, Dr. Sci. (Biology), Russia

E.A. Salina, Professor, Dr. Sci. (Biology), Russia

Y.N. Shavrukov, Dr. Sci. (Biology), Australia

R.I. Sheiko, Corr. Member of the NAS of Belarus,
Dr. Sci. (Agricul.), BelarusS.V. Shestakov, Full Member of the RAS, Dr. Sci. (Biology),
RussiaV.A. Stepanov, Full Member of the RAS, Dr. Sci. (Biology),
RussiaI.A. Tikhonovich, Full Member of the RAS, Dr. Sci. (Biology),
Russia

A.V. Vasiliev, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

M.I. Voevoda, Full Member of the RAS, Dr. Sci. (Medicine),
RussiaN.K. Yankovsky, Full Member of the RAS, Dr. Sci. (Biology),
Russia

I.K. Zakharov, Professor, Dr. Sci. (Biology), Russia

I.A. Zakharov-Gezekhus, Corr. Member of the RAS,
Dr. Sci. (Biology), Russia

Молекулярная и клеточная биология

- 93 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Обработка глифосатом приводит к накоплению в клетках растений малых дискретных 5'- и 3'-концевых фрагментов 18S рРНК.
А.В. Жигайлов, А.С. Низкородова, К.О. Шарипов, Н.С. Полимбетова, Б.К. Исакаев (на англ. языке)

- 99 **ОБЗОР**
Контроль мейотического кроссинговера в селекции растений.
С.Р. Стрельникова, Р.А. Комахин

Генетика растений

- 111 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Метаболомный подход в изучении *Dactylis glomerata* L. из коллекции ВИР.
Н.Ю. Малышева, Т.В. Шеленга, А.Е. Соловьева, Т.Б. Нагиев, Н.В. Ковалева, Л.Л. Малышев (на англ. языке)

- 119 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Изменчивость содержания алкалоидов в семенах люпина узколистного у образцов коллекции ВИР в условиях Северо-Запада Российской Федерации.
М.А. Вишнякова, А.В. Саликова, Т.В. Шеленга, Г.П. Егорова, Л.Ю. Новикова

Селекция растений на иммунитет и продуктивность

- 129 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
ДНК-маркерная идентификация локуса устойчивости к милдью *Rpv10* в генотипах винограда. Е.Т. Ильницкая, М.В. Макаркина, С.В. Токмаков, Л.Г. Наумова (на англ. языке)

- 135 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Анализ стабильности урожайности семян генотипов нута (*Cicer arietinum* L.) с помощью экспериментальных и биологических подходов.
Р. Каримзаде, П. Пезешкпур, А. Мирзаи, М. Барзали, П. Шарифи, М.Р. Сафари Мотлах (на англ. языке)

Филогенетика

- 146 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Митохондриальный геном *Dendrobaena tellermanica* Perel, 1966 (Annelida: Lumbricidae).
С.В. Шеховцов, Г.В. Васильев, Р. Латиф, Т.В. Полубоярова, С.Е. Пельтек, И.Б. Панопорт

- 153 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Использование первичной структуры района ITS1–ITS2 для видовой идентификации у некоторых представителей водных макрофитов рода *Stuckenia*.
А.В. Мглинец, О.Э. Костерин (на англ. языке)

Генетика животных

- 162 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Геногеографическое исследование киргизского горного меринуса с использованием микросателлитных маркеров. А.Б. Бектуров, Ж.Т. Исакова, В.Н. Кипень, Т.Д. Чортонбаев, С.Б. Мукуева, С.К. Осмоналиев, К.А. Айтбаев

- 169 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Генетический профиль популяции домашней кошки (*Felis catus* L.) острова Аошима (Япония). С.К. Холин

- 177 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Изменения в социальном предпочтении места и плотность дофаминергических нейронов в вентральном тегментуме у *Clsnt2*-КО мышей. И.Н. Рожкова, С.В. Окотруб, Е.Ю. Брусенцев, Е.Е. Ульданова, Э.А. Чуйко, В.А. Напримеров, Т.В. Липина, Т.Г. Амстиславская, С.Я. Амстиславский

Molecular and cell biology

- 93 ORIGINAL ARTICLE
Glyphosate treatment mediates the accumulation of small discrete 5'- and 3'-terminal fragments of 18S rRNA in plant cells. A.V. Zhigailov, A.S. Nizkorodova, K.O. Sharipov, N.S. Polimbetova, B.K. Iskakov
- 99 REVIEW
Control of meiotic crossing over in plant breeding. S.R. Strelnikova, R.A. Komakhin

Plant genetics

- 111 ORIGINAL ARTICLE
Metabolomic approach to investigate *Dactylis glomerata* L. from the VIR collection. N.Yu. Malysheva, T.V. Shelenga, A.E. Solovyeva, T.B. Nagiev, N.V. Kovaleva, L.L. Malyshev
- 119 ORIGINAL ARTICLE
Alkaloid content variability in the seeds of narrow-leaved lupine accessions from the VIR collection under the conditions of the Russian Northwest. M.A. Vishnyakova, A.V. Salikova, T.V. Shelenga, G.P. Egorova, L.Yu. Novikova

Plant breeding for immunity and performance

- 129 ORIGINAL ARTICLE
DNA marker identification of downy mildew resistance locus *Rpv10* in grapevine genotypes. E.T. Il'nitskaya, M.V. Makarkina, S.V. Tokmakov, L.G. Naumova
- 135 ORIGINAL ARTICLE
Stability analysis for seed yield of chickpea (*Cicer arietinum* L.) genotypes by experimental and biological approaches. R. Karimizadeh, P. Pezeshkpour, A. Mirzaee, M. Barzali, P. Sharifi, M.R. Safari Motlagh

Phylogenetics

- 146 ORIGINAL ARTICLE
The mitochondrial genome of *Dendrobaena tellermanica* Perel, 1966 (Annelida: Lumbricidae) and its phylogenetic position. S.V. Shekhovtsov, G.V. Vasiliev, R. Latif, T.V. Poluboyarova, S.E. Peltek, I.B. Rapoport
- 153 ORIGINAL ARTICLE
The use of the primary structure of the ITS1–ITS2 region for species identification in some submerged aquatic macrophytes of the genus *Stuckenia*. A.V. Mglinets, O.E. Kosterin

Animal genetics

- 162 ORIGINAL ARTICLE
A genogeographic study of the Kyrgyz mountain merino via microsatellite markers. A.B. Bekturov, Zh.T. Isakova, V.N. Kipen, T.Dzh. Chortonbaev, S.B. Mukeeva, S.K. Osmonaliev, K.A. Aitbaev
- 169 ORIGINAL ARTICLE
Genetic profile of domestic cat (*Felis catus* L.) population of Aoshima Island (Japan). S.K. Kholin
- 177 ORIGINAL ARTICLE
Alterations in the social-conditioned place preference and density of dopaminergic neurons in the ventral tegmental area in *Clnt2*-KO mice. I.N. Rozhkova, S.V. Okotrub, E.Yu. Brusentsev, K.E. Uldanova, E.A. Chuyko, V.A. Naprimerov, T.V. Lipina, T.G. Amstislavskaya, S.Ya. Amstislavsky

Glyphosate treatment mediates the accumulation of small discrete 5'- and 3'-terminal fragments of 18S rRNA in plant cells

A.V. Zhigailov , A.S. Nizkorodova, K.O. Sharipov, N.S. Polimbetova, B.K. Iskakov

M.A. Aitkhozhin Institute of Molecular Biology and Biochemistry of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Almaty, Kazakhstan
 andrzhig@gmail.com

Abstract. Under many kinds of stress, eukaryotic cells rapidly decrease the overall translation level of the majority of mRNAs. However, some molecular mechanisms of protein synthesis inhibition like phosphorylation of eukaryotic elongation factor 2 (eEF2), which are known to be functional in animals and yeast, are not implemented in plants. We suggest that there is an alternative mechanism for the inhibition of protein synthesis in plant cells and possibly, in other eukaryotes, which is based on the discrete fragmentation of 18S rRNA molecules within small ribosomal subunits. We identified four stress-induced small RNAs, which are 5'- and 3'-terminal fragments of 18S rRNA. In the present work, we studied the induction of 18S rRNA discrete fragmentation and phosphorylation of the α -subunit of eukaryotic initiation factor 2 (eIF2 α) in germinated wheat embryos in the presence of glyphosate, which imitates the condition of amino acid starvation. Using northern and western blotting, we have shown that stress-induced 18S rRNA fragments started to accumulate in wheat embryos at glyphosate concentrations that did not evoke eIF2 α phosphorylation. It was also found that cleavage of 18S rRNA near the 5'-terminus began much earlier than eIF2 α phosphorylation, which became noticeable only at higher concentration (500 μ M) of glyphosate. This result suggests that discrete fragmentation of 18S rRNA may constitute a regulatory mechanism of mRNA translation in response to stress and may occur in plant cells in parallel with and independently of eIF2 α phosphorylation. The identified small 5'- and 3'-terminal fragments of 18S rRNA that accumulate during various stresses may serve as stress resistance markers in the breeding of economically important plant crops.

Key words: wheat embryos; 18S rRNA; discrete fragmentation; 40S ribosomal subunits; glyphosate; eIF2 α phosphorylation; stress; starvation.

For citation: Zhigailov A.V., Nizkorodova A.S., Sharipov K.O., Polimbetova N.S., Iskakov B.K. Glyphosate treatment mediates the accumulation of small discrete 5'- and 3'-terminal fragments of 18S rRNA in plant cells. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):93-98. DOI 10.18699/VJGB-23-13

Обработка глифосатом приводит к накоплению в клетках растений малых дискретных 5'- и 3'-концевых фрагментов 18S рРНК

А.В. Жигайлов , А.С. Низкородова, К.О. Шарипов, Н.С. Полимбетова, Б.К. Искаков

Институт молекулярной биологии и биохимии им. М.А. Айтхожина Министерства науки и высшего образования Республики Казахстан, Алматы, Казахстан
 andrzhig@gmail.com

Аннотация. При многих видах стресса эукариотические клетки быстро снижают общий уровень трансляции большинства мРНК. Однако некоторые молекулярные механизмы ингибирования синтеза белка, такие как фосфорилирование эукариотического фактора элонгации трансляции (eEF2), функционируют у животных и дрожжей, но не реализуются у растений. Мы предполагаем, что существует альтернативный механизм ингибирования синтеза белка в клетках растений и, возможно, других эукариот, основанный на дискретной фрагментации молекул 18S рРНК внутри малых субъединиц рибосомы. Мы идентифицировали четыре малые РНК, индуцированные стрессом, которые представляют собой 5'- и 3'-концевые фрагменты 18S рРНК. В настоящей работе мы исследовали индукцию дискретной фрагментации 18S рРНК и фосфорилирование α -субъединицы эукариотического фактора инициации 2 (eIF2 α) в проросших зародышах пшеницы в присутствии глифосата, имитирующего состояние аминокислотного голодания. Используя нозерн- и вестерн-блоттинг, мы показали, что индуцированные стрессом фрагменты 18S рРНК начинают накапливаться в зародышах пшеницы при концентрациях глифосата, не вызывающих фосфорилирования eIF2 α . Также установлено, что расщепление 18S рРНК вблизи 5'-конца начинается гораздо раньше, чем становится заметным фосфорилирование eIF2 α при высокой концентрации глифосата (500 мкМ). Этот результат указывает на то, что дискретная фрагментация 18S рРНК может представлять собой регуляторный механизм трансляции мРНК в ответ на стресс и происходить в растительных клетках параллельно с фосфорилированием eIF2 α и независимо от него. Выявленные 5'- и 3'-концевые малые фрагменты 18S рРНК, накапливающиеся при различных стрессах, могут служить маркерами стрессоустойчивости в процессе селекции хозяйственно важных культур растений.

Ключевые слова: зародыши пшеницы; 18S рРНК; дискретная фрагментация; 40S рибосомные субъединицы; глифосат; фосфорилирование eIF2 α ; стресс; голодание.

Introduction

Protein biosynthesis is a very energy-intensive process, so under stress conditions, the translation of most cellular mRNAs is inhibited in order to save energy and resources and to ensure preferential synthesis of stress proteins. Several molecular mechanisms of protein synthesis inhibition have been described in mammalian and yeast cells. One of these mechanisms is the eukaryotic translation elongation factor 2 (eEF2) phosphorylation, which is carried out by a highly specific protein kinase in response to a sharp decrease in cytosolic ATP concentration levels. Phosphorylation inactivates mammalian eEF2 by preventing it from binding to the ribosome (Ballard et al., 2021). However, plants do not exhibit endogenous kinase activity for eEF2 either under normal conditions (Smailov et al., 1993), or under stress (Gallie et al., 1998).

The second mechanism known in animals to reduce the level of mRNA translation is triggered under conditions of amino acid starvation and is mediated by eIF4E-binding proteins (4E-BPs), which prevent eIF4E from binding to the m⁷G-cap structure of mRNA (Hernandez et al., 2010). However, no clear homologs of these eIF4E-BPs have yet been found in plants (Echevarria-Zomeno et al., 2013), nor were any orthologues of the 4E-BPs genes found in plant genomes (Browning, Bailey-Serres, 2015).

Another important mechanism of eukaryotic protein synthesis inhibition is the phosphorylation of the α -subunit of eukaryotic initiation factor 2 (eIF2 α) by specific protein kinases. This process in mammalian and yeast cells leads to the blocking of GDP \rightarrow GTP exchange protein eIF2B and to a sharp inhibition of the mRNA translation initiation (Baird, Wek, 2012). However, recycling of the ternary complex in plant cells can occur without the participation of eIF2B (Shaikhin et al., 1992), and eIF2 α phosphorylation in plant systems *in vitro* does not lead to strong inhibition of protein synthesis (Zhigailov et al., 2020). In addition, of the four protein kinases (mPKR, mHCR, mPERK, mGCN2) that phosphorylate the eIF2 α in mammalian cells, only pGCN2-kinase was found in plants, and the phosphorylation of eIF2 α in plants is not a universal response to all stress types (Immanuel et al., 2012; Zhigailov et al., 2020).

Thus, the mechanisms of protein biosynthesis suppression due to the phosphorylation of translational factors, which are well described for mammals and yeast, are either used to a limited extent or are not realized at all in plant cells. We suggest that another mechanism of protein synthesis inhibition can function in plants, which is triggered by certain abiotic and biotic stresses. In our understanding, this mechanism is associated with the cleavage at certain sites of the 18S rRNA as part of 40S ribosomal subunits (40S RS). Previously, we described the process of 18S rRNA cleavage, leading to 5'-terminal fragments formation of 132–134 nt. (Zhanybekova et al., 1996) and of 54–57 nt. (Zhigailov et al., 2014), as well as a 3'-terminal fragment of 100 nt. (Zhigailov et al., 2013). Our data are quite consistent with the data of full-transcriptome analysis, which showed that breaks in 28S-, 18S-, and 5.8S-rRNA do not occur randomly, but discretely, which leads to the fact that some fragments of ribosomal RNA are detected in the cell significantly more often than other fragments (Chen et al., 2017).

The process of RNA cleavage is widely used by cells during the processing of ribosomal RNA from their precursor during

ribosome biogenesis (Henras et al., 2015). In addition, in pro- and eukaryotic organisms, the mechanism of protein biosynthesis suppression is realized due to the cleavage of the 28S rRNA molecule from the large (60S) ribosomal subunit along the sarcin-ricin loop with the cleavage of the 3'-terminal Endo-rRNA-fragment (Endo, 1988). Toxins of plants (ricin, abrin, and modecin), fungi (α -sarcin) and bacteria (Shiga toxin) act this way (Kast et al., 2014). Possibly similar endonucleases and/or glycosylases (that mediate abasic site formation as in the case of ribosome inactivating proteins, RIPs) are activated in plant cells during stress, but targeting 18S rRNA in 40S RS instead of 28S rRNA in 60S RS, and thus leading to temporary or permanent suppression of mRNA translation.

In this work, we have shown that in the case of glyphosate-mediated amino acid starvation, when the only specific eIF2 α kinase of plants (pGCN2-kinase) is activated, in addition to plant eIF2 α phosphorylation, another protective mechanism is triggered in plant cells, namely, discrete fragmentation of 18S rRNA. It was shown that the accumulation in plants of 18S rRNA 5'-terminal fragments of 75 nucleotides (75nt-5'18S) and 134 nucleotides (134nt-5'18S) begins earlier than the activation of pGCN2 kinase and becomes noticeable at relatively low concentrations of glyphosate when plant eIF2 α phosphorylation does not occur at all.

Materials and methods

Plant material and treatment. Wheat (*Triticum aestivum* L. cv. Kazakhstanskaya 10) seeds were sterilized in 70 % (v/v) ethanol for 2 min, then in 2 % (w/v) NaOCl for 20 min, and washed thoroughly with sterile water. Seeds were germinated at 26 °C on sterile filter paper soaked in water. After 18 hours, viable embryos were isolated by spatula from swollen seeds and placed in 1 % glucose solution containing 50 U/ml penicillin, 50 μ g/ml chloramphenicol, and 50 μ g/ml nystatin. After this, embryos were divided into equal portions (1 g), which were subjected to treatment with glyphosate (simulation of amino acid starvation) or without any additives (control).

Synthesis of probes. DIG-labeling of *de novo* synthesized oligodeoxyribonucleotides 5'18S (5'-ACAAGCATATGACTACTGGCAGGATCAACCAGGTA) and 3'18S (5'-CAATGATCCTTCCGCAGGTTACCTACGGAAACCT) was carried out using DIG Oligonucleotide 3'-End Labeling Kit (Roche) according to the manufacturer's manual. Probes (5'18S-DIG and 3'18S-DIG) were used for northern blotting.

Northern blotting. Total RNA was extracted from plant tissues with Tri-reagent (Sigma Aldridge) and analyzed on 10 % PAGE with 8 M urea in Tris-borate buffer (1xTBE: 89 mM Tris-borate, 2 mM EDTA, pH 8.3). RNAs were blotted to a nylon membrane (Roche) equilibrated in 0.1x TBE using a semi-dry blotter (Sci-Plas) at 250 mA for 30 min. The membrane was dried and irradiated with UV light for 2 min at 10 mJ/cm² in a crosslinker (UVP). Hybridization of DIG-labeled probes and subsequent chemiluminescent band detection was performed with DIG Luminescent Detection Kit for Nucleic Acids (Roche) according to the manufacturer's procedure. The hybridization temperature was 55 °C. Anti-Digoxigenin-AP Fab fragment conjugates (Roche) were used to detect bound DIG-labeled probes. The blots were developed using a commercial alkaline phosphatase substrate CSPD (Roche).

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting. Frozen embryos were ground to a powder in a mortar and then homogenized in Laemmli sample buffer (Laemmli, 1970). Proteins were separated by 12.5 % SDS-PAGE with 0.1 % SDS. The separated proteins were transferred to a nitrocellulose membrane (GVS) that was afterwards stained with Ponceau S (Sigma-Aldrich). The antibodies against human phospho-eIF2 α (S51) produced in rabbit (CellSignaling Technology, 1:1000) were used for the immune-detection of phosphorylated *T. aestivum* (*Ta*) eIF2 α (*TaeIF2*(α P)). Then horseradish peroxidase-conjugated anti-rabbit secondary antibodies produced in donkey (ECL, 1:2000 dilution) were used.

Results

The effect of glyphosate concentration on 18S rRNA fragmentation in wheat embryos. Since the mechanisms of mRNA translation inhibition mediated by 4E-BPs and eEF2K are not implemented in plants, it is believed that the main response in plants to amino acid starvation is eIF2 α phosphorylation with pGCN2 kinase (Zhang et al., 2008). To test whether the process of discrete fragmentation of 18S rRNA is also induced under these conditions, the herbicide glyphosate was used. Glyphosate targets 5-enolpyruvylshikimate 3-phosphate synthase, which catalyzes the key penultimate reaction in the shikimate pathway (Padgett et al., 1995). Therefore, it inhibits the synthesis of many aromatic plant metabolites including the amino acids tryptophan, tyrosine, and phenylalanine and leads to pGCN2 kinase activation and phosphorylation of the plant eIF2 α (Zhang et al., 2008). Germinated wheat embryos were treated with glyphosate at various concentrations, after which the content of 18S rRNA small fragments and the phosphorylation status of *TaeIF2* α were assessed in their cells. The results are present in Figure 1.

Phosphorylation of *TaeIF2* α becomes noticeable only at relatively high concentrations (0.5 and 5 μ M) of glyphosate (tracks 4 and 5 on Fig. 1, *d*; right panel), at which wheat embryos stopped to grow (variants 4 and 5 on Fig. 1, *a*). The appearance of 3'-terminal fragments 100nt-3'18S and 70nt-3'18S was observed at the same concentrations of glyphosate (tracks 4 and 5 on Fig. 1, *c*; right panel). At the same time, 5'-terminal fragments of 18S rRNA, 134nt-5'18S and 75nt-5'18S, began to accumulate in noticeable amounts even at very low concentrations (5 μ M) of glyphosate (see Fig. 1, *b*). The results of semi-quantitative optical densitometry analysis for this experiment are presented in Table 1. Since the 134nt-5'18S fragment can be a precursor of 75nt-5'18S, and the 100nt-3'18S fragment can act as a precursor for 70nt-3'18S, it is reasonable to estimate the sum of these small 18S rRNA fragments.

The dynamics of glyphosate influence on 18S rRNA fragmentation in wheat embryos. Then, we assessed how quickly wheat embryos respond to glyphosate treatment by measuring the time dependence of *TaeIF2* α phosphorylation and of discrete fragmentation of 18S rRNA. For this, a glyphosate concentration of 500 μ M was chosen, which induced quite effective phosphorylation of *TaeIF2* α , as well as a significant increase in the content of 18S rRNA small fragments: 134nt-5'18S, 75nt-5'18S, 100nt-3'18S and 70nt-3'18S (see Fig. 1, Table 1). The results of the experiment are shown in Figure 2.

The results of semi-quantitative optical densitometry analysis of the data presented in Figure 2 are shown in Table 2.

Data presented in Figure 2 and Table 2 show that *TaeIF2* α phosphorylation begins 45 min after the start of glyphosate treatment (a faintly visible band on track 4 on Fig. 2, *c*; right panel), and *TaeIF2* α P becomes quite noticeable after 60 min of such treatment (track 5 in Fig. 2, *c*; immunoblot).

The 3'-terminal fragmentation of 18S rRNA is observed after 3 hours after the start of glyphosate treatment: fragments 100nt-3'18S and 70nt-3'18S become detectable as faintly visible bands on track 7 of Figure 2, *b* (right panel). The amount of these 3'-coterminial fragments is significantly lower than after 10 hours of the same treatment with glyphosate (compare with track 4 on Fig. 1, *c*; right panel).

As for fragmentation from the 5'-terminus of 18S rRNA, the amount of both 5'-coterminial fragments, 134nt-5'18S and 75nt-5'18S, is significantly increased as early as by the 15th min after the start of treatment of wheat embryos with glyphosate (see Fig. 2, *a*; right panel). Notably, the amount of the fragment 134nt-5'18S is higher than that of 75nt-5'18S fragment during 30–45 min of incubation with glyphosate. By 60 min of incubation their amounts become almost equal and after that, the amount of 75nt-5'18S fragment becomes higher (by 90 min) and even obviously prevalent by 180 min (see Fig. 2, *a*; right panel). Similar interrelation can be seen in Figure 1, *b* (right panel) regarding the applied concentrations of glyphosate.

These observations suggest that cleavage at 134th nucleotide may happen more quickly and this site is more susceptible at the beginning of stress. The cleavage site at 75th nucleotide becomes more prevalent with an increase of stress duration and severity. The cleavage sites at the 3'-terminal segment of 18S rRNA occur only at very high severity and duration of stress. Therefore, there seemingly exist several different mechanisms for the cleavage at 5'- and 3'-termini of 18S rRNA, which may result in several different consequences for the functioning of 40S RS.

Discussion

No phosphorylation of eIF2 α was observed in plants under osmotic and oxidative stresses (Lageix et al., 2008), heat shock (Gallie et al., 1997; Echevarria-Zomeno et al., 2013) and during unfolded protein response in plants (Kamauchi et al., 2005). At the same time, during these stresses a significant decrease in the translation level of most mRNAs is observed with exception only for those templates that are responsible for the synthesis of stress proteins (Altschuler, Mascarenhas, 1982; Ruberti et al., 2015). Most likely, in plants, other mechanisms of protein biosynthesis suppression are realized, than eIF2 α phosphorylation (Yu et al., 2021). In addition, eIF2 α phosphorylation is not the only possible mechanism of response to some types of stress in different eukaryotic cells. For example, when yeast cells are exposed to harsh ultraviolet light, phosphorylation of eIF2 α is observed, as well as a suppression of the overall level of protein synthesis. However, inhibition of mRNA translation upon exposure to UV light occurs even in cells containing a mutant form of eIF2 α that is not capable of phosphorylation (Knutsen et al., 2015).

We postulate that the process of discrete fragmentation of 18S rRNA observed under glyphosate mediated amino

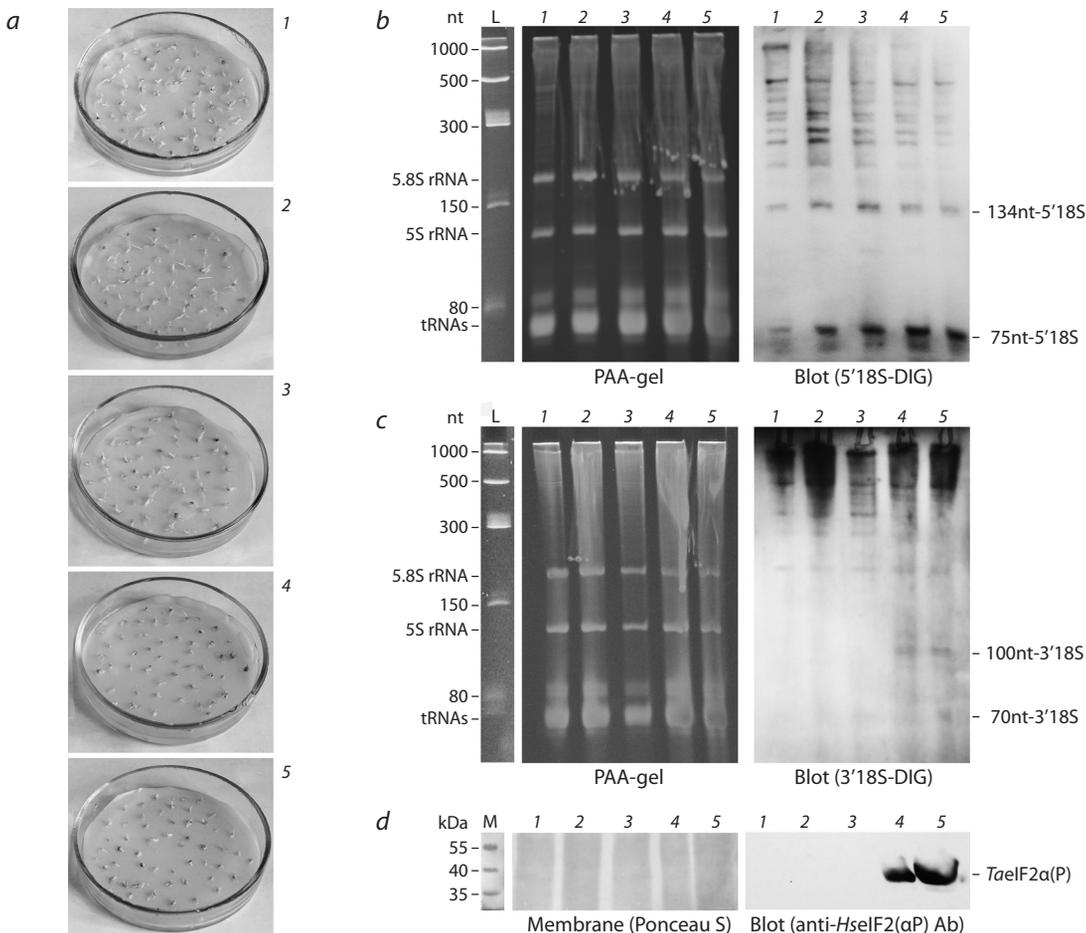


Fig. 1. The effect of glyphosate on *TaelF2a* phosphorylation and 18S rRNA fragmentation in germinated wheat embryos. *a*, The appearance of wheat embryos exposed to different concentration of glyphosate; *b*, Northern blot analysis using 5'18S-DIG probe (right panel). Left panel – ethidium bromide stained PAA-gel; *c*, Northern blot analysis using 3'18S-DIG probe (right panel). Left panel – ethidium bromide stained PAA-gel; *d*, Phosphorylation status of *TaelF2a* in wheat embryos. Presented are the membrane stained with Ponceau S (left panel) and blot-membrane developed using anti-*HsElf2*(αP) antibodies (right panel).

For all variants, embryos were first germinated at 26 °C for 18 h and then incubated at 26 °C for 10 h in the absence or presence of glyphosate at the following concentration: 1 – 0 μM (control); 2 – 5 μM; 3 – 50 μM; 4 – 0.5 μM; 5 – 5 μM. L – Low Range ssRNA ladder; M – PageRuler Plus Protein Ladder.

Table 1. Optical densitometry analysis for assessing the 18S rRNA fragments in wheat embryos following treatment with glyphosate at various concentrations

Glyphosate concentration, μM	Normalized signal					
	$S_{(134nt-5'18S)} / S_{(5S\ RNA)}$	$S_{(75nt-5'18S)} / S_{(5S\ RNA)}$	$S_{(134nt-5'18S + 75nt-5'18S)} / S_{(5S\ RNA)}$	$S_{(100nt-3'18S)} / S_{(5S\ RNA)}$	$S_{(70nt-3'18S)} / S_{(5S\ RNA)}$	$S_{(100nt-3'18S + 70nt-3'18S)} / S_{(5S\ RNA)}$
0	0.14 ± 0.01	0.28 ± 0.07	0.42 ± 0.07	0.15 ± 0.01	0.17 ± 0.03	0.33 ± 0.03
5	0.29 ± 0.04*	0.8 ± 0.15*	1.1 ± 0.15*	0.19 ± 0.03	0.19 ± 0.02	0.38 ± 0.05
50	0.44 ± 0.1**	0.78 ± 0.12*	1.2 ± 0.1**	0.25 ± 0.04	0.27 ± 0.03	0.52 ± 0.06
500	0.2 ± 0.02*	0.9 ± 0.19*	1.1 ± 0.19*	0.8 ± 0.1**	0.58 ± 0.09*	1.4 ± 0.2**
5000	0.2 ± 0.02	1.18 ± 0.21*	1.4 ± 0.2**	1.2 ± 0.1**	0.86 ± 0.15*	2.01 ± 0.18**

Here and in Table 2: * $p < 0.05$; ** $p < 0.001$ (when compared with control). Densitometry analysis was performed using "ImageJ 1.42q" software.

acid starvation may lead to a decrease in the level of mRNA translation. This molecular mechanism can be realized in parallel with the known mechanism of translational regulation mediated by eIF2α phosphorylation and independently of it.

Understanding the molecular mechanisms of plant adaptation to stresses can make it possible to increase the efficiency of breeding work to obtain genetic lines and varieties of economically important plant species that are characterized by increased resistance to certain stresses.

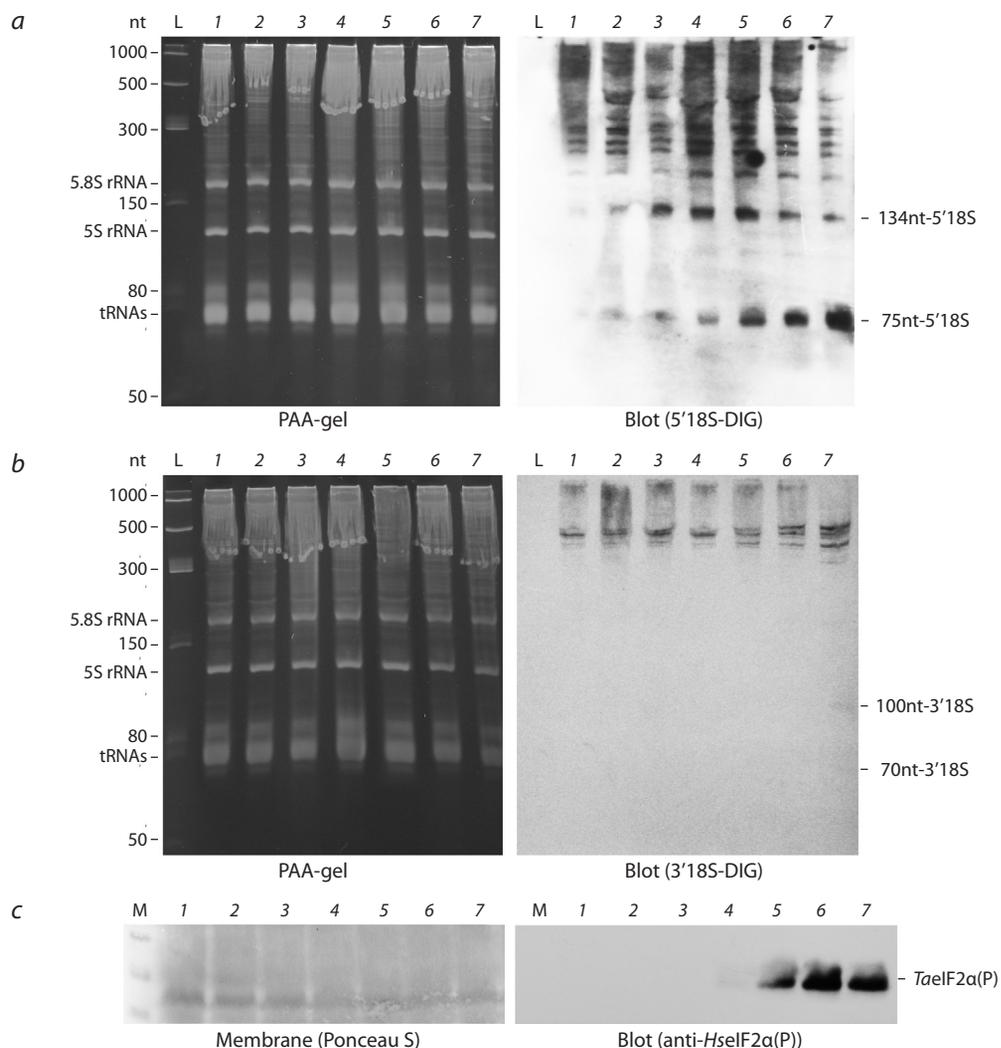


Fig. 2. The dynamics of glyphosate action on the *TaelF2α* phosphorylation and 18S rRNA fragmentation in germinated wheat embryos. *a*, Northern blotting analysis (right panel) using 5'18S-DIG probe. Left panel – ethidium bromide stained PAA-gel; *b*, Northern blotting analysis (right panel) using 3'18S-DIG probe. Left panel – ethidium bromide stained PAA-gel; *c*, Phosphorylation status of *TaelF2α* in wheat embryos exposed to glyphosate treatment. Presented are the membrane stained with Ponceau S (left panel) and blot-membrane developed using anti-*HsElf2α*(P) antibodies (right panel).

The embryos were first germinated at 26 °C for 18 h and then incubated at 26 °C in the presence of 0.5 μM glyphosate during the following periods: 1 – 0 min; 2 – 15 min; 3 – 30 min; 4 – 45 min; 5 – 60 min; 6 – 90 min; 7 – 180 min. L – Low Range ssRNA ladder; M – PageRuler Plus Protein Ladder.

Table 2. Densitometry analysis for assessing the 18S rRNA fragments in wheat embryos treated with 500 μM glyphosate for different time periods

Time, min	Normalized signal					
	$S_{(134nt-5'18S)} / S_{(5S RNA)}$	$S_{(75nt-5'18S)} / S_{(5S RNA)}$	$S_{(134nt-5'18S + 75nt-5'18S)} / S_{(5S RNA)}$	$S_{(100nt-3'18S)} / S_{(5S RNA)}$	$S_{(70nt-3'18S)} / S_{(5S RNA)}$	$S_{(100nt-3'18S + 70nt-3'18S)} / S_{(5S RNA)}$
0	0.23 ± 0.02	0.17 ± 0.01	0.81 ± 0.02	0.08 ± 0.01	0.09 ± 0.01	0.17 ± 0.01
15	0.85 ± 0.14*	0.93 ± 0.01**	3.54 ± 0.1**	0.08 ± 0.01	0.1 ± 0.01	0.19 ± 0.01
30	1.84 ± 0.28**	1.45 ± 0.06**	6.5 ± 0.34**	0.08 ± 0.01	0.12 ± 0.01	0.2 ± 0.02
45	2.4 ± 0.35**	1.3 ± 0.06**	7.3 ± 0.4**	0.08 ± 0.01	0.12 ± 0.02	0.19 ± 0.03
60	2.4 ± 0.46**	3.25 ± 0.4**	11.1 ± 0.8**	0.07 ± 0.01	0.09 ± 0.01	0.16 ± 0.02
90	1.42 ± 0.29*	3.7 ± 0.24**	10.0 ± 0.5**	0.09 ± 0.01	0.13 ± 0.02	0.22 ± 0.03
180	1.06 ± 0.25*	8.0 ± 1.16**	18.0 ± 1.3**	0.33 ± 0.5**	0.22 ± 0.03*	0.55 ± 0.06**

Conclusion

This paper presents data indicating that in plant cells the imitiation of amino acid starvation induces, in addition to eIF2 α phosphorylation, another cellular response that involves the cleavage of the 18S rRNA molecule with the formation of discrete 5'- and 3'-terminal fragments. At the same time, 3'-terminal fragments of 18S rRNA appear only at lethal concentrations of glyphosate and after a prolonged period of stress (3 hours or more). In contrast, 5'-terminal fragments of 18S rRNA began to accumulate in wheat embryos at relatively low glyphosate concentrations, at which wheat embryos could continue development, and already 15 min after the start of glyphosate treatment. Thus, the process of 18S rRNA fragmentation in wheat embryo 40S RS is triggered even under conditions where eIF2 α phosphorylation does not occur. We suggest that such cleavage of the 18S rRNA molecule, which is activated during amino acid starvation, may result in either global or selective suppression of mRNA translation.

References

Altschuler M., Mascarenhas J.P. Heat shock proteins and effects of heat shock in plants. *Plant Mol. Biol.* 1982;1(2):103-115. DOI 10.1007/BF00024974.

Baird Th.D., Wek R.C. Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. *Adv. Nutr.* 2012;3:307-321. DOI 10.3945/an.112.002113.

Ballard D.J., Peng H.-Y., Das J.K., Kumar A., Wang L., Ren Y., Xiong X., Ren X., Yang J.-M., Song J. Insights into the pathologic roles and regulation of eukaryotic elongation factor-2 kinase. *Front. Mol. Biosci.* 2021;8:839. DOI 10.3389/fmolb.2021.727863.

Browning K.S., Bailey-Serres J. Mechanism of cytoplasmic mRNA translation. *Arabidopsis Book.* 2015;13:e0176.

Chen Z., Sun Y., Yang X., Wu Z., Guo K., Niu X., Wang Q., Ruan J., Bu W., Gao S. Two featured series of rRNA-derived RNA fragments (rRFs) constitute a novel class of small RNAs. *PLoS One.* 2017;12:e0176458. DOI 10.1371/journal.pone.0176458.

Echevarria-Zomeno S., Yanguéz E., Fernandez-Bautista N., Castro-Sanz A.B., Ferrando A., Castellano M.M. Regulation of translation initiation under biotic and abiotic stresses. *Int. J. Molec. Sci.* 2013;14:4670-4683. DOI 10.3390/ijms14034670.

Endo Y. Mechanism of action of ricin and related toxins on the inactivation of eukaryotic ribosomes. *Cancer Res. Treat.* 1988;37:75-89. DOI 10.1007/978-1-4613-1083-9_5.

Gallie D.R., Le H., Caldwell C., Tanguay R., Hoang N.X., Browning K.S. The phosphorylation state of translation initiation factors is regulated developmentally and following heat shock in wheat. *J. Biol. Chem.* 1997;272:1046-1053. DOI 10.1074/jbc.272.2.1046.

Gallie D.R., Le H., Caldwell C., Browning K.S. Analysis of translation elongation factors from wheat during development and following heat shock. *Biochem. Biophys. Res. Comm.* 1998;245:295-300. DOI 10.1006/bbrc.1998.8427.

Henras A.K., Plisson-Chastang C., O'Donohue M.F., Chakraborty A., Gleizes P.E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA.* 2015;6(2):225-242. DOI 10.1002/wrna.1269.

Hernandez G., Altmann M., Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. *Trends Biochem. Sci.* 2010;35(2):63-73. DOI 10.1016/j.tibs.2009.10.009.

Immanuel T.M., Greenwood D.R., MacDiarmid R.M. A critical review of translation initiation factor eIF2 α kinases in plants – regulating protein synthesis during stress. *Funct. Plant Biol.* 2012;39(9):717-735. DOI 10.1071/FP121116.

Kamauchi S., Nakatani H., Nakano C., Urade R. Gene expression in response to endoplasmic reticulum stress in *Arabidopsis thaliana*. *FEBS J.* 2005;272(13):3461-3476. DOI 10.1111/j.1742-4658.2005.04770.x.

Kast A., Klassen R., Meinhardt F. rRNA fragmentation induced by a yeast killer toxin. *Mol. Microbiol.* 2014;91(3):606-617. DOI 10.1111/mmi.12481.

Knutsen J.H., Rødland G.E., Bøe C.A., Håland T.W., Sunnerhagen P., Grallert B., Boye E. Stress-induced inhibition of translation independently of eIF2 α phosphorylation. *J. Cell Sci.* 2015;128(23):4420-4427. DOI 10.1242/jcs.176545.

Laemmli U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970;227(5259):680-685. DOI 10.1038/227680a0.

Lageix S., Lanet E., Pouch-Pelissier M.N., Espagnol M.C., Robaglia C., Deragon J.M., Pelissier T. *Arabidopsis* eIF2 α kinase GCN2 is essential for growth in stress conditions and is activated by wounding. *BMC Plant Biol.* 2008;8:134. DOI 10.1186/1471-2229-8-134.

Padgett S.R., Kolacz K.H., Delannay X., Re D.B., LaVallee B.J., Tinius C.N., Rhodes W.K., Otero Y.I., Barry G.F., Eichholtz D.A., Peschke V.M., Nida D.L., Taylor N.B., Kishore G.M. Development, identification and characterization of a Glyphosate-tolerant soybean line. *Crop Sci.* 1995;35(5):1451-1461. DOI 10.3389/fpls.2016.01009.

Ruberti C., Kim S.J., Stefano G., Brandizzi F. Unfolded protein response in plants: one master, many questions. *Curr. Opin. Plant Biol.* 2015;27:59-66. DOI 10.1016/j.pbi.2015.05.016.

Shaikhin S.M., Smailov S.K., Lee A.V., Kozhanov E.V., Iskakov B.K. Interaction of wheat germ translation initiation factor 2 with GDP and GTP. *Biochimie.* 1992;74(5):447-454. DOI 10.1016/0300-9084(92)90085-s.

Smailov S.K., Lee A.V., Iskakov B.K. Study of phosphorylation of translation elongation factor 2 (EF-2) from wheat germ. *FEBS Lett.* 1993;321(2-3):219-223. DOI 10.1016/0014-5793(93)80112-8.

Zhang Y., Wang Y., Kanyuka K., Parry M.A., Powers S.J., Halford N.G. GCN2-dependent phosphorylation of eukaryotic translation initiation factor-2 α in *Arabidopsis*. *J. Exp. Bot.* 2008;59(11):3131-3141. DOI 10.1093/jxb/ern169.

Zhanybekova S.S., Polimbetova N.S., Nakisbekov N.O., Iskakov B.K. Detection of a new small RNA, induced by heat shock, in wheat seed ribosomes. *Biochemistry (Moscow).* 1996;61:862-870.

Zhigailov A.V., Alexandrova A.M., Nizkorodova A.S., Stanbekova G.E., Kryldakov R.V., Karpova O.V., Polimbetova N.S., Halford N.G., Iskakov B.K. Evidence that Phosphorylation of the α -subunit of eIF2 does not essentially inhibit mRNA translation in wheat germ cell-free system. *Front. Plant Sci.* 2020;11:936. DOI 10.3389/fpls.2020.00936.

Zhigailov A.V., Polimbetova N.S., Borankul R.I., Iskakov B.K. Investigation of discrete fragmentation of 18S rRNA within 40S ribosomal subparticles of plant cells. *Vestnik KazNU. Biological Series.* 2013;2:81-87. (in Russian)

Zhigailov A.V., Polimbetova N.S., Doshchanov Kh.I., Iskakov B.K. Detection in plant cells of a new 75-nucleotide cytoplasmic RNA corresponding to the 5'-terminal fragment of 18S RNA. *Vestnik KazNU. Biological and Medical Series.* 2014;1:191-194. (in Russian)

Yu Ch.-Y., Cho Y., Sharma O., Kanehara K. What's unique? The unfolded protein response in plants. *J. Exp. Botany.* 2021:erab513. DOI 10.1093/jxb/erab513.

ORCID ID

A.V. Zhigailov orcid.org/0000-0002-9646-033X
A.S. Nizkorodova orcid.org/0000-0002-1597-7207

K.O. Sharipov orcid.org/0000-0001-5946-5521
N.S. Polimbetova orcid.org/0000-0002-2806-3009
B.K. Iskakov orcid.org/0000-0002-5204-4377

Acknowledgements. Current work was carried out in the framework of scientific grant AP14869357 and program OR11465447 funded by the Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

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

Received March 9, 2022. Revised August 31, 2022. Accepted September 1, 2022.

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

Control of meiotic crossing over in plant breeding

S.R. Strelnikova , R.A. Komakhin

All-Russia Research Institute of Agricultural Biotechnology, Moscow, Russia
 recombination@iab.ac.ru

Abstract. Meiotic crossing over is the main mechanism for constructing a new allelic composition of individual chromosomes and is necessary for the proper distribution of homologous chromosomes between gametes. The parameters of meiotic crossing over that have developed in the course of evolution are determined by natural selection and do not fully suit the tasks of selective breeding research. This review summarizes the results of experimental studies aimed at increasing the frequency of crossovers and redistributing their positions along chromosomes using genetic manipulations at different stages of meiotic recombination. The consequences of inactivation and/or overexpression of the *SPO11* genes, the products of which generate meiotic double-strand breaks in DNA, for the redistribution of crossover positions in the genome of various organisms are discussed. The results of studies concerning the effect of inactivation or overexpression of genes encoding RecA-like recombinases on meiotic crossing over, including those in cultivated tomato (*Solanum lycopersicum* L.) and its interspecific hybrids, are summarized. The consequences of inactivation of key genes of the mismatch repair system are discussed. Their suppression made it possible to significantly increase the frequency of meiotic recombination between homeologues in the interspecific hybrid yeast *Saccharomyces cerevisiae* × *S. paradoxus* and between homologues in arabidopsis plants (*Arabidopsis thaliana* L.). Also discussed are attempts to extrapolate these results to other plant species, in which a decrease in reproductive properties and microsatellite instability in the genome have been noted. The most significant results on the meiotic recombination frequency increase upon inactivation of the *FANCM*, *TOP3a*, *RECQ4*, *FIGL1* crossover repressor genes and upon overexpression of the *HEI10* crossover enhancer gene are separately described. In some experiments, the increase of meiotic recombination frequency by almost an order of magnitude and partial redistribution of the crossover positions along chromosomes were achieved in arabidopsis while fully preserving fecundity. Similar results have been obtained for some crops.

Key words: meiosis; DNA; reparation; recombination; crossing over; plant breeding.

For citation: Strelnikova S.R., Komakhin R.A. Control of meiotic crossing over in plant breeding. *Vavilovskii Zhurnal Genetiki i Selekcii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):99-110. DOI 10.18699/VJGB-23-15

Контроль мейотического кроссинговера в селекции растений

С.Р. Стрельникова , Р.А. Комахин

Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии, Москва, Россия
 recombination@iab.ac.ru

Аннотация. Мейотический кроссинговер является основным механизмом конструирования нового аллельного состава индивидуальных хромосом и необходим для равнозначного распределения гомологичных хромосом между гаметами. Сложившиеся в ходе эволюции параметры мейотического кроссинговера определены естественным отбором и не полностью соответствуют задачам селекционных исследований. В настоящем обзоре суммированы результаты экспериментальных работ, направленных на повышение частоты кроссоверов и перераспределение их позиций вдоль хромосом с помощью генетических манипуляций на разных этапах мейотической рекомбинации. Обсуждаются последствия инактивации и/или сверхэкспрессии генов *SPO11*, продукты которых генерируют мейотические двуцепочечные разрывы в ДНК, для перераспределения позиций кроссоверов в геноме различных организмов. Обобщены результаты исследований по влиянию инактивации или сверхэкспрессии генов RecA-подобных рекомбиназ на мейотический кроссинговер, в том числе у культурного томата (*Solanum lycopersicum* L.) и его межвидовых гибридов. Обсуждаются последствия инактивации ключевых генов системы мисмэтч-репарации. Их подавление позволило достоверно повысить частоту мейотической рекомбинации между гомеологами у межвидового гибрида дрожжей *Saccharomyces cerevisiae* × *S. paradoxus* и между гомологами у растений арабидопсиса (*Arabidopsis thaliana* L.). Рассматриваются попытки экстраполировать эти результаты на другие виды растений, у которых отмечены снижение репродуктивных свойств и микросателлитная нестабильность в геноме. Отдельно описаны наиболее значимые результаты по увеличению частоты мейотической рекомбинации при инактивации генов-репрессоров кроссинговера *FANCM*, *TOP3a*, *RECQ4*, *FIGL1* и при сверхэкспрессии гена-энхансера кроссинговера *HEI10*. В некоторых экспериментах удалось практически на порядок повысить частоту мейотической рекомбинации и частично перераспределить позиции кроссоверов вдоль хромосом при полном сохранении плодовитости у арабидопсиса. Сходные результаты были получены для некоторых сельскохозяйственных культур.

Ключевые слова: мейоз; ДНК; репарация; рекомбинация; кроссинговер; селекция.

Introduction

Meiosis is a cell division underlying sexual reproduction, which allows species to maintain a stable set of chromosomes in a number of generations due to the proper segregation of homologous chromosomes in prophase I of meiosis. At the same time, meiosis is also a source of genetic variability, which forms due to the recombination of whole chromosomes, the exchange of regions between homologous chromosomes during crossing over, and conversion events between unpaired nucleotide bases at the site of repair of programmed double-strand breaks (DSBs) in DNA (Mercier et al., 2015).

The entire traditional scheme of selective breeding of plant varieties and hybrids is based on the use of meiotic crossing over as the main mechanism for creating chromosomes with new combinations of alleles that are transmitted to offspring (Zhuchenko, Korol, 1985). The indispensability of meiotic crossing over for selective breeding is evident in the introgression of individual economically valuable genes from the chromosomes of wild species into the chromosomes of cultivated plants (De Muylt et al., 2009). Gaining control over the distribution of meiotic crossing over points and over the frequency of crossover exchanges will allow to construct a new allelic composition of chromosomes more effectively (Wijnker, de Jong, 2008; Lambing et al., 2017; Blary, Jenczewski, 2019).

The relevance of the study of meiotic crossing over is emphasized by many scientific reviews devoted to general issues of meiosis (Kleckner, 1996; Harrison et al., 2010; Osman et al., 2011; Crismani et al., 2013), meiotic recombination (Mézard et al., 2007; De Muylt et al., 2009; Gray, Cohen, 2016; Blary, Jenczewski, 2019; Bogdanov, Grishaeva, 2020), metabolic pathways, and crossover mechanisms (Mézard et al., 2007, 2015; De Muylt et al., 2009; Mercier et al., 2015; Gray, Cohen, 2016; Wang, Copenhaver, 2018; Bogdanov, Grishaeva, 2020), genetic control of meiotic division (Mercier et al., 2015; Gray, Cohen, 2016; Simanovsky, Bogdanov, 2018), identification and functional analysis of genes involved in meiosis (Mercier et al., 2015), epigenetic control over meiotic recombination (Yelina et al., 2015; Taagen et al., 2020), and the effect of ploidy on meiotic recombination (De Muylt et al., 2009; Lambing et al., 2017). Unlike previously published review articles, this review is devoted to practical issues of controlling the frequency and distribution of crossover exchanges between homologous chromosomes during meiosis in plants.

The role of meiotic crossing over in evolution and selection

Modern views on the molecular mechanisms of crossing over in meiosis are detailed in a number of previously published scientific literature reviews (Mercier et al., 2015; Mézard et al., 2015; Gray, Cohen, 2016; Wang, Copenhaver, 2018). The historical retrospective of the development of the theory of meiotic recombination of chromosomal DNA based on DSB repair and the experimental discovery of the “core” set of proteins: SPO11, RAD51, ZMM complex and others responsible for meiotic crossing over in most eukaryotes is presented in detail in a recent review by Yu.F. Bogdanov and T.M. Grishaeva (2020). Therefore, in this article, let us briefly note that meiotic crossing over is controlled by meiosis-specific genes,

namely, meiotic recombination genes (Youds, Boulton, 2011; Bogdanov, Grishaeva, 2020). These genes are usually suppressed in somatic cells that divide by mitosis. The transition of diploid cells from division by mitosis to division by meiosis occurs as a result of acts of negative regulation, since the genes initiating meiosis turn off the genetic program of mitosis, and then the previously silent genetic program of meiosis is turned on (Turner, 2007; Bogdanov, Grishaeva, 2020).

In flowering plants, meiotic crossing over occurs in specialized cells (microsporocytes and megasporocytes) and consists of successive processes, including the creation of programmed DSBs in DNA and their repair by the mechanism of homologous recombination with the preferred use of a homologous chromosome as a template. Crossover products are crossover chromosomes that carry new combinations of allelic variants of genes (Mieulet et al., 2018).

The site of crossover exchange manifests itself in the form of a crossing between chromosomes observable under a microscope, called chiasm. In addition to crossover exchange, chiasmata also perform a structural or mechanical function: they hold chromosomes in the form of bivalents during the prophase and metaphase I of meiosis; as a result, it is the bivalents, and not individual chromosomes, that line up on the equator of the division spindle in metaphase I, which creates an opportunity for segregation of homologous chromosomes in the first division of meiosis and cell haploidization (Bogdanov, Grishaeva, 2020).

It is known that many cellular proteins required for DSB repair in the DNA of somatic cells are also involved in meiotic crossing over. However, their functions may change depending on localization, post-translational modifications, and/or interactions with meiosis-specific proteins (Villeneuve, Hillers, 2001). This means that at some point in evolution, crossing over deviated from the function of repairing only somatic damage and acquired a function specific to meiosis. Therefore, meiotic crossing over is an evolutionary adaptation of somatic repair functions for successful sexual reproduction during the transition from the diploid to haploid phase of the life cycle. It is assumed that the evolutionarily established parameters of meiotic crossing over are regulated by natural selection to maintain the maximum adaptability of organisms to changing environmental conditions in a series of sexual generations (Zhuchenko, Korol, 1985; Wijnker, de Jong, 2008). From this point of view, the parameters of the evolutionarily established mechanism of meiotic crossing over can limit the selection process, which requires the creation of maximum genetic diversity among sexual offspring, even to the detriment of adaptability to natural habitat conditions.

The frequency of occurrence of crossovers and their distribution along chromosomes are the determining factors promoting the new and selectable genetic variability in meiosis (Zhuchenko, Korol, 1985; Wijnker, de Jong, 2008). Many studies have shown that the position of meiotic crossovers along chromosomes is non-random, strictly predetermined, non-uniform, and does not depend on genome size (Mercier et al., 2015; Mézard et al., 2015; Lambing et al., 2017; Wang, Copenhaver, 2018; Blary, Jenczewski, 2019).

According to widespread belief, there are three conceptual levels of regulation of the frequency and distribution of cross-

overs: obligatory crossover exchange in bivalent (crossover insurance), crossover interference, and crossover homeostasis (Simanovsky, Bogdanov, 2018; Bogdanov, Grishaeva, 2020). A possible reason for the obligatory crossover exchange in each bivalent is the mechanical function of the chiasmata (Roeder, 1997). The interference provides a nonrandom distribution of crossovers along the chromosome and their location at a greater distance from one another than is expected in case of a random distribution if there is more than one crossover per bivalent (Jones, Franklin, 2006).

Crossover homeostasis is the ability of meiotic cells to maintain the level of the number of crossovers per chromosome inherent in a given biological species, even if the number of DSBs decreases by an order of magnitude (Bogdanov, Grishaeva, 2020). In particular, in budding yeast (*Saccharomyces cerevisiae* (Desm.) Meyen ex E.C. Hansen), a fivefold decrease in the number of DSBs per cell was achieved, but the number of crossovers remained normal and unchanged (Martini et al., 2006). In the latter case, it is unclear whether this rule is true for all organisms. For example, in male house mice (*Mus musculus*), a 20–50 % decrease in the number of DSBs does not lead to a decrease in the number of crossovers and fertility (Cole et al., 2012a). However, a 60 % decrease in the number of DSBs already provokes homologue asynapsis and sterility in male mice (Kauppi et al., 2013).

Thus, in selective breeding practice, the relatively low frequency of meiotic crossovers and their determinism along chromosomes lead to the necessity for the analysis of large populations in order to identify rare recombinant genotypes that combine the desired economically valuable genes. It has also been argued that regions of chromosomes that are rarely used for crossover create additional problems for breeders, as deleterious mutations accumulate in regions of low recombination (Rodgers-Melnick et al., 2015).

Stimulation of meiotic crossing over at the stage of creation of DNA double-strand breaks

Numerous and genetically programmed DSBs in DNA molecules are precursors of mutual genetic exchange between homologous chromosomes in prophase I of meiosis. During prophase I of meiosis, hundreds of DSBs are created along the chromosomes, generated by the evolutionarily conservative endonuclease SPO11 and some associated proteins (Keeney et al., 1997; Wang, Copenhaver, 2018).

SPO11 genes have been described in all eukaryotes whose genomes have been studied and whose protein products are similar to the A subunit of archaeal DNA topoisomerase VI (Nichols et al., 1999; Hartung, Puchta, 2001; Wu et al., 2004). In yeast, insects, and vertebrates, the *SPO11* gene is represented by one copy, while the plant genome has three copies of *SPO11* (Hartung, Puchta, 2001; Stacey et al., 2006). In *Arabidopsis thaliana* L., the *SPO11-1* and *SPO11-2* genes are required for meiotic recombination (Grelon et al., 2001; Stacey et al., 2006), while *SPO11-3* is involved in somatic endoreplication (Hartung et al., 2002). It was previously shown that the B subunit of DNA topoisomerase VI is also required for the full formation of meiotic DSBs (Robert et al., 2016; Vrielynck et al., 2016). It forms a complex with two

orthologues SPO11-1 and SPO11-2, and is absolutely necessary for the formation of the SPO11-1/SPO11-2 heterodimer in *Arabidopsis* (Vrielynck et al., 2016).

It is assumed that the mechanism of meiotic DSB formation with the participation of SPO11 proteins is conservative, but its regulation may differ due to a set of auxiliary proteins that are less conservative in organisms from different kingdoms (De Muyt et al., 2009).

The distribution of DSBs can be considered as the first possible level of determination of future crossover formation sites. In particular, in budding yeast, from 160 to 200 DSB points are formed in each cell, and the repair of most of them leads to crossing over (Mercier et al., 2015). In other organisms, the number of DSB points significantly exceeds the number of crossing over points. For example, in *Arabidopsis*, there are from 150 to 300 DSB points producing about 10 crossover exchanges per genome (Kurzbaue et al., 2012; Choi et al., 2013); in maize (*Zea mays* L.), almost 500 DSB points lead to the formation of about 20 crossover exchanges (Anderson et al., 2003). As a consequence, the site of crossover realization is selected from a wide range of potential DSB formation sites in the genome.

There is an opinion that the “hot spots” of crossing over are closely associated with the “hot spots” for the formation of DSB (Wang, Copenhaver, 2018). High-resolution analysis in budding yeast shows that crossover hotspots and meiotic DSBs are concentrated in promoter-adjacent regions with low nucleosome density (Mancera et al., 2008). In humans (*Homo sapiens*) and house mice, “hot spots” of crossing over occur in certain DNA sequences bound by the PRDM9 protein in gene and intergenic regions not associated with transcription initiation sites (Baudat et al., 2010; Smagulova et al., 2011).

PRDM9 is a DNA-binding protein that catalyzes H3 histone (H3K4 modification) methylation, which initiates the formation of DSBs away from transcription start sites (Baudat et al., 2010; Smagulova et al., 2011). Plants do not have a PRDM9 homologue, but crossover “hot spots” do exist. “Hot spots” in *Arabidopsis* are characterized by the fact that crossing over in them occurs up to 50 times more actively than on average for the genome (Choi et al., 2013; Yelina et al., 2015). At the same time, crossover “hot spots” could be associated with active transcription of RNA polymerase II, low nucleosome density, low DNA methylation level, as well as with intergenic regions, promoters, transcription start and stop sites, transposons, or insertion-deletion regions. The search for DNA motifs associated with crossover hotspots in *Arabidopsis* revealed their significant enrichment in CTT, CCN, and poly A sequences (Choi et al., 2013; Wijnker et al., 2013).

Chromosomal regions associated with an increased frequency of crossing over have also been found in maize (Liu et al., 2009), wheat (*Triticum aestivum* L.) (Santenac et al., 2009, 2011), and cultivated tomato (*Solanum lycopersicum* L.) (Demirci et al., 2017), which indicates the preservation of common mechanisms in different plant species. In most organisms, DSBs can occur along the entire length of chromosomes, however, it is surprising that 80 % of crossing over points are concentrated in about 25 % of genome regions (Blary, Jenczewski, 2019). For example, 82 % of crossovers are concentrated at the distal ends of wheat chromosome 3B,

which is 19 % of its total length (Darrier et al., 2017). Therefore, despite a significant amount of species-specific information, it is currently impossible to identify a general pattern, a common or key factor in the localization of all DSBs in the genome (Bogdanov, Grishaeva, 2020). The success of practical selection can be directly related to the expansion of the range of genome regions liable to crossing over in ways including the creation of additional DSBs in regions that are rarely used to initiate DSBs or the redistribution of DSB regions.

It is known that in *spo11Δ*-mutant budding yeast lacking their own functional *SPO11* alleles, expression of the chimeric *GAL4BD-SPO11* gene initiated additional DSBs at the binding site of the Gal4 protein (Peciña et al., 2002). Later, it was shown that SPO11 chimeric proteins fused with various DNA-binding protein modules (transcription factors, Cas9 nuclease, etc.) can stimulate crossing over in regions of the yeast genome with low natural recombination activity (Sarno et al., 2017). In the latter case, the authors propose their own strategy for increasing the genetic variability of gametes in plant breeding.

However, it is difficult to use higher organisms with a knockout of their own *SPO11* genes in selective breeding work. In arabidopsis, mutations in the *SPO11-1* gene lead to a complete loss of synapsis of homologues in prophase I and their random segregation, a formation of a significant level of nonfunctional gametes, and a decrease in meiotic recombination by an order of magnitude (Grelon et al., 2001). In mice, the *Spo11^{-/-}* genotype with a complete absence of DSB demonstrates chromosome asynapsis and sterility (Baudat et al., 2010). Expression of the recombinant isoform of the mouse's own *Spo11β* gene made it possible to prove that the SPO11 protein level is crucial for chromosome synapsis and successful completion of meiosis (Kauppi et al., 2013). In the *mei-W68¹ (spo11)* mutants of the drosophila fly (*Drosophila melanogaster*), expression of the native *SPO11* gene restores the wild-type phenotype (Shingu et al., 2012). Expression of the arabidopsis *SPO11-1* and *SPO11-2* genes or rice (*Oryza sativa* L.) *SPO11A*, *SPO11B*, and *SPO11D* genes leads to an increase in the amount of DSB in *mei-W68¹* mutants, but this is not enough for the normal completion of meiosis (Shingu et al., 2012). The totality of the presented results shows that in higher organisms within the framework of the proposed strategy, probably, only overexpression of recombinant *SPO11* genes could become a way of redistributing exchanges between homologous chromosomes.

Previously, to test this assumption, transgenic tomato plants that express the *SPO11* genes from budding yeast or arabidopsis under the control of a strong constitutive 35S CaMV viral promoter were created (Komakhina et al., 2020). Using genetic analysis, it was shown that overexpression of both recombinant *SPO11* genes partially disrupts the monogenic inheritance of marker alleles of the *Wv:wv* locus of chromosome 2 among tomato offspring. Segregation disruption at the *Wv:wv* locus could be the result of gene conversion due to the preferential formation of DSB in one of the *Wv* or *wv* alleles in transgenic plants. Overexpression of the *SPO11* genes reduced the frequency of meiotic recombination in the region between the *wv* and *d* genes of tomato chromosome 2 by 17–18 % compared to the non-transgenic control. At the

same time, a negative correlation was found between the expression level of the recombinant *SPO11* genes and the frequency of recombination in the analyzed *wv-d* region of chromosome 2.

Unfortunately, the effect of the expression of recombinant *SPO11* genes on the frequency of meiotic recombination in other regions of the tomato genome remained unexplored. In general, it has been shown that the strategy of meiotic recombination induction using additional SPO11 activity, previously successfully implemented in yeast, may have limitations in plant (Komakhina et al., 2020) and insect cells (Shingu et al., 2012).

Later, a debatable opinion that DSB “hot spots” do not necessarily become crossover “hot spots” was expressed. This opinion is substantiated by the fact that a small absolute number of DSBs in “cold regions” can paradoxically turn into a relatively high frequency of realized crossing over (Bogdanov, Grishaeva, 2020).

Stimulation of meiotic crossing over at the stage of homology search during repair of DNA double-strand breaks

During meiosis, DSBs resulting from the activity of SPO11 endonuclease are processed to 3'-single-stranded DNA ends, which then cooperatively bind RecA-like recombinases RAD51 and DMC1 (Brown, Bishop, 2014; Mercier et al., 2015). As a result, nucleoprotein filaments are formed that carry out single end invasion into the sister chromatid or the homologous chromosome (Girard et al., 2015). The 3'-single-stranded DNA ends invading the double-stranded DNA molecule are then elongated by DNA synthesis and ligation, which leads to the formation of a D-loop (displacement loop), from which a double Holliday junction is then formed (Brown, Bishop, 2014; Wang, Copenhagen, 2018).

During meiosis, DSB repair can shift towards predominant use of the homologous chromosome as a template, a process called interhomolog bias (Brown, Bishop, 2014). This process is a prerequisite for crossing over between homologous chromosomes and requires the involvement of a specific meiotic mechanism that prevents sister chromatids from being used for repair (Brown, Bishop, 2014). In particular, in arabidopsis, the meiosis-specific DMC1 protein is presumably responsible for the increased probability of DSB repair using a homologous chromosome (Kurzbaue et al., 2012).

In budding yeast and arabidopsis during meiosis, the catalytic activity of RAD51 is not necessary for the formation of interhomologous crossing over that confirms the preferential role of the DMC1 protein in this process (Cloud et al., 2012; Da Ines et al., 2013). In arabidopsis, the RAD51 protein functions within a backup pathway for DSB recovery during meiosis in the event of DMC1 dysfunction (Kurzbaue et al., 2012). In the absence of DMC1, meiotic DSBs are restored by the RAD51 protein using the sister chromatid as a template, which leads to the absence of synapsis between homologues and the appearance of univalents (Couteau et al., 1999). The presence of the DMC1 protein suppresses RAD51 activity in arabidopsis (Uanschou et al., 2013); the same is observed in meiosis in budding yeast, in which the DMC1 protein suppresses RAD51 activity (Lao et al., 2013). At the same time,

in the *figl1* mutants of *Arabidopsis*, which exhibit an increased frequency of crossover exchanges, a twofold increase in the number of RAD51 foci was found in cells at the leptotene/zygotene stages of meiosis, while the number of foci of meiosis-specific DMC1 did not change or increased insignificantly (Girard et al., 2015; Fernandes et al., 2018a). Recent results do not rule out that the role of RAD51 recombinase in meiotic crossing over may be somewhat wider than commonly believed.

It is currently assumed that the choice in favor of crossing over or its absence is made during DSB processing and before the formation of the double Holliday structure (Hunter, Kleckner, 2001; Bogdanov, Grishaeva, 2020). The molecular mechanism that makes this choice continues to be discussed, but the fact of early choice is considered established (Hunter, Kleckner, 2001; Bishop, Zickler, 2004; Youds, Boulton, 2011; Gray, Cohen, 2016; Bogdanov, Grishaeva, 2020).

Structural and biochemical differences between RAD51 and DMC1 proteins are not very large (Sheridan et al., 2008). However, a large number of protein factors have been found that are required for proper loading, stabilization, and/or activation of these eukaryotic recombinases (Mercier et al., 2015).

It is known that bacterial recombinase RecA has 40 to 60 % homology with eukaryotic recombinases, but unlike them, it is universal and capable of different and even unique functions without the participation of helper proteins and with greater efficiency (Baumann, West, 1998; Lanzov, 2007). It was shown that the expression of the *recA* gene from *Escherichia coli* triples the number of DSBs restored by the mechanism of homologous recombination and more than doubles the number of sister chromatid exchanges in the somatic cells of tobacco plants (*Nicotiana tabacum* L.) (Reiss et al., 1996, 2000). This suggested that the expression of the *recA* gene in plant cells in prophase I of meiosis can also change the number and distribution of crossover exchanges between homologous chromosomes (Komakhin et al., 2010). It was later shown that the expression of the *recA* gene from *E. coli* under the control of a strong and constitutive CaMV35S promoter in cultivated tomato leads to an increase in the frequency of meiotic recombination between the *wv* and *d* genes of chromosome 2 by 50 % compared with the non-transgenic control (Komakhin et al., 2012).

The molecular mechanism that allowed to increase the frequency of meiotic recombination in the transgenic tomato remained unclear at the time these results were published. Later, it became known that the yeast Top3 topoisomerase negatively affects meiotic crossing over since it specifically destroys the D-loops formed by the yeast Rad51/Rad54 proteins (Fasching et al., 2015). However, D-loops formed by the bacterial RecA protein proved to be resistant to destruction by the Top3 protein. It has also been found that *Arabidopsis* plants carrying *top3a* mutant alleles show a 1.5 to 2.5-fold increase in meiotic recombination frequency (Séguéla-Arnaud et al., 2015). Probably, in transgenic tomato plants expressing the *recA* gene, an increase in the frequency of meiotic recombination could be due to the formation of D-loops by the bacterial RecA protein, which could not be destroyed by the tomato TOP3 α protein, resulting in an increase of the recombination frequency.

An attempt to apply this experimental approach to increase crossover exchanges between chromosomes of different tomato species showed an ambiguous result (Komakhin et al., 2019). In particular, none of the three combinations of crossing a cultivated tomato expressing the *recA* gene and wild tomato species *S. cheesmaniae*, *S. pimpinellifolium*, and *S. habrochaites* showed a significant increase in the frequency of recombination between the marker genes of chromosome 2. It is assumed that the factor limiting recombination between chromosomes from different species is the mismatch repair system, which eliminates mismatched bases in DNA at the DSB repair site (Chambers et al., 1996; Emmanuel et al., 2006; Strelnikova et al., 2021). This assumption is based on the fact that in interspecific tomato hybrids, due to the increased level of DNA polymorphism between chromosomes of different species, one should expect a more active resistance of mismatch repair to meiotic crossing over than in interline hybrids of cultivated tomato.

Stimulation of meiotic crossing over at the stage of correction of unpaired bases at the site of DNA double-strand breaks repair

During meiotic crossing over between homologous chromosomes, regions of heteroduplex DNA containing unpaired bases can arise locally. The mismatch repair system eliminates these regions.

The mismatch repair system is a highly conservative way of maintaining DNA integrity that exists in all organisms. The first step of this pathway in eukaryotes, mismatch recognition, is performed by homologues of prokaryotic MutS proteins, viz. MSH proteins. Eight of them were described in eukaryotes, from MSH1 to MSH8. MSH7 is found only in plants (Culligan, Hays, 2000), while MSH8 is found in the phylum Euglenozoa (Sachadyn, 2010). MSH proteins recognize unpaired bases as heterodimers. The heterodimer designated MutSa (MSH2-MSH6) repairs mismatches or 1–2 nucleotide loops (Acharya et al., 1996; Genschel et al., 1998). The MutSb heterodimer (MSH2-MSH3) recognizes larger loops containing up to 14 nucleotides (Modrich, 1991; Marti et al., 2002). Plants form an additional heterodimeric complex known as MutSc (MSH2-MSH7) (Culligan, Hays, 2000), which is involved in meiotic recombination (Lloyd et al., 2007). In meiosis, the mismatch repair system is able to destroy heteroduplex DNA and suppress crossing over (Cole et al., 2012b).

Inactivation of the *MSH2* gene in interspecific yeast hybrids *S. cerevisiae* \times *S. paradoxus* increases the recombination frequency between homeologous chromosomes up to 5.5 times and also increases the viability of spores (Hunter et al., 1996). In *Arabidopsis* plants, knockout of the *MSH2* gene (mutation *msh2-1*) increases microsatellite instability and somatic recombination, which indicates a decrease in the efficiency of the mismatch repair system in plant cells (Leonard et al., 2003). In another study, it was shown that the *msh2-1* mutation increased by 40 % the frequency of meiotic recombination between marker genes of fluorescent proteins in an isogenic background of *Arabidopsis* (*Landsberg erecta* ecotype) (Emmanuel et al., 2006).

These results allowed to apply the strategy of suppressing mismatch repair by inhibiting the expression of the *MSH2* and

MSH7 genes to increase the frequency of crossover exchanges in other plant species. In particular, in cultivated tomato, the inhibition of the expression of the *MSH2* and *MSH7* genes was performed by three independent scientific groups at different times either using RNA interference (RNAi) (Tam et al., 2011; Sarma et al., 2018; Strelnikova et al., 2021) or using a dominant-negative construct with the mutant *MSH2-DN2* protein gene from *Arabidopsis* (Tam et al., 2011).

The use of a dominant-negative construct or inhibition of the *MSH7* transcript by RNAi allowed a non-substantial increase by 17.8 % in the frequency of meiotic recombination between homeologues in a cultivated tomato heterozygous by chromosome 8 from *S. lycopersicoides* Dunal (Tam et al., 2011). At the same time, silencing of the *MSH2* gene transcript with RNAi delivered pronounced negative consequences for the fertility of tomato plants (Sarma et al., 2018; Strelnikova et al., 2021), especially when using the strong pro-SmAMP2 plant promoter to control the expression of the RNAi construct (Strelnikova et al., 2021). In recent experiments, it was convincingly shown that the highly effective RNAi of the *MSH2* gene leads to phenotypic anomalies in cultivated tomato plants: growth and flowering retardation and formation of a reduced number of seeds (Sarma et al., 2018; Strelnikova et al., 2021). In cases where the RNAi of the *MSH2* gene was moderate, tomato plants were fertile, but no increase in the frequency of meiotic recombination was found (Tam et al., 2011; Strelnikova et al., 2021).

These results show that in tomato plants, in contrast to *Arabidopsis* plants, suppression of the *MSH2* gene by RNAi to increase the frequency of meiotic recombination has significant limitations. Probably, there is a certain level of expression of the *MSH2* gene, which is critical for the viability of tomato plants. This may be due to the fact that, in contrast to *Arabidopsis* plants, the *MSH2* gene in tomato performs an additional cellular function necessary for plant fertility. This may be the reason why spontaneous or induced *msh2* mutants have not yet been described among various tomato species.

It should be noted that the repression of mismatch repair has a negative effect on the stability of the genome and the reproductive properties of many other plant species besides tomato. It was shown that a knockout mutation of the *MSH2* gene in *Arabidopsis* plants after several generations led to an intensive accumulation of various mutations in the genome, a partial loss of fertility, and a decrease in the number of seeds (Leonard et al., 2003; Hoffman et al., 2004). Another study showed that a mutation in the *MLH1* gene, which is also a part of the mismatch repair system, leads to reproductive defects in *Arabidopsis* plants (Dion et al., 2007).

Inhibition of *MSH2* gene expression using two different strategies leads to numerous phenotypic anomalies and microsatellite instability in somatic potato hybrids (Rakosy-Tican et al., 2019). RNAi of the *MSH7* gene in transgenic barley plants (*Hordeum vulgare* L.) leads to a decrease in the number of seeds and pollen viability (Lloyd et al., 2007). In wheat plants, the *msh7-3D* mutation also reduces pollen viability but does not affect plant fertility (Serra et al., 2018). Overall, these results confirm that the strategy of stimulating meiotic recombination by suppressing mismatch repair in different plant species can lead to impaired reproductive functions.

Stimulation of meiotic recombination at the stage of D-loop resolution

As already mentioned in the section “Stimulation of meiotic crossing over at the stage of creation of DNA double-strand breaks”, in most eukaryotes, the number of DSBs significantly exceeds the number of crossovers. This suggests that there are negatively acting metabolic mechanisms that prevent the resolution of part of the DSBs through the crossover pathway.

The choice of the DSB repair mechanism in favor of a crossover or non-crossover pathway occurs at the early stages of DSB repair, when a single-stranded DNA-protein filament invades a homologous DNA molecule and causes the formation of a D-loop in it (Hunter, Kleckner, 2001; Bishop, Zickler, 2004; Bogdanov, Grishaeva, 2020). D-loops that arise during the homology search step using RAD51 and DMC1 recombinases can be transformed via various metabolic pathways, leading either to crossovers between homologous chromosomes or to non-reciprocal exchange (without crossing over) between them.

Currently, two ways of crossing over implementation, leading to the appearance of either class I or class II crossovers, are most fully described (Gray, Cohen, 2016). Class I crossovers are products of the activity of a group of proteins collectively referred to as ZMM (Zip1, Zip2, Zip3, Zip4, Msh4 and Msh5, Mer3) that stabilize intermediate D-loops, promoting the formation of a double Holliday structure (Hunter, 2015). The MLH1 and MLH3 proteins in combination with EXO1 promote the transformation of the Holliday structure into class I crossovers (Ranjha et al., 2014).

Class I crossovers are not randomly distributed along chromosomes, as they reduce the likelihood of adjacent crossovers in close proximity of them (Wang et al., 2015). This phenomenon is commonly referred to as interference. In addition, the D-loops (as a recombination intermediate) can be converted by structure-specific endonucleases, including the MUS81 enzyme, producing class II crossovers that are not subject to interference (Berchowitz et al., 2007; Wang, Copenhagen, 2018; Bogdanov, Grishaeva, 2020). There are known double mutants of *Arabidopsis* at both *msh4* and *mus81* genes which control crossing over pathways I and II, respectively; despite this, these mutants show a residual 5–10 % of crossovers (Higgins et al., 2008). However, the mechanism that generates these residual crossovers is unclear; possibly, it is active only when the main crossover pathways I and II are disrupted (Osman et al., 2011; Mercier et al., 2015; Lambing et al., 2017).

There are organisms in which only one of the two major pathways of crossover formation is present. In particular, in fission yeast (*Schizosaccharomyces pombe* Lindner) and mold (*Aspergillus nidulans* P. Michel ex Haller) only pathway II is present, which is not susceptible to interference. In contrast, in the soil nematode (*Caenorhabditis elegans* Dougherty), only interfering pathway I is known. In plants, both crossing over pathways were found, but in different proportions. For example, in *Arabidopsis* and tomato, class I crossovers amount to 70 to 90 %, and the rest belong to class II (Lhuissier et al., 2007; Higgins et al., 2008; Macaisne et al., 2011; Anderson et al., 2014).

Recently, using molecular genetic studies, arabidopsis was found to contain protein factors that act against the conversion of DSBs into class II crossovers: DNA helicase FANCM (Fanconi anemia complementation group M) (Crismani et al., 2012; Girard et al., 2014), FIGL1 (AAA-ATPase FIDGETIN-LIKE1) (Fasching et al., 2015; Girard et al., 2015), BTR complex of DNA helicases RECQ4A and RECQ4B and topoisomerase TOP3 α (Séguéla-Arnaud et al., 2015).

Previously, the *At1g35530* gene was found in the arabidopsis genome, the mutation in which allows to suppress the *zip4(s)1* and *msh5* mutations associated with disturbances in meiotic division (Crismani et al., 2012). It turned out that the *At1g35530* gene encodes a DNA helicase homologous to the human FANCM helicase. In yeast, FANCM orthologues and their cofactors form a conservative complex involved in the formation of non-crossover products during meiosis through disruption of D-loops (Gari et al., 2008). Arabidopsis plants with the mutant *fancm* gene demonstrate an increase in the frequency of meiotic recombination from 2 to 3.6 times in all eight studied genome regions and are indistinguishable from wild-type plants in terms of growth and fertility (Crismani et al., 2012). Additional crossovers are independent of ZMM proteins and occur via the MUS81 pathway typical to class II (Crismani et al., 2012; Girard et al., 2014).

Thus, it was demonstrated for the first time that *FANCM* in plants is a strong negative regulator of crossing over. However, subsequent studies showed that the *fancm* mutation was effective only in arabidopsis inbred lines of the Columbia-0 or Landsberg erecta ecotypes; in hybrids of the Columbia-0 \times Landsberg erecta combination, the *fancm* mutation does not increase the frequency of meiotic recombination (Girard et al., 2015). In addition, the *fancm* mutation effectively restores the formation of bivalents in the *zmm* mutants in the Columbia-0, Landsberg erecta, or Wassilewskija inbred lines, but not in the Columbia-0 \times Landsberg erecta and Columbia-0 \times Wassilewskija hybrids.

Later, in arabidopsis plants, the conservative FIDGETIN-LIKE1 AAA-ATPase (FIGL1) was identified, which also acted as a negative regulator of crossover formation (Girard et al., 2015). It is known that FIGL1 belongs to the FIDGETIN subfamily and is involved in DNA repair (Yuan, Chen, 2013). In arabidopsis, FIGL1, like FANCM (Crismani et al., 2012), limits the formation of crossovers across the entire genome (Girard et al., 2015). In particular, in a single arabidopsis *figl1-1* mutant, the frequency of meiotic recombination increased in each of the six tested sites by an average of 72 % (in single *fancm-1* mutants the frequency of meiotic recombination on average tripled (Crismani et al., 2012)) and a noticeable increase in the frequency of recombination took place in the distal regions of chromosomes. A six-fold increase in the frequency of meiotic recombination was found in arabidopsis *figl1-1 fancm-1* double mutants compared to wild-type plants in six tested genome regions, while maintaining the progression of meiotic division and fecundity (Girard et al., 2015).

Recent results indicate that the effects of the *figl1-1* and *fancm-1* mutations are synergistic, thus affecting different metabolic pathways to limit crossing over. It was also found that two *figl1 fancm* mutations in Columbia-0 \times Landsberg

erecta hybrids resulted in a 2.5-fold increase in meiotic recombination frequency in the four sites tested compared to wild-type hybrids. This was higher than either of the *figl1* or *fancm* mutants alone (1.8 and 1.2 times, respectively), confirming that *figl1* and *fancm* have a multiplicative effect also in the hybrid genetic environment.

It is assumed that in arabidopsis, the FIGL1 protein negatively affects the dynamics of two conservative recombinases DMC1 and RAD51, counteracting the invasion of single-stranded DNA ends into the homologous chromosome, and thus prevents the interaction between homologous chromosomes (Girard et al., 2015). The available data allow us to conclude that FIGL1 and FANCM represent two sequential barriers against crossing over, the first of which limits the invasion of DNA strands into the homologous chromosome, and the second, due to helicase activity, unwinds intermediate DNA structures that arise during the formation of the D-loop (Girard et al., 2015). This model is supported by direct evidence for physical interaction of the FIGL1 protein via its FRBD domain with RAD51 and DMC1 proteins and an increase in DMC1 foci in arabidopsis *figl1* mutants (Fernandes et al., 2018a).

The complex of BTR (BLOOM-TOP3-RMI1-RMI2) in humans and Sgs1-Top3-Rmi1 in budding yeast is highly conservative and plays a major role in the formation of non-crossover products by resolving the double Holliday structure or by disrupting D-loops (Fasching et al., 2015). In particular, during the reaction, two Holliday structures migrate towards each other using the BLOOM/Sgs1 helicase. The structure thus generated is then removed using the TOP3 α /Top3 topoisomerase and its cofactors (Berchowitz et al., 2007; Higgins et al., 2008; Macaisne et al., 2011). The same protein complex promotes D-loop unwinding, which results in the formation of exclusively non-crossover products (Crismani et al., 2012; Girard et al., 2014; Mercier et al., 2015).

In arabidopsis, the genome contains three members of the BTR complex: TOP3 α and RMI1 as single genes, and the Sgs1 homologue as two paralogous genes, RECQ4A and RECQ4B (Séguéla-Arnaud et al., 2015, 2017). In a recent study, arabidopsis plants carrying different *top3a* mutant alleles were shown to make a 1.5 to 2.5-fold increase in meiotic recombination frequency. Arabidopsis *recq4a recq4b* double mutants show a 6.2-fold increase in meiotic recombination frequency compared to wild-type plants (Séguéla-Arnaud et al., 2015). Moreover, the increase in frequency occurred due to the appearance of class II crossovers that are not subject to interference. The effects of the *top3a* and *recq4a recq4b* mutations in arabidopsis were enhanced against the background of the *fancm* mutation. Compared with wild-type plants, the frequency of meiotic recombination increases on average by 4.8 times in the *top3a fancm* double mutant and by 9 times in the *recq4a recq4b fancm* triple mutant. These results allowed the authors to state that there are at least two independent pathways for the negative regulation of crossing over in arabidopsis. From the point of view of selective breeding studies, it was important that, despite a significant increase in recombination, the *top3a fancm* and *recq4a recq4b fancm* mutants grew normally, were fully fertile and did not show defects in meiotic division (Séguéla-Arnaud et al., 2015, 2017).

The same authors (Fernandes et al., 2018a) showed that when the *recq4* mutation and the *figl1* mutation are combined in one arabidopsis plant, the frequency of crossing over increases by 7.8 times and the genetic map lengthens from 389 to 3037 cM. It has also been shown that the increase in the number of crossing over events occurs unevenly along the chromosomes and increases from the centromere to the telomere. Finally, female recombination was higher than male recombination in the *recq4 figl1* double mutant (3200 versus 2720 cM), although in wild-type plants recombination in male meiosis is much higher than in female meiosis (490 versus 290 cM). These results suggest that the factors that make female meiosis less recombinogenic than male meiosis do not operate in the context of this double mutant.

Almost simultaneously with studies of the role of *FANCM* in arabidopsis plants, an attempt was made to increase the frequency of meiotic recombination in its close relatives of agricultural importance: diploid turnip plants (*Brassica rapa* L.) and tetraploid rapeseed plants (*B. napus* L.) (Blary et al., 2018). In this work, it was found that the *braA.fanm-1* missense mutation in the turnip *BraA.FANCM* gene is able to partially complement the *braA.msh4-1* mutation in the meiosis-specific *BraA.MSH4* gene of the turnip, which in turn reduces the number of bivalents in metaphase and gives rise to univalents.

Turnip double mutants *braA.fanm-1 braA.msh4-1* show a 3-fold increase in the number of crossovers, equal to the increase previously observed in arabidopsis (Crismani et al., 2012). In rape plant mutants carrying the *bnA.fanm-1* nonsense mutation in the A genome and the *bnC.fanm-1* or *bnC.fanm-2* missense mutation in the C genome, a certain increase (1.3 times) in the frequency of meiotic recombination was observed. The authors attribute this result to the residual activity of *FANCM* mutant variants from the C genome in tetraploid rapeseed plants (AACC genome) (Blary et al., 2018).

Also, the influence of *FANCM*, *RECQ4* and *FIGL1* factors on the frequency of meiotic recombination was studied in other important agricultural crops: rice, pea (*Pisum sativum* L.) and cultivated tomato (Mieulet et al., 2018). Mutations in the *recq4* orthologue genes increase the frequency of meiotic recombination from 2.7 to 3.7 times in all studied plant species. Mutations in *fanm* orthologue genes slightly increase the frequency of meiotic recombination, from 1.6 to 2.3 times in pea and rice, but not in tomato, which showed no changes. It was shown that in lettuce (*Lactuca sativa* L.), knockout of the *FANCM* gene orthologue using CRISPR/Cas9 genome editing leads to a decrease in the viability of pollen and a decrease in the number of seeds (Li et al., 2021). In lettuce *fanm* mutants, 78 % of meiocytes in metaphase I have univalents. These results indicate that *FANCM* in lettuce, in contrast to arabidopsis plants, likely has an additional function in meiosis. Notably, homozygous knockout of *figl1* orthologs in tomato, pea, and rice plants induces sterility (Zhang et al., 2017; Mieulet et al., 2018).

Thus, the *recq4* mutation increases the frequency of crossing over by about 3 times in all studied crops (rice, pea, and tomato), so manipulation of the *RECQ4* gene may be a ver-

satile tool to increase meiotic recombination in plants. However, the presented results also indicate that the meiotic effects found in the model object are not always reproduced in agricultural crops.

It has been shown that the frequency of interfering class I crossovers in arabidopsis can be influenced by overexpression of the *HEI10* gene (an analogue of *Human Enhancer of Invasion 10*), which encodes a meiosis-specific E3 ligase associated with quantitative variation in the frequency of crossing over between arabidopsis ecotypes (Ziolkowski et al., 2017). In particular, the frequency of meiotic recombination in transgenic arabidopsis plants of both Columbia-0 and Landsberg erecta ecotypes or their hybrid Columbia-0 × Landsberg erecta significantly increases and shows a positive correlation with the expression level of the *HEI10* transgene. The population of transgenic plants based on the Columbia-0 ecotype with overexpression of the *HEI10* gene contained more than twice as many crossovers, which was revealed using the MLH1 protein, a marker for class I crossovers. A simultaneous increase in the number of copies of the *HEI10* gene and knockout of the *RECQ4A* and *RECQ4B* genes in arabidopsis lead to a 5-fold increase in meiotic recombination in chromosome arms and to a 1.5-fold increase in pericentromeric heterochromatin (Serra et al., 2018). Thus, the combination of overexpression of the *HEI10* gene with suppression of the expression of the *RECQ4A* and *RECQ4B* genes for the first time made it possible to simultaneously increase the number of class I and II crossovers.

Conclusion

Over the past two decades, numerous studies have been performed that allowed to reveal key elements of the control of meiotic crossing over, which can be used to increase the frequency of crossover exchanges and redistribute their positions along the chromosomes. The experiments on overexpression of the *HEI10* crossover enhancer gene and inactivation of the *FANCM*, *RECQ4*, and *FIGL1* crossover repressor genes in arabidopsis plants turned out to be the most promising. Combining these experimental approaches has significantly increased the frequency and distribution of class I and II crossovers. The results obtained in arabidopsis opened up the possibility of manipulating the process of meiotic recombination in agricultural plant species. However, the results obtained on the model object are not always reproducible on agricultural crops. Obviously, additional efforts are needed to reveal the features of the functioning of orthologues of these genes in various plant genomes.

References

- Acharya S., Wilson T., Gradia S., Kane M.F., Guerrette S., Marsichky G.T., Kolodner R., Fishel R. hMSH2 forms specific mismatch-binding complexes with hMSH3 and hMSH6. *Proc. Natl. Acad. Sci. USA.* 1996;93(24):13629-13634. DOI 10.1073/pnas.93.24.13629.
- Anderson L.K., Doyle G.G., Brigham B., Carter J., Hooker K.D., Lai A., Rice M., Stack S.M. High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics.* 2003;165(2): 849-865. DOI 10.1093/genetics/165.2.849.
- Anderson L.K., Lohmiller L.D., Tang X., Hammond D.B., Javernick L., Shearer L., Basu-Roy S., Martin O.C., Falque M. Combined fluorescent and electron microscopic imaging unveils the specific proper-

- ties of two classes of meiotic crossovers. *Proc. Natl. Acad. Sci. USA*. 2014;111(37):13415-13420. DOI 10.1073/pnas.1406846111.
- Baudat F., Buard J., Grey C., Fledel-Alon A., Ober C., Przeworski M., Coop G., de Massy B. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science*. 2010; 327(5967):836-840. DOI 10.1126/science.1183439.
- Baumann P., West S.C. Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem. Sci.* 1998;23(7):247-251. DOI 10.1016/S0968-0004(98)01232-8.
- Berchowitz L.E., Francis K.E., Bey A.L., Copenhaver G.P. The role of *AtMUS81* in interference-insensitive crossovers in *A. thaliana*. *PLoS Genet.* 2007;3(8):e132. DOI 10.1371/journal.pgen.0030132.
- Bishop D.K., Zickler D. Early decision: meiotic crossover interference prior to stable strand exchange and synapsis. *Cell*. 2004;117(1): 9-15. DOI 10.1016/S0092-8674(04)00297-1.
- Blary A., Gonzalo A., Eber F., Bérard A., Bergès H., Bessoltane N., Charif D., Charpentier C., Cromer L., Fourment J., Genevriez C., Le Paslier M.-C., Lodé M., Lucas M.-O., Nesi N., Lloyd A., Chèvre A.-M., Jenczewski E. FANCM limits meiotic crossovers in *Brassica* crops. *Front. Plant Sci.* 2018;9:368. DOI 10.3389/fpls.2018.00368.
- Blary A., Jenczewski E. Manipulation of crossover frequency and distribution for plant breeding. *Theor. Appl. Genet.* 2019;132(3):575-592. DOI 10.1007/s00122-018-3240-1.
- Bogdanov Y.F., Grishaeva T.M. Meiotic recombination. The metabolic pathways from DNA double-strand breaks to crossing over and chiasmata. *Russ. J. Genet.* 2020;56(2):159-176. DOI 10.1134/S1022795420020039.
- Brown M.S., Bishop D.K. DNA strand exchange and RecA homologs in meiosis. *Cold Spring Harb. Perspect. Biol.* 2014;7(1):a016659. DOI 10.1101/cshperspect.a016659.
- Chambers S.R., Hunter N., Louis E.J., Borts R.H. The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. *Mol. Cell. Biol.* 1996; 16(11):6110-6120. DOI 10.1128/MCB.16.11.6110.
- Choi K., Zhao X., Kelly K.A., Venn O., Higgins J.D., Yelina N.E., Hardcastle T.J., Ziolkowski P.A., Copenhaver G.P., Franklin F.C.H., McVean G., Henderson I.R. *Arabidopsis* meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* 2013;45(11):1327-1336. DOI 10.1038/ng.2766.
- Cloud V., Chan Y.-L., Grubb J., Budke B., Bishop D.K. Rad51 is an accessory factor for Dmcl-mediated joint molecule formation during meiosis. *Science*. 2012;337(6099):1222-1225. DOI 10.1126/science.1219379.
- Cole F., Kauppi L., Lange J., Roig I., Wang R., Keeney S., Jasin M. Homeostatic control of recombination is implemented progressively in mouse meiosis. *Nat. Cell Biol.* 2012a;14(4):424-430. DOI 10.1038/ncb2451.
- Cole F., Keeney S., Jasin M. Preaching about the converted: how meiotic gene conversion influences genomic diversity. *Ann. N.Y. Acad. Sci.* 2012b;1267(1):95-102. DOI 10.1111/j.1749-6632.2012.06595.x.
- Couteau F., Belzile F., Horlow C., Grandjean O., Vezon D., Doutriaux M.-P. Random chromosome segregation without meiotic arrest in both male and female meocytes of a *dmc1* mutant of *Arabidopsis*. *Plant Cell*. 1999;11(9):1623-1634. DOI 10.1105/tpc.11.9.1623.
- Crismani W., Girard C., Froger N., Pradillo M., Santos J.L., Chelysheva L., Copenhaver G.P., Horlow C., Mercier R. FANCM limits meiotic crossovers. *Science*. 2012;336(6088):1588-1590. DOI 10.1126/science.1220381.
- Crismani W., Girard C., Mercier R. Tinkering with meiosis. *J. Exp. Bot.* 2013;64(1):55-65. DOI 10.1093/jxb/ers314.
- Culligan K.M., Hays J.B. *Arabidopsis* MutS homologs – atMSH2, atMSH3, atMSH6, and a novel atMSH7 – form three distinct protein heterodimers with different specificities for mismatched DNA. *Plant Cell*. 2000;12(6):991-1002. DOI 10.1105/tpc.12.6.991.
- Da Ines O., Degroote F., Goubely C., Amiard S., Gallego M.E., White C.I. Meiotic recombination in *Arabidopsis* is catalysed by DMC1, with RAD51 playing a supporting role. *PLoS Genet.* 2013; 9(9):e1003787. DOI 10.1371/journal.pgen.1003787.
- Darrier B., Rimbart H., Balfourier F., Pingault L., Josselin A.-A., Servin B., Navarro J., Choulet F., Paux E., Sourdille P. High-resolution mapping of crossover events in the hexaploid wheat genome suggests a universal recombination mechanism. *Genetics*. 2017;206(3):1373-1388. DOI 10.1534/genetics.116.196014.
- De Muyt A., Pereira L., Vezon D., Chelysheva L., Gendrot G., Chambon A., Lainé-Choinard S., Pelletier G., Mercier R., Nogué F., Grelon M. A high throughput genetic screen identifies new early meiotic recombination functions in *Arabidopsis thaliana*. *PLoS Genet.* 2009;5(9):e1000654. DOI 10.1371/journal.pgen.1000654.
- Demirci S., van Dijk A.D.J., Sanchez Perez G., Afitos S.A., de Ridder D., Peters S.A. Distribution, position and genomic characteristics of crossovers in tomato recombinant inbred lines derived from an interspecific cross between *Solanum lycopersicum* and *Solanum pimpinellifolium*. *Plant J.* 2017;89(3):554-564. DOI 10.1111/tj.13406.
- Dion É., Li L., Jean M., Belzile F. An *Arabidopsis MLH1* mutant exhibits reproductive defects and reveals a dual role for this gene in mitotic recombination. *Plant J.* 2007;51(3):431-440. DOI 10.1111/j.1365-3113X.2007.03145.x.
- Emmanuel E., Yehuda E., Melamed-Bessudo C., Avivi-Ragolsky N., Levy A.A. The role of *AtMSH2* in homologous recombination in *Arabidopsis thaliana*. *EMBO Rep.* 2006;7(1):100-105. DOI 10.1038/sj.embor.7400577.
- Fasching C.L., Cejka P., Kowalczykowski S.C., Heyer W.-D. Top3-Rmi1 dissolve Rad51-mediated D loops by a topoisomerase-based mechanism. *Mol. Cell*. 2015;57(4):595-606. DOI 10.1016/j.molcel.2015.01.022.
- Fernandes J.B., Duhamel M., Seguéla-Arnaud M., Froger N., Girard C., Choinard S., Solier V., Winne N.D., Jaeger G.D., Gevaert K., Andrey P., Grelon M., Guerois R., Kumar R., Mercier R. FIGL1 and its novel partner FLIP form a conserved complex that regulates homologous recombination. *PLoS Genet.* 2018a;14(4):e1007317. DOI 10.1371/journal.pgen.1007317.
- Fernandes J.B., Séguéla-Arnaud M., Larchevêque C., Lloyd A.H., Mercier R. Unleashing meiotic crossovers in hybrid plants. *Proc. Natl. Acad. Sci. USA*. 2018b;115(10):2431-2436. DOI 10.1073/pnas.1713078114.
- Gari K., Décaillot C., Stasiak A.Z., Stasiak A., Constantinou A. The fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. *Mol. Cell*. 2008;29(1):141-148. DOI 10.1016/j.molcel.2007.11.032.
- Genschel J., Littman S.J., Drummond J.T., Modrich P. Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα. *J. Biol. Chem.* 1998;273(31): 19895-19901. DOI 10.1074/jbc.273.31.19895.
- Girard C., Chelysheva L., Choinard S., Froger N., Macaisne N., Lemhemdi A., Mazel J., Crismani W., Mercier R. AAA-ATPase FIDGETIN-LIKE 1 and helicase FANCM antagonize meiotic crossovers by distinct mechanisms. *PLoS Genet.* 2015;11(7):e1005369. DOI 10.1371/journal.pgen.1005369.
- Girard C., Crismani W., Froger N., Mazel J., Lemhemdi A., Horlow C., Mercier R. FANCM-associated proteins MHF1 and MHF2, but not the other Fanconi anemia factors, limit meiotic crossovers. *Nucleic Acids Res.* 2014;42(14):9087-9095. DOI 10.1093/nar/gku614.
- Gray S., Cohen P.E. Control of meiotic crossovers: from double-strand break formation to designation. *Annu. Rev. Genet.* 2016;50(1):175-210. DOI 10.1146/annurev-genet-120215-035111.
- Grelon M., Vezon D., Gendrot G., Pelletier G. *AtSPO11-1* is necessary for efficient meiotic recombination in plants. *EMBO J.* 2001; 20(3):589-600. DOI 10.1093/emboj/20.3.589.

- Harrison C.J., Alvey E., Henderson I.R. Meiosis in flowering plants and other green organisms. *J. Exp. Bot.* 2010;61(11):2863-2875. DOI 10.1093/jxb/erq191.
- Hartung F., Angelis K.J., Meister A., Schubert I., Melzer M., Puchta H. An archaeobacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants. *Curr. Biol.* 2002;12(20):1787-1791. DOI 10.1016/S0960-9822(02)01218-6.
- Hartung F., Puchta H. Molecular characterization of homologues of both subunits A (SPO11) and B of the archaeobacterial topoisomerase 6 in plants. *Gene.* 2001;271(1):81-86. DOI 10.1016/S0378-1119(01)00496-6.
- Higgins J.D., Buckling E.F., Franklin F.C.H., Jones G.H. Expression and functional analysis of *AtMUS81* in Arabidopsis meiosis reveals a role in the second pathway of crossing-over. *Plant J.* 2008;54(1):152-162. DOI 10.1111/j.1365-313X.2008.03403.x.
- Hoffman P.D., Leonard J.M., Lindberg G.E., Bollmann S.R., Hays J.B. Rapid accumulation of mutations during seed-to-seed propagation of mismatch-repair-defective *Arabidopsis*. *Genes Dev.* 2004;18(21):2676-2685. DOI 10.1101/gad.1217204.
- Hunter N. Meiotic recombination: the essence of heredity. *Cold Spring Harb. Perspect. Biol.* 2015;7(12):a016618. DOI 10.1101/cshperspect.a016618.
- Hunter N., Chambers S.R., Louis E.J., Borts R.H. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* 1996;15(7):1726-1733. DOI 10.1002/j.1460-2075.1996.tb00518.x.
- Hunter N., Kleckner N. The single-end invasion: an asymmetric intermediate at the double-strand break to double-holliday junction transition of meiotic recombination. *Cell.* 2001;106(1):59-70. DOI 10.1016/S0092-8674(01)00430-5.
- Jones G.H., Franklin F.C.H. Meiotic crossing-over: obligation and interference. *Cell.* 2006;126(2):246-248. DOI 10.1016/j.cell.2006.07.010.
- Kauppi L., Barchi M., Lange J., Baudat F., Jasin M., Keeney S. Numerical constraints and feedback control of double-strand breaks in mouse meiosis. *Genes Dev.* 2013;27(8):873-886. DOI 10.1101/gad.213652.113.
- Keeney S., Giroux C.N., Kleckner N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell.* 1997;88(3):375-384. DOI 10.1016/S0092-8674(00)81876-0.
- Kleckner N. Meiosis: how could it work? *Proc. Natl. Acad. Sci. USA.* 1996;93(16):8167-8174. DOI 10.1073/pnas.93.16.8167.
- Komakhin R.A., Komakhina V.V., Milyukova N.A., Goldenkova-Pavlova I.V., Fadina O.A., Zhuchenko A.A. Transgenic tomato plants expressing *recA* and *NLS-recA-licBM3* genes as a model for studying meiotic recombination. *Russ. J. Genet.* 2010;46(12):1440-1448. DOI 10.1134/S1022795410120069.
- Komakhin R.A., Komakhina V.V., Milyukova N.A., Zhuchenko A.A. Analysis of the meiotic recombination frequency in transgenic tomato hybrids expressing *recA* and *NLS-recA-licBM3* genes. *Russ. J. Genet.* 2012;48(1):23-31. DOI 10.1134/S1022795411110093.
- Komakhin R.A., Milyukova N.A., Strelnikova S.R., Krinitsina A.A., Komakhina V.V., Zhuchenko A.A. Inheritance of marker genes among progeny of interspecific tomato hybrids expressing the *recA* *Escherichia coli* gene. *Russ. J. Genet.* 2019;55(4):433-443. DOI 10.1134/S1022795419040069.
- Komakhina V.V., Krinitsina A.A., Milyukova N.A., Komakhin R.A. Expression of recombinant *SPO11* genes locally alters crossing over in tomato. *Russ. J. Genet.* 2020;56(9):1079-1089. DOI 10.1134/S1022795420090124.
- Kurzbaier M.-T., Uanschou C., Chen D., Schlögelhofer P. The recombinases DMC1 and RAD51 are functionally and spatially separated during meiosis in *Arabidopsis*. *Plant Cell.* 2012;24(5):2058-2070. DOI 10.1105/tpc.112.098459.
- Lambing C., Franklin F.C.H., Wang C.-J.R. Understanding and manipulating meiotic recombination in plants. *Plant Physiol.* 2017;173(3):1530-1542. DOI 10.1104/pp.16.01530.
- Lanzov V.A. *RecA* homologous DNA transferase: functional activities and a search for homology by recombining DNA molecules. *Mol. Biol.* 2007;41(3):417-426. DOI 10.1134/S0026893307030077.
- Lao J.P., Cloud V., Huang C.-C., Grubb J., Thacker D., Lee C.-Y., Dresser M.E., Hunter N., Bishop D.K. Meiotic crossover control by concerted action of Rad51-Dmc1 in homolog template bias and robust homeostatic regulation. *PLoS Genet.* 2013;9(12):e1003978. DOI 10.1371/journal.pgen.1003978.
- Leonard J.M., Bollmann S.R., Hays J.B. Reduction of stability of *Arabidopsis* genomic and transgenic DNA-repeat sequences (microsatellites) by inactivation of atMSH2 mismatch-repair function. *Plant Physiol.* 2003;133(1):328-338. DOI 10.1104/pp.103.023952.
- Lhuissier F.G.P., Offenberg H.H., Wittich P.E., Vischer N.O.E., Heyting C. The Mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. *Plant Cell.* 2007;19(3):862-876. DOI 10.1105/tpc.106.049106.
- Li X., Yu M., Bolaños-Villegas P., Zhang J., Ni D., Ma H., Wang Y. Fanconi anemia ortholog FANCM regulates meiotic crossover distribution in plants. *Plant Physiol.* 2021;186(1):344-360. DOI 10.1093/plphys/kiab061.
- Liu S., Yeh C.-T., Ji T., Ying K., Wu H., Tang H.M., Fu Y., Nettleton D., Schnable P.S. *Mu* transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genet.* 2009;5(11):e1000733. DOI 10.1371/journal.pgen.1000733.
- Lloyd A.H., Milligan A.S., Langridge P., Able J.A. *TaMSH7*: A cereal mismatch repair gene that affects fertility in transgenic barley (*Hordeum vulgare* L.). *BMC Plant Biol.* 2007;7(1):67. DOI 10.1186/1471-2229-7-67.
- Macaisne N., Vignard J., Mercier R. SHOC1 and PTD form an XPF-ERCC1-like complex that is required for formation of class I crossovers. *J. Cell Sci.* 2011;124(16):2687-2691. DOI 10.1242/jcs.088229.
- Mancera E., Bourgon R., Brozzi A., Huber W., Steinmetz L.M. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature.* 2008;454(7203):479-485. DOI 10.1038/nature07135.
- Marti T.M., Kunz C., Fleck O. DNA mismatch repair and mutation avoidance pathways. *J. Cell. Physiol.* 2002;191(1):28-41. DOI 10.1002/jcp.10077.
- Martini E., Diaz R.L., Hunter N., Keeney S. Crossover homeostasis in yeast meiosis. *Cell.* 2006;126:285-295. DOI 10.1016/j.cell.2006.05.044.
- Mercier R., Mézard C., Jenczewski E., Macaisne N., Grelon M. The molecular biology of meiosis in plants. *Annu. Rev. Plant Biol.* 2015;66(1):297-327. DOI 10.1146/annurev-arplant-050213-035923.
- Mézard C., Tagliaro Jahns M., Grelon M. Where to cross? New insights into the location of meiotic crossovers. *Trends Genet.* 2015;31(7):393-401. DOI 10.1016/j.tig.2015.03.008.
- Mézard C., Vignard J., Drouaud J., Mercier R. The road to crossovers: plants have their say. *Trends Genet.* 2007;23(2):91-99. DOI 10.1016/j.tig.2006.12.007.
- Mieulet D., Aubert G., Bres C., Klein A., Droc G., Vieille E., Rond-Coissieux C., Sanchez M., Dalmais M., Mauxion J.-P., Rothan C., Guiderdoni E., Mercier R. Unleashing meiotic crossovers in crops. *Nat. Plants.* 2018;4(12):1010-1016. DOI 10.1038/s41477-018-0311-x.
- Modrich P. Mechanisms and biological effects of Mismatch repair. *Annu. Rev. Genet.* 1991;25:229-253. DOI 10.1146/annurev.ge.25.120191.001305.
- Nichols M.D., DeAngelis K., Keck J.L., Berger J.M. Structure and function of an archaeal topoisomerase VI subunit with homology

- to the meiotic recombination factor Spo11. *EMBO J.* 1999;18(21): 6177-6188. DOI 10.1093/emboj/18.21.6177.
- Osman K., Higgins J.D., Sanchez-Moran E., Armstrong S.J., Franklin F.C.H. Pathways to meiotic recombination in *Arabidopsis thaliana*. *New Phytol.* 2011;190(3):523-544. DOI 10.1111/j.1469-8137.2011.03665.x.
- Peciña A., Smith K.N., Mézard C., Murakami H., Ohta K., Nicolas A. Targeted stimulation of meiotic recombination. *Cell.* 2002;111(2): 173-184. DOI 10.1016/S0092-8674(02)01002-4.
- Rakosy-Tican E., Lőrincz-Besenyey E., Molnár I., Thieme R., Hartung F., Sprink T., Antonova O., Famelauer I., Angenon G., Aurori A. New phenotypes of potato co-induced by mismatch repair deficiency and somatic hybridization. *Front. Plant Sci.* 2019;10:3. DOI 10.3389/fpls.2019.00003.
- Ranjha L., Anand R., Cejka P. The *Saccharomyces cerevisiae* Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. *J. Biol. Chem.* 2014;289(9):5674-5686. DOI 10.1074/jbc.M113.533810.
- Reiss B., Klemm M., Kosak H., Schell J. *RecA* protein stimulates homologous recombination in plants. *Proc. Natl. Acad. Sci. USA.* 1996;93(7):3094-3098. DOI 10.1073/pnas.93.7.3094.
- Reiss B., Schubert I., Köpchen K., Wendeler E., Schell J., Puchta H. *RecA* stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by *Agrobacterium*. *Proc. Natl. Acad. Sci. USA.* 2000;97(7): 3358-3363. DOI 10.1073/pnas.97.7.3358.
- Robert T., Nore A., Brun C., Maffre C., Crimi B., Guichard V., Bourbon H.-M., de Massy B. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science.* 2016;351(6276):943-949. DOI 10.1126/science.aad5309.
- Rodgers-Melnick E., Bradbury P.J., Elshire R.J., Glaubitz J.C., Acharya C.B., Mitchell S.E., Li C., Li Y., Buckler E.S. Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl. Acad. Sci. USA.* 2015;112(12):3823-3828. DOI 10.1073/pnas.1413864112.
- Roeder G.S. Meiotic chromosomes: it takes two to tango. *Genes Dev.* 1997;11(20):2600-2621. DOI 10.1101/gad.11.20.2600.
- Sachadyn P. Conservation and diversity of MutS proteins. *Mutat. Res. Mol. Mech. Mutagen.* 2010;694(1):20-30. DOI 10.1016/j.mrfmmm.2010.08.009.
- Saintenac C., Falque M., Martin O.C., Paux E., Feuillet C., Sourdille P. Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). *Genetics.* 2009;181(2):393-403. DOI 10.1534/genetics.108.097469.
- Saintenac C., Faure S., Remay A., Choulet F., Ravel C., Paux E., Balfourier F., Feuillet C., Sourdille P. Variation in crossover rates across a 3-Mb contig of bread wheat (*Triticum aestivum*) reveals the presence of a meiotic recombination hotspot. *Chromosoma.* 2011; 120(2):185-198. DOI 10.1007/s00412-010-0302-9.
- Sarma S., Pandey A.K., Sharma K., Ravi M., Sreelakshmi Y., Sharma R. MutS-Homolog2 silencing generates tetraploid meiocytes in tomato (*Solanum lycopersicum*). *Plant Direct.* 2018;2(1):e00017. DOI 10.1002/pld3.17.
- Sarno R., Vicq Y., Uematsu N., Luka M., Lapierre C., Carroll D., Bastianelli G., Serero A., Nicolas A. Programming sites of meiotic crossovers using Spo11 fusion proteins. *Nucleic Acids Res.* 2017; 45(19):e164. DOI 10.1093/nar/gkx739.
- Séguéla-Arnaud M., Choinard S., Larchevêque C., Girard C., Froger N., Crismani W., Mercier R. RMI1 and TOP3 α limit meiotic CO formation through their C-terminal domains. *Nucleic Acids Res.* 2017; 45(4):1860-1871. DOI 10.1093/nar/gkw1210.
- Séguéla-Arnaud M., Crismani W., Larchevêque C., Mazel J., Froger N., Choinard S., Lemhemdi A., Macaisne N., Leene J.V., Gevaert K., Jaeger G.D., Chelysheva L., Mercier R. Multiple mechanisms limit meiotic crossovers: TOP3 α and two BLM homologs antagonize crossovers in parallel to FANCM. *Proc. Natl. Acad. Sci. USA.* 2015; 112(15):4713-4718. DOI 10.1073/pnas.1423107112.
- Serra H., Lambing C., Griffin C.H., Topp S.D., Nageswaran D.C., Underwood C.J., Ziolkowski P.A., Séguéla-Arnaud M., Fernandes J.B., Mercier R., Henderson I.R. Massive crossover elevation via combination of *HEI10* and *recq4a recq4b* during *Arabidopsis* meiosis. *Proc. Natl. Acad. Sci. USA.* 2018;115(10):2437-2442. DOI 10.1073/pnas.1713071115.
- Sheridan S.D., Yu X., Roth R., Heuser J.E., Sehorn M.G., Sung P., Egelman E.H., Bishop D.K. A comparative analysis of Dmc1 and Rad51 nucleoprotein filaments. *Nucleic Acids Res.* 2008;36(12):4057-4066. DOI 10.1093/nar/gkn352.
- Shingu Y., Tokai T., Agawa Y., Toyota K., Ahmed S., Kawagishi-Kobayashi M., Komatsu A., Mikawa T., Yamamoto M.-T., Wakasa K., Shibata T., Kusano K. The double-stranded break-forming activity of plant SPO11s and a novel rice SPO11 revealed by a *Drosophila* bioassay. *BMC Mol. Biol.* 2012;13(1):1. DOI 10.1186/1471-2199-13-1.
- Simanovsky S.A., Bogdanov Yu.F. Genetic control of meiosis in plants. *Russ. J. Genet.* 2018;54(4):389-402. DOI 10.1134/S1022795418030122.
- Smagulova F., Gregoretti I.V., Brick K., Khil P., Camerini-Otero R.D., Petukhova G.V. Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature.* 2011;472(7343): 375-378. DOI 10.1038/nature09869.
- Stacey N.J., Kuromori T., Azumi Y., Roberts G., Breuer C., Wada T., Maxwell A., Roberts K., Sugimoto-Shirasu K. Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J.* 2006; 48(2):206-216. DOI 10.1111/j.1365-313X.2006.02867.x.
- Strelnikova S.R., Krinitsina A.A., Komakhin R.A. Effective RNAi-mediated silencing of the *Mismatch repair MSH2* gene induces sterility of tomato plants but not an increase in meiotic recombination. *Genes.* 2021;12(8):1167. DOI 10.3390/genes12081167.
- Taagen E., Bogdanove A.J., Sorrells M.E. Counting on crossovers: controlled recombination for plant breeding. *Trends Plant Sci.* 2020;25(5):455-465. DOI 10.1016/j.tplants.2019.12.017.
- Tam S.M., Hays J.B., Chetelat R.T. Effects of suppressing the DNA mismatch repair system on homeologous recombination in tomato. *Theor. Appl. Genet.* 2011;123(8):1445-1458. DOI 10.1007/s00122-011-1679-4.
- Turner J.M.A. Meiosis 2007 – Where have we got to and where are we going? *Chromosome Res.* 2007;15(5):517-521. DOI 10.1007/s10577-007-1152-z.
- Uanschou C., Ronceret A., Harder M.V., Muylt A.D., Vezon D., Pereira L., Chelysheva L., Kobayashi W., Kurumizaka H., Schlögelhofer P., Grelon M. Sufficient amounts of functional HOP2/MND1 complex promote interhomolog DNA repair but are dispensable for intersister DNA repair during meiosis in *Arabidopsis*. *Plant Cell.* 2013;25(12):4924-4940. DOI 10.1105/tpc.113.118521.
- Villeneuve A.M., Hillers K.J. Whence meiosis? *Cell.* 2001;106(6):647-650. DOI 10.1016/S0092-8674(01)00500-1.
- Vrielynck N., Chambon A., Vezon D., Pereira L., Chelysheva L., Muylt A.D., Mézard C., Mayer C., Grelon M. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science.* 2016; 351(6276):939-943. DOI 10.1126/science.aad5196.
- Wang S., Zickler D., Kleckner N., Zhang L. Meiotic crossover patterns: Obligatory crossover, interference and homeostasis in a single process. *Cell Cycle.* 2015;14(3):305-314. DOI 10.4161/15384101.2014.991185.
- Wang Y., Copenhaver G.P. Meiotic recombination: Mixing it up in plants. *Annu. Rev. Plant Biol.* 2018;69(1):577-609. DOI 10.1146/annurev-arplant-042817-040431.
- Wijnker E., de Jong H. Managing meiotic recombination in plant breeding. *Trends Plant Sci.* 2008;13(12):640-646. DOI 10.1016/j.tplants.2008.09.004.

- Wijnker E., James G.V., Ding J., Becker F., Klasen J.R., Rawat V., Rowan B.A., de Jong D.F., de Snoo C.B., Zapata L. The genomic landscape of meiotic crossovers and gene conversions in *Arabidopsis thaliana*. *eLife*. 2013;2:e01426. DOI 10.7554/eLife.01426.
- Wu H., Gao J., Sharif W.D., Davidson M.K., Wahls W.P. Purification, folding, and characterization of Rec12 (Spo11) meiotic recombinase of fission yeast. *Protein Expr. Purif.* 2004;38(1):136-144. DOI 10.1016/j.pep.2004.07.012.
- Yelina N.E., Lambing C., Hardcastle T.J., Zhao X., Santos B., Henderson I.R. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in *Arabidopsis*. *Genes Dev.* 2015;29(20):2183-2202. DOI 10.1101/gad.270876.115.
- Youds J.L., Boulton S.J. The choice in meiosis – defining the factors that influence crossover or non-crossover formation. *J. Cell Sci.* 2011;124(4):501-513. DOI 10.1242/jcs.074427.
- Yuan J., Chen J. FIGNL1-containing protein complex is required for efficient homologous recombination repair. *Proc. Natl. Acad. Sci. USA.* 2013;110(26):10640-10645. DOI 10.1073/pnas.1220662110.
- Zhang P., Zhang Y., Sun L., Sinumporn S., Yang Z., Sun B., Xuan D., Li Z., Yu P., Wu W., Wang K., Cao L., Cheng S. The rice AAA-ATPase OsFIGNL1 is essential for male meiosis. *Front. Plant Sci.* 2017;8:1639. DOI 10.3389/fpls.2017.01639.
- Zhuchenko A.A., Korol A.B. Recombination in Evolution and Breeding. Moscow: Nauka Publ., 1985. (in Russian)
- Ziolkowski P.A., Underwood C.J., Lambing C., Martinez-Garcia M., Lawrence E.J., Ziolkowska L., Griffin C., Choi K., Franklin F.C.H., Martienssen R.A., Henderson I.R. Natural variation and dosage of the HEI10 meiotic E3 ligase control *Arabidopsis* crossover recombination. *Genes Dev.* 2017;31(3):306-317. DOI 10.1101/gad.295501.116.

ORCID ID

S.R. Strelnikova orcid.org/0000-0003-2641-7069
R.A. Komakhin orcid.org/0000-0001-5963-8111

Acknowledgements. The work was carried out at the expense of state task No. 0431-2022-0004.

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

Received May 19, 2022. Revised September 11, 2022. Accepted September 26, 2022.

Metabolomic approach to investigate *Dactylis glomerata* L. from the VIR collection

N.Yu. Malysheva¹, T.V. Shelenga¹, A.E. Solovyeva¹, T.B. Nagiev², N.V. Kovaleva², L.L. Malyshev¹ 

¹ Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia

² Leningrad Research Agriculture Institute Branch of Russian Potato Research Centre, Leningrad region, Russia

 l.malyshev@vir.nw.ru

Abstract. The perennial grass cocksfoot (*Dactylis glomerata* L.) is a valuable early highly nutritious crop used as green fodder in agricultural production. The species is widespread across the Eurasian continent; it is characterized by plasticity and high ecological and geographical variability. The article considers the metabolic profiles of 15 accessions of the cocksfoot from the collection of the N.I. Vavilov Institute of Plant Genetic Resources (VIR). The material is represented by varieties and wild forms of various origin: the European part of the Russian Federation, Norway and Finland. The study was carried out using gas-liquid chromatography coupled with mass spectrometry. The study and comparison of groups of metabolites of cocksfoot accessions of various ecological and geographical origin was carried out. Statistical processing included the calculation of the main parameters of variability, factor analysis of the correlation system (*Q*- and *R*-technique), cluster analysis by Ward's method and discriminant analysis. The variability of the quantitative and qualitative composition of the substances identified was revealed. Based on statistical processing of the results obtained, five groups of cocksfoot accessions were identified, differing in the profile of metabolites. One of the groups with a similar composition of metabolites consisted of accessions from one ecological and geographical region; another, of accessions of different origin. Significant differences were noted in the metabolomic profiles of a late-maturing wild cocksfoot accession from the Republic of Karelia at the booting stage from early- and mid-maturing accessions at the heading stage; it contained the largest number of free amino acids and the smallest number of identified primary and secondary metabolites. Wild-growing accession k-44020 from Norway surpassed other wild-growing accessions in the content of free amino acids, sugars and phosphates at the heading stage. Wild-growing accessions differed from breeding varieties with a high content of proline and threonine, indicators of high resistance to lack of moisture and high air temperature.

Key words: *Dactylis glomerata*; genetic resources; metabolomic profiling; character polymorphism.

For citation: Malysheva N.Yu., Shelenga T.V., Solovyeva A.E., Nagiev T.B., Kovaleva N.V., Malyshev L.L. Metabolomic approach to investigate *Dactylis glomerata* L. from the VIR collection. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2023;27(2):111-118. DOI 10.18699/VJGB-23-16

Метаболомный подход в изучении *Dactylis glomerata* L. из коллекции ВИР

Н.Ю. Малышева¹, Т.В. Шеленга¹, А.Е. Соловьева¹, Т.Б. Нагиев², Н.В. Ковалева², Л.Л. Малышев¹ 

¹ Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия

² Ленинградский научно-исследовательский институт сельского хозяйства «Белогорка» –

филиал Федерального исследовательского центра картофеля имени А.Г. Лорха, Ленинградская область, Россия

 l.malyshev@vir.nw.ru

Аннотация. Многолетний злак ежа сборная (*Dactylis glomerata* L.) – ценная ранняя высокопитательная культура, используемая в качестве зеленого корма в сельскохозяйственном производстве. Вид широко распространен на территории евразийского континента, характеризуется пластичностью и высокой эколого-географической изменчивостью. В статье рассмотрены метаболитные профили 15 образцов ежи сборной из коллекции Всероссийского института генетических ресурсов растений им. Н.И. Вавилова (ВИР). Материал представлен сортами и дикорастущими формами различного происхождения: европейская часть РФ, Норвегия и Финляндия. Исследования проводили с помощью газо-жидкостной хроматографии, сопряженной с масс-спектрометрией. Выполнено изучение и сравнение групп метаболитов образцов ежи сборной различного эколого-географического происхождения. Статистическая обработка включала вычисление основных параметров изменчивости, факторный анализ системы корреляций (*Q*- и *R*-техника), кластерный анализ по методу Варда и дискриминантный анализ. Выявлена изменчивость количественного и качественного состава идентифицированных веществ. На основе статистической обработки полученных результатов выделили пять групп образцов ежи, различающихся по профилю метаболитов. В одной группе с похожим составом метаболитов оказались образцы из одного эколого-географического региона, в другой – образцы различного происхождения. Отмечены значительные отличия метаболомных профилей позднеспелого дикорастущего образца ежи из Карелии в фазе выхода в

трубку от ранне- и среднеспелых образцов в фазе колошения: он содержал наибольшее количество свободных аминокислот и наименьшее число выявленных первичных и вторичных метаболитов. Дикорастущий образец к-44020 из Норвегии в фазе колошения превзошел остальные дикорастущие образцы по содержанию свободных аминокислот, сахаров и фосфатов. Дикорастущие образцы отличались от селекционных сортов высоким содержанием пролина и треонина, устойчивостью к недостатку влаги и высокой температуре воздуха.

Ключевые слова: *Dactylis glomerata*; генетические ресурсы; метаболомное профилирование; полиморфизм признаков.

Introduction

Dactylis glomerata L. is widely distributed in Eurasia and North Africa. This culture is the fourth most important forage crop in the world, due to high yield and stress factors resistance (Stewart, Ellison, 2011). It is the earliest hay-type fodder crop in Northern Europe. The world collection of the N.I. Vavilov Institute of Plant Genetic Resources (VIR) presents varieties and wild populations of *D. glomerata* from various ecological and geographical areas. The material is represented by the tetraploid subspecies *D. glomerata* subsp. *glomerata* ($2n = 28$) with a high level of genetic diversity (Last et al., 2013). The main criteria in fodder crops breeding are high productivity, intensity of regrowth, and resistance to abiotic stress factors (Tulinov et al., 2019). Quality characteristics are rarely taken into account (Yakovleva et al., 2015).

Plants are able to synthesize a huge number of compounds having a variety of functions. Investigation of individual characters of their quantitative and qualitative composition determines the economic using of the culture (Maslennikov et al., 2012, 2013). N.I. Vavilov Institute, has experience of using metabolomic profiling in studying plant genetic resources from the VIR collection (Shelenga et al., 2014). The biochemical composition of the cocksfoot has not been studied enough. The recently conducted study of *D. glomerata* growing on the Aeolian Islands (Italy) by M. Mandrone et al. (2022) confirmed the relevance of its evaluation as a promising pasture crop that yields a good harvest of green mass under stressful conditions (drought, low temperatures, low pH soils). In the countries of North America, Europe and Oceania, *D. glomerata* is effectively used to combat soil erosion, desertification, for restoration of green areas after fires and logging. The authors also note the lack of information about metabolomic studies of *D. glomerata*. The study of the diversity of *D. glomerata* genotypes from the collection of VIR reveals accessions with optimal feed properties: high values of organic acids, essential fatty and amino acids, monosaccharides, polyols (inositol and its isomers), phytosterols, low concentrations of anti-nutrients (raffinose). Also it reveals accessions in metabolomic profiles (MP) which were dominated by substances – factors of resistance to abiotic stress (FSS, free amino acids – precursors of phenylpropanoids: phenylalanine, tyrosine, tryptophan; pipercolic acid, oxypoline (a structural compound of extensin, which is part of the matrix of the plant cell wall) (Solovyeva et al., 2019), oligosaccharides, monoacylglycerols, galactinol, mannitol, glycosides) and can be used in programs for breeding new varieties resistant to environmental stresses, as well as varieties with improved feed (Rasmussen et al., 2012; Solovyeva et al., 2020).

The purpose of our research was *D. glomerata* metabolomic profiles evaluation to assess the biochemical variability of varieties and wild populations, degree of similarity, differences, and identify the promising sources for breeding.

Materials and methods

The material for research was 15 cultivar and wild accessions of *D. glomerata* from the VIR collection zoned in different regions of the Russian Federation, Norway and Finland (Table). The green mass of 14 accessions was collected at the heading stage, one late-maturing accession – at the booting stage. Samples preparation, GC-MS analysis, results and processing were carried out according to the protocol in three analytical replications (Loskutov et al., 2020). Statistical data processing was performed using application the software package Statistica 12.0 and included calculation of the main parameters of variation – mean, standard error, minimum and maximum, upper and lower level of the confidence interval of the mean at $p = 0.05$ and coefficient of variation; correlation analysis; cluster analysis by Ward's method and Q - and R -technique of the analysis of principal components and discriminant analysis.

Results and discussion

Cocksfoot green mass chemical composition

In total, 125 components from amino acids, organic acids, phenol-containing compounds, sugars, free fatty acids, polyols, glycosides, lactones, phosphates, sterols, and paraffins groups were identified.

List of accessions of cocksfoot (*Dactylis glomerata* L.)

VIR Catalogue	Variety	Origin	Ripeness
k-36566	Tammisto	Finland	Middle
k-36682	VIK 61	Moscow region	Middle
k-36684	Dvina	Arkhangelsk region	Middle
k-38088	Wild	Pskov region	Early
k-43142	Wild	Yaroslavl region	Middle
k-44020	Wild	Norway	Early
k-44021	Wild	Norway	Middle
k-44349	Wild	Leningrad region	Middle
k-44354	Wild	Republic of Komi	Middle
k-27863	Leningradskaya 853	Leningrad region	Middle
k-35060	Neva	Leningrad region	Middle
k-38648	Petrozavodskaya	Republic of Karelia	Middle
k-45034	Khlynovskaya	Kirov region	Middle
k-48628	Triada	Leningrad region	Middle
i-152589	Wild	Republic of Karelia	Late

Amino acids. The main nitrogenous substances of herbaceous plants are proteins, free amino acids and their amides, nucleic acids, nucleotides, and nitrogenous bases. Free amino acids are an important group of compounds involved in the synthesis of specific tissue proteins and other components necessary for organisms (Shkrobotko et al., 2009), contributing to maintaining the functional stability under stress conditions (Sampieva et al., 2010). Free amino acids, having a wide spectrum of pharmacological action, give other substances an easily digestible and harmless form, while enhancing their effect (Shilova et al., 2008). The green mass of the cocksfoot was found to contain 19 free amino acids, including six essential (valine, leucine, isoleucine, phenylalanine, tyrosine, tryptophan), and the nucleoside adenosine (Suppl. Material¹). Nine of them (valine, alanine, leucine, isoleucine, glycine, threonine, serine, aspartic and glutamic acids and their derivatives – asparagine and glutamine; ornithine) were aliphatic; three (phenylalanine, tyrosine and tryptophan) – aromatic and two (proline and oxyproline) – heterocyclic amino acids. Phenylalanine, tyrosine and tryptophan are precursors of phenylpropanoids. Oxyproline is one of the main compounds of the cell matrix, the extensin protein, indirectly indicating the stress resistance of the accessions. Extensin is a glycoprotein with a high content of oxyproline and oligosaccharide side chains from arabinose. Pípecolic acid and proline also belong to the factors of plant protection from stress (Lotova, 2007; Solovyeva et al., 2019, 2020). A quite high content of pípecolic acid, which is related to non-protein amino acids, was detected. The predominant amino acids in the green mass of cocksfoot are oxyproline and glutamine (23.05 and 13.29 % of the total amino acids, respectively). The content of essential amino acids in the accessions is 19.65 %, where valine predominates (5.56 %). In combination with other BAS (biology active compounds: phenol-containing compounds (PhC), polysaccharides, organic acids (OA), macro- and microelements), it emphasizes the economic value of the green mass of cocksfoot and perspectives in breeding for improvement of feed quality. The total content of free amino acids varied from 91.77 to 346.08 conventional units (CU) (average 207.18). The highest values were determined in three accessions (more than 300 CU): Tammisto k-36566 (Finland), wild k-44020 (Norway) and wild i-152589 (the Republic of Karelia). The lowest were found in wild accession k-44349 (Leningrad region), high content of essential amino acids: k-48628 (59.30; Triada, Leningrad region), anti-stress factors: FSS precursor amino acids: i-152589 (25.14; wild, the Republic of Karelia) and k-44354 (24.10; wild, the Republic of Komi), pípecolic acid: k-27863 (27.58; Leningradskaya 853, Leningrad region), oxyproline: k-44354 (68.79; wild, the Republic of Komi; proline: k-44349 (24.27; wild, Leningrad region).

Organic acids. Fruits and roots are characterized by the predominance of free OA; in grass, buds and leaves it is usually in the form of acidic salts. The most common types OA of aliphatic series are malic, citric, succinic, oxalic, phytic, acetic, tartaric, lactic, gallic and others. The value of OA in the diet is determined by their energy value and active participation in metabolism (Latypova et al., 2014). Up to 60 % of organic acids was malic acid (see Suppl. Material). In second place

were inorganic phosphoric and fumaric acids, the content of which is 248.64 and 102.22 CU, respectively. The content of succinic, threonic, citric, ribonic, lactic, glyceric and ketogluconic acids varied in the range from 17.82 to 75.72 CU. The concentration of gluconic, oxalic, maleic, glucaric, erythronic, pyruvic, dehydroabietinic, azelaic, tartaric and aconitic acids did not exceed 10 CU, citraconic and methylmalonic acids – 0.11 CU. The use of green vegetable mass with a high content of malic, tartaric, citric, lactic and ascorbic acids in animal husbandry and poultry farming as the main feed or feed additive improves the absorption of nutrients, and also has antibacterial effect, which has a positive effect on the weight gain of farm birds and animals (Rasmussen et al., 2012; Solovyeva et al., 2019, 2020; Khan et al., 2022).

Lactone and phosphate forms of OA have been also identified. Lactone forms (erythrono-1,4-lactone, glucono-1,4-lactone, glucono-1,5-lactone, on average 127.46 CU) are biologically active forms of organic acids, capable of binding heavy metals, protecting the cell from damage. The presence of phosphate forms (gluconic acid-6-phosphate, on average 2.71 CU) (see Suppl. Material) reflects the activity of metabolic processes in the plant (Cañete-Rodríguez et al., 2016). Data analysis shows a significant content of OA in the green mass of cocksfoot. On average, it was 1819.48 CU, depending on the variety; it varied from 1074.83 to 2579.87 CU (see Suppl. Material). The lowest values of OA were observed in wild accessions i-152589 (the Republic of Karelia) and k-38088 (Pskov region). The highest – in wild accessions k-44349 (2454.37; Leningrad region), k-44021 and k-44020 (2311.65, 2579.87; Norway), high content of malic: k-44349 (1667.06; wild, Leningrad region), k-27863 (1524.16; Leningradskaya 853, Leningrad region), tartaric: k-43142 (3.51; wild, Yaroslavl region) citric: k-44354 (115.68; wild, the Republic of Komi), k-35060 (101.77; Neva, Leningrad region), lactic acids: k-44349 (42.43; wild, Leningrad region), k-44354 (40.10; wild, the Republic of Komi), k-27863 (40.11; Leningradskaya 853, Leningrad region).

Phenol-containing compounds. PhC is one of the most numerous classes of natural compounds with biological activity. The intensity of their accumulation depends on stress factors, plant age and light conditions (Misin et al., 2010; Maslennikov et al., 2013). The accumulation of PhC is closely related to their function and development phase (Sazhina, Misin, 2011). It is noted that PhC have pronounced antibacterial activity, therefore, accessions of forage crops with a high content of PhC can be used not only to create new stress-resistant varieties, but also as an effective supplement to the daily diet in animal husbandry and poultry (Mahfuz et al., 2021). High values of caffeic acid in plant tissues contribute to protection against the penetration of fungal pathogens into them (Balmer et al., 2013). A total of 19 PhC were found in the green mass of cocksfoot (see Suppl. Material): free phenolcarboxylic acids (benzoic – 1.37, nicotinic – 0.55, 4-hydroxybenzoic – 0.65, protocatechuic – 2.20 and 2,3-dihydroxybenzoic – 0.09; average content – 4.86 CU), quinones (hydroquinone – 1.34, resorcinol – 1.19, pyrogallol – 3.04, and plumbagin – 1.01; average – 5.57), acyclic PhC (shikimic – 342.76 and quinic acids – 829.91; average content – 1172.67 CU) and phenylpropanoids (E)-4-hydroxycoric – 22.85, (E)-ferulic – 6.51, caffeic – 24.96, chlorogenic – 21.05, cryptochlorogenic – 4.26,

¹ Supplementary Material is available in the online version of the paper: <https://vavilovj-icg.ru/download/pict-2023-27/appx2.pdf>

neochlorogenic – 10.05 acids, coniferyl alcohol – 17.22 and α -tocopherol – 1.04; average – 106.99 CU). The predominant PhC were quinic and shikimic acids (64.28 and 26.55 % of the total PhC), which indicates the activity of the shikimate pathway of PhC synthesis, and may be associated with environmental stress impacts (Misin et al., 2010). The amount of PhC in cocksfoot on average was 1292.06 CU and varied from 165.27 to 1788.16 CU. In our study, the highest accumulation of PhC was in three accessions: varieties Leningradskaya 253 (k-27863), Tammisto (k-36566) and wild accession from Norway (k-44020) (1788.16, 1771.25, and 1783.67 CU), caffeic acid in Tammisto (45.26 CU; k-36566, Finland).

Carbohydrate composition. In the vegetative organs of forage grasses, the main products of photosynthesis are carbohydrates. Their nutritional value is determined by the amount of easily soluble carbohydrates – monosaccharides and sucrose. In our study, the total amount of sugars in cocksfoot averaged about 15 % of the dry mass, 71 % of which were represented by monosaccharides. Fifteen sugars were identified in the studied cocksfoot accessions: 11 monosaccharides, four oligosaccharides – three disaccharides and one trisaccharide (see Suppl. Material). The sugar content in cocksfoot averaged 4.00 (1.07–7.27) %. The majority of sugars were represented by monosaccharides – 2.85 (0.77–5.48) %, hexoses – 2.84 (0.76–5.47) %, pentoses – 0.014 %. Oligosaccharides were represented by disaccharides – sucrose, maltose and rutinose; trisaccharides – raffinose. The amount of oligosaccharides was 1.15 (0.30–1.78) %, where sucrose was 1.13 %. Metabolically active derivatives of sugars are lactone (glucose-1,4-lactone), phosphate (glucose-1-phosphate) and methyl forms (methylmannoside, methylpentafuranoside, methylglucofuranoside). The amount of sugar derivatives in cocksfoot was 366.2 (28.31–790.37) CU (see Suppl. Material). A number of sugars, such as glucose, sucrose, and raffinose, can accumulate under the influence of stress factors and reflect the activity of plant protection mechanisms from their effects. The nutritional value of feed is associated with a high sugar content, but raffinose has anti-nutritional properties (Solovyeva et al., 2019, 2020). The highest sugar content was determined in the wild accession k-44020 (7.27 %; Norway), monosaccharides (5.48 %), glucose (1755.90 CU), sucrose (1770.5 CU) in k-44349 (wild, Leningrad region), raffinose in k-43142 (27,18; wild, Yaroslavl region); the lowest – in the wild accession i-152589 (1.07 %; the Republic of Karelia), raffinose: k-48628 (2,09; Triada, Leningrad region).

Free fatty acids, acylglycerols and alkanes. The lipid complex of plants is represented by structural and reserve forms. Most of the lipids are found in the tissues of leaves and inflorescences; a lesser part is in the roots and stems of plants. During vegetation, the content of lipids decreases in the green mass, especially in the reproductive phase of development (Novikov, 2012). Eleven free fatty acids (FA) were identified in the green mass of cocksfoot: saturated (pelargonic, undecylic, palmitic, stearic, begenic, lignoceric, cerotinic), unsaturated (oleic, linoleic, linolenic), hydroxyoctadecanoic acids and monoacylglycerols (MAG 1-C16:0; MAG 1-C18:0); and four alkanes (pentacosane, octacosane, nonacosane, hentriacontane) (see Suppl. Material). The high content of FA in the green mass of feed and feed additives has a positive effect on the growth and development of cattle (Shurson et

al., 2015; Leiva, Granados-Chinchilla, 2020). The presence of monoacylglycerols and alkanes in plant tissues is associated with stress resistance (Solovyeva et al., 2019, 2020). The amount of free FA varied from 82.70 to 297.30 (on average 185.69 CU). Lipids of forage grasses have a lot of unsaturated FA 53 % of the total amount of FA, including 39 % essential, so the cocksfoot has a high nutritional value for livestock feeding. The amount of monoacylglycerols ranged from 8.82 to 19.37 CU (on average 14.43), alkanes – from 3.82 to 30.07 CU (10.86) (see Suppl. Material). A high accumulation of FA was observed in wild accessions k-38088 (297.30 CU; Pskov region) and k-44021 (256.78; Norway). Variety Tammisto (k-36566) was distinguished by the content of essential FA (126.56 CU), acylglycerols are in k-44354 (19.37; wild, the Republic of Komi), paraffins – in k-35060 (30.07; Neva, Leningrad region). The lowest FA values were observed in a wild accession from the Republic of Karelia (82.71 CU).

Polyols and phytosterols. Thirteen polyatomic alcohols were found in cocksfoot accessions. The range of variability of identified polyatomic alcohols varied from 119.58 to 269.02 CU (on average 179.87), most of them were related to sugar alcohols: glycerol, erythritol, trietol, xylitol, arabinitol, sorbitol, dulcitol, inositol (presented in three forms – chiro-inositol, methylinositol and myo-inositol) and galactinol, their amount was 164.22 CU. The composition of alcohols also included amino alcohol (ethanolamine) and acyclic diterpene alcohol – phytol. The share of inositols was 25 % of the total amount of alcohols. Phytosterols (campesterol, stigmaterol, β -sitosterol) were detected as well – 34.09 CU (range from 15.92 to 53.70) (see Suppl. Material). Among phytosterols, β -sitosterol prevailed (24.57 CU). In addition, the phosphate forms of glycerol and inositol, and the products of glycerophospholipid metabolism (glycerol-5-phosphate, myo-inositol-2-phosphate, in total – 20.26 CU) were identified. A high content of phytosterols, the quantitative and qualitative composition of polyols characterizes not only the feed value of the green mass (inositol and its derivatives), but also resistance to stress factors (dulcitol, galaktinol) (Noiraud et al., 2001; Solovyeva et al., 2019, 2020). The study revealed accessions with high alcohol content: Petrozavodskaya (269.02) and Tammisto (247.04), wild accession k-44021 (256.75) from Norway, inositol and its derivatives are in k-43142 (114.10; wild, Yaroslavl region) and k-44020 (105.70; wild, Norway), dulcitol is in k-44020 (63.90; wild, Norway), phytosterols and galaktinol – in k-36682 (53.70 and 85.79; VIK 61, Moscow region).

Glycosides. Biologically active secondary metabolites of plants include glycosides, playing an important role in plant protecting and interacting with other organisms. Antirrhinoside and its derivatives are iridoids. They protect the plant from pathogens and pest insects: they repel leaf-eating and non-pollinating insects. Derivative of antirrhinoside, antirride, has antimicrobial and fungicidal activity (Matveeva, Sokornova, 2017). Lupeol (triterpenoid) has an estrogenic, androgynous, antimicrobial, and anticancer effect, and is used as a chemotherapy drug for a number of diseases (Gallo, Sarachine, 2009). Five glycosides were found: methylpentofuranoside, methylmannoside, methylglucofuranoside, antirrhinoside and lupeol (see Suppl. Material). The first three glycosides were discussed earlier in the section “Carbohydrate composition”.

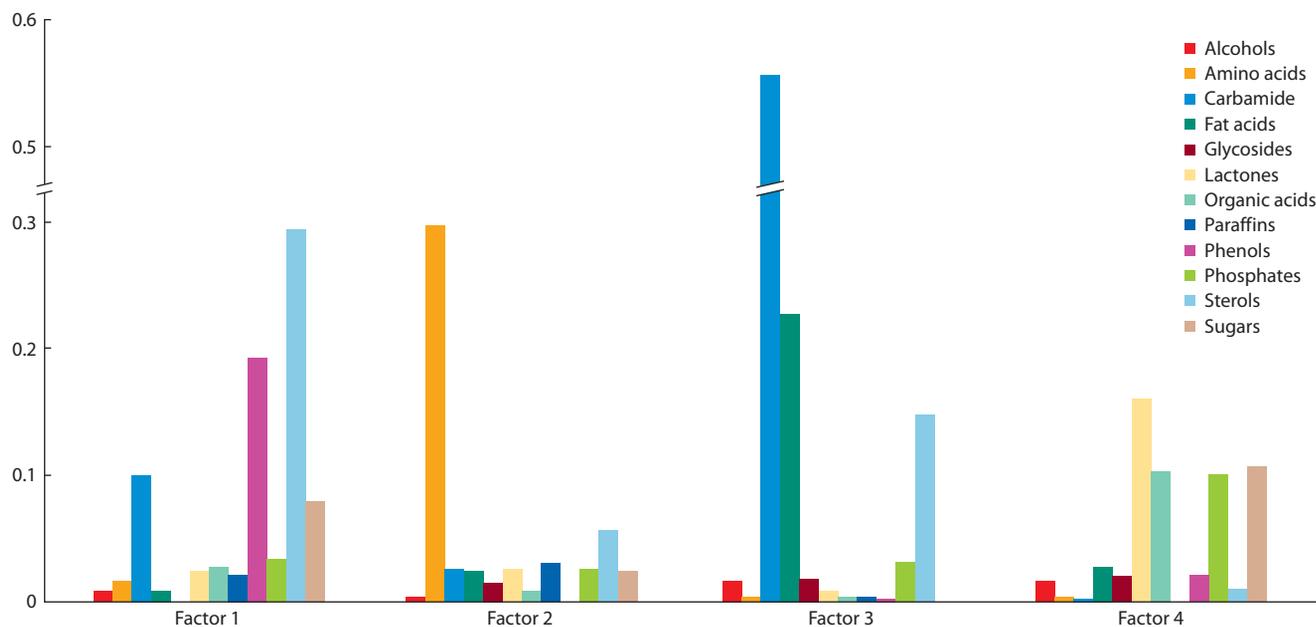


Fig. 1. Average determination of groups of metabolites by the first four principal components of variation.

Antirrhinoside and lupeol were not found in all the studied accessions. The maximum amount of antirrhinoside was found in variety Leningradskaya 853 (k-27863; 63.48 CU), lupeol – in wild accession k-38088 from Pskov region (14.69 CU).

Variability of metabolome in the studied accessions of cocksfoot

This study revealed significant variability in the metabolomic profiles of *D. glomerata* L. accessions. In the course of factor analysis of the correlation matrix, 13 factors were identified, covering a total of 99.3 % of variability. First four factors provide 70.6 % variability, the other nine, only 28.7 %. Factor 1 (27.4 % variance) correlates with the content of 48 out of 126 compounds: 17 with an average ($0.49 > D > 0.25$) and 31 with a high degree of determination ($D \geq 0.50$), where coefficient of determination $D = r^2$, and r is the loading of character on the axis. According to this factor 12 PhC, 4 phosphates, 2 lactones, 2 sterols vary. Factor 2 (15.4 %) determines the variability of 26 compounds: 13 with an average and 13 with a high degree of determination. It is associated with the variability of 13 amino acids. By factor 3 (12.7 %), the content of 19 compounds varies (5 with an average and 14 with a high degree of determination). By this factor varies the content of 6 fatty acids and urea. Factor 4 (15.3 %) is associated with the variability of the content of 28 compounds: 13 with an average and 15 with a high degree of determination. The largest number of compounds that vary by this factor are OA (9) and sugars (7). The following nine factors are associated with the variation of a limited number of compounds. Factor 5 (4.8 %) is strongly correlated with H-quinone, nonacosan and pentacosane and glycerol. Factor 6 (5.3 %) is associated with variation of the amino acid tyrosine, ribose, altrose, sorbose and galactose sugars, and DHO-benzoic PhC. Factor 7 (3.0 %) is associated with the variation of alcohol trietol, alkane hentriacontane and methylpentofuranoside glycoside. Methylphosphate, dulcitol, methyl-inositol, and sterol β -stigmasterol

vary by factor 8 (3.8 %). Factor 9 (4.1 %) causes variation of citric OA and pelargonic FA. Factor 10 (2.0 %) determines the variation of non-protein pipercolic amino acid, glycoside methylglucofuranoside, myo-inositol-2-phosphate, linoleic FA and PhC ferulic acid. Variation of the octacosane alkane and the me-malonic OA occurs by factor 11 (2.7 %). Factor 12 (1.8 %) determines the variability of the alcohol xylitol. The variation in the content of all metabolites is poorly related to factor 13 (1.1 %). Some compounds vary by two factors.

Thus, in the system of inter-population correlations between metabolites, four large pleiades of traits are distinguished (Fig. 1). The first pleiad is related primarily to the variation in the content of phenols and sterols; the second describes the variation in the content of amino acids, the third – fatty acids and urea; the fourth – lactones, organic acids and saccharides. Another eight factors describe the variability of relatively independent traits that are poorly correlated with traits from the main pleiades.

When using the Q -technique of factor analysis, only two groups of accessions are allocated: wild accession i-152589 from the Republic of Karelia (Factor 2) and all other accessions (Factor 1). The Ward's method was used for the cluster analysis procedure. Based on the results of cluster analysis of metabolic profiles, five groups of accessions characterized by similar metabolomic profiles were identified: wild accession from Pskov region (k-38088); wild accession from the Republic of Karelia (i-152589); wild accessions from Norway (k-44020) and (k-44021) and Leningrad region (k-44349); varieties Dvina, Khlynovskaya, Petrozavodskaya, Triada and wild accessions from Yaroslavl region (k-43142) and the Republic of Komi (k-44354); varieties Leningradskaya 853, Neva, Tammisto and VIK 61 (Fig. 2).

The group affiliation of the studied accessions of cocksfoot has a significant effect on the content of 100 metabolites out of 136 identified, i.e. the features of the MP of each of the groups. According to the results of the classical discriminant

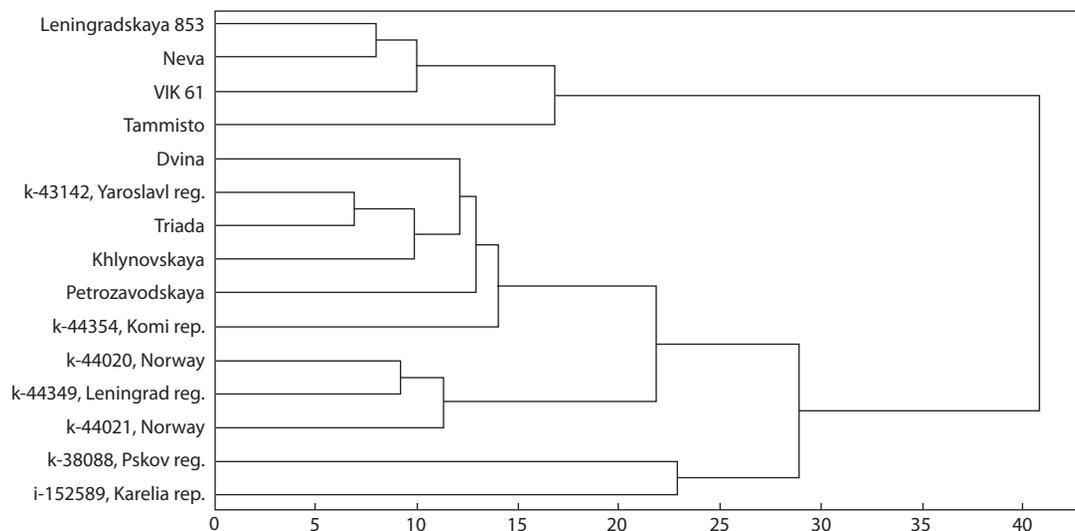


Fig. 2. Classification of cocksfoot accessions by the content of metabolites (cluster analysis, Ward's method).

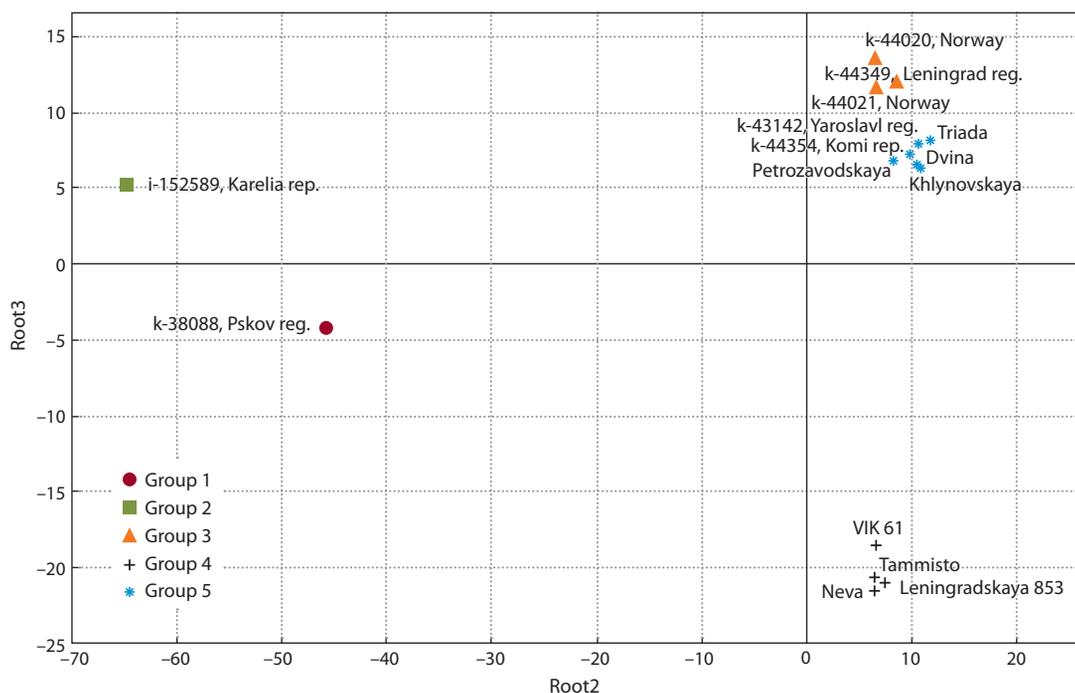


Fig. 3. Differentiation of cocksfoot accessions by the content of metabolites (general discriminant analysis).

analysis, “information value”, the following components reliably distinguishing the groups were identified: lupeol, monosaccharides, arabinose, raffinose, ethanolamine, erythritol, arabinitol. Four variables were identified that ensure the separation of accessions taken in the study: Root1 (arabinitol, ethanolamine), Root2 (arabinose, erythritol), Root3 (sum of monosaccharides), Root4 (sum of monosaccharides, arabinitol). The most obvious separation of accessions was obtained in the Root2 and Root3 axes (Fig. 3).

The first group is characterized by a high content of fatty acids, PhC and polyatomic alcohols, and a low content of glycosides; the second group – by a high content of most amino acids and a low content of sugars and sterols. The third

group, consisting of three wild accessions, is characterized by an increased content of organic acids and sugars. Early-maturing accession from Norway (k-44020) also showed a high content of free amino acids, as well as a late-maturing wild accession from the Republic of Karelia (i-152589). The first one attracts attention as material for creating late-maturing variety, herbal mixes with legumes. Its cutting ripeness occurs during the budding of clover and alfalfa in the North-West of the Russian Federation. These three groups of wild cocksfoot are characterized by a high content of proline and threonine, amino acids that are associated with resistance to stress in response to adverse abiotic factors (Ibragimova et al., 2010; Pandyan et al., 2018).

Wild cocksfoot accessions showed a higher resistance and responded to this stressful situation by accumulating proline. On the other hand, a lesser proline and threonine accumulation and lesser resistance to drought of varieties is a consequence of the process of “domestication”, when the selection in the population was carried out only for economically valuable traits. In this case, the resistance of varieties to stress factors may decrease. The fourth group is characterized by average values of the content of most compounds. Varieties Dvina and Petrozavodskaya created from local wild populations; variety Khlynovskaya – by “free-limited” cross-pollination of local Dedinovskaya from Moscow region. Wild accessions from this group are from the Republic of Komi (k-44354) and Yaroslavl region (k-43142). In this case, it is impossible to explain the grouping of accessions from geographically remote locations into one group. In the fifth group, consisting exclusively of selective varieties, there is a high content of sugars and PhC and a low content of phosphates. Two varieties from the fifth group (Leningradskaya 853 and Neva) were derived from wild populations of cocksfoot from Leningrad region, VIK 61 – by re-pollination of a wild accession from the foothills of the Caucasus with wild accessions from the non-Chernozem zone; the pedigree of variety Tammisto from Finland is unknown. In this group, there were only two varieties originating from one common region. A generalized discriminant analysis model was used to evaluate the degree of differentiation of the selected groups of accessions by metabolomic profiles. The final discriminant functions included nine indicators: the content of lupeol, erythrono-1,4-lactone, glucono-1,4-lactone, methylmalonic acid, pyrogallol, glucosamine, maltose, ethanolamine, and the sum of PhC. The predicted classification based on the constructed functions gives 100 % correct solutions. Thus, the proposed hypothesis about the similarity of metabolomic sections in accessions from a common territory and having similar genotypes is only partially confirmed.

The study of the metabolomic profiles of this culture has been scarcely carried out, as was noted in the work devoted to the study of the features of MP of *D. glomerata*, conducted by M. Mandrone et al. The researchers noted the importance of studying MP to identify the effectiveness of the response of a plant organism to environmental stress, as well as phylogenetic features of culture. The association of high concentrations of valine, asparagine, phenylalanine, fumaric acid and PhC with stressful growth conditions of *D. glomerata*, in particular with drought and increased salt content in the soil of volcanic rocks and the presence of volcanic gases, was noted (Mandrone et al., 2022). The comparison of the data obtained by Italian researchers with our results is rather conditional, since other research methods were used in the work of M. Mandrone et al.: UHPLC–MS, NMR analysis and spectrophotometry. However, they established the prevalence of fumaric acid in the MP of *D. glomerata*, which coincides with our data. The researchers also stressed that increased stress exposure leads to an increase in the accumulation of PhC. S. Rasmussen et al. (2012) noted that there is an increase in quinic and shikimic acid, phytosterols and raffinose in the MP forms of *Lolium perenne*, resistant to drought. In the current study, quinic and shikimic acids were established as dominant in the group of PhC of MP. Accessions of *D. glomerata* with the highest con-

centration of phytosterols, raffinose and quinic and shikimic acid were identified as potentially resistant to stress. In the article of D. Balmer et al. (2013), it is shown that high values of caffeic acid in the tissues of cereal crops protect the plant from fungal pathogens. That was taken into account when we distinguished economically significant *D. glomerata* accessions. The same compounds dominate among organic acids, oligosaccharides, phytosterols and PhC in the MP of cocksfoot and oat seedlings and green mass of peavine previously studied by us (Solovyeva et al., 2019, 2020; Loskutov et al., 2021). In the MP of oat and cocksfoot seedlings in the group of polyols and FA, the main substances are isomers of inositol and linoleic and palmitic acids, peavine and cocksfoot – glycosides: methylglucoside (Loskutov et al., 2021). An iridoid glycoside – antirrhinoside was detected in both the green mass of peavine and cocksfoot (Solovyeva et al., 2019, 2020). There are significant differences in the qualitative composition of the other groups. These differences in MP of different cultures make it possible to assert that MP reflects the specific features of a culture.

Conclusion

As a result of the study, new data on the qualitative and quantitative composition of MP of *D. glomerata* was obtained. With the help of discriminant analysis, the most significant indicators of the MP of the green mass of *D. glomerata* were established. Accessions combining feed value with stability indicators were identified (i-152589, k-27863, 35060, 36566, 43142, 44020, 44349, 44354), as well as those with high indicators of feed value (k-38088, 38648, 44021, 48628) and anti-stress factors (k-27863, 36682, 38088), suitable for breeding highly nutritious varieties resistant to abiotic factors. The study confirms the potential of *D. glomerata* as a promising forage crop. We have confirmed that the optimal plant stage for feeding animals is the stage of heading, when the content of nutrients is high and at the same time the stems of plants are not yet coarsened. But additional research is required to identify changes in metabolites at different stages of cocksfoot’s life cycle.

References

- Balmer D., Flors V., Glauser G., Mauch-Mani B. Metabolomics of cereals under biotic stress: current knowledge and techniques. *Front. Plant Sci.* 2013;4:82. DOI 10.3389/fpls.2013.00082.
- Cañete-Rodríguez A.M., Santos-Duenas I.M., Jimenez-Hornero J.E., Ehrenreich A., Liebl W., Garcia-Garcia I. Gluconic acid: Properties, production methods and applications – an excellent opportunity for agroindustrial by-products and waste biovalorization. *Process Biochem.* 2016;51(12):1891-1903. DOI 10.1016/j.procbio.2016.08.028.
- Gallo M.B.C., Sarachine M.J. Biological activities of lupeol. *Int. J. Res. Pharm. Biomed. Sci.* 2009;3:46-66.
- Ibragimova S.S., Gorelova V.V., Kochetov A.V., Shumnyi V.K. The role of various metabolites in the formation of plant stress resistance. *Vestnik Novosibirskogo Gosudarstvennogo Universiteta. Seriya: Biologiya, Klinicheskaya Meditsina = Bulletin of Novosibirsk State University. Series: Biology, Clinical Medicine.* 2010;8(3):98-103. (in Russian)
- Khan R.U., Naz S., Raziq F., Qudratullah Q., Khan N.A., Laudadio V., Tufarelli V., Ragni M. Prospects of organic acids as safe alternative to antibiotics in broiler chickens diet. *Environ. Sci. Pollut. Res.* 2022;29(22):32594-32604. DOI 10.1007/s11356-022-19241-8.

- Last L., Widmer F., Fjellstad W., Stoyanova S., Kölliker R. Genetic diversity of natural orchardgrass (*Dactylis glomerata* L.) populations in three regions in Europe. *BMC Genet.* 2013;14:102. DOI 10.1186/1471-2156-14-102.
- Latypova G.M., Ivanova D.F., Davletshina R.Ya., Urazlina O.I. Composition of organic acids in plants of the genus primrose. *Sibirskii Medicinskii Zhurnal = Siberian Medical Journal.* 2014;3:96-98. (in Russian)
- Leiva A., Granados-Chinchilla F. Fatty acid profiling in animal feeds and related food matrixes using a fast GC/MS method and in situ derivatization. *JAEFES.* 2020;4(1):7-89. DOI 10.31015/jaefs.2020.9.
- Loskutov I., Shelenga T., Blinova E., Gnutikov A., Konarev A. Metabolomic profiling in evaluation of cultivated oat species with different ploidy level. *BIO Web Conf.* 2021;36:01026. DOI 10.1051/bioconf/20213601026.
- Loskutov I.G., Shelenga T.V., Konarev A.V., Vargach Yu.Yu., Porokhovina E.A., Blinova E.V., Gnutikov A.A., Rodionov A.V. A new approach to structuring the varietal diversity of naked and filmy forms of cultivated oats (*Avena sativa* L.). *Ekologicheskaya Genetika = Ecological Genetics.* 2020;18(1):27-41. DOI 10.17816/ecogen 12977. (in Russian)
- Lotova L.I. Botany. Morphology and anatomy of higher plants. Moscow: KomKniga Publ., 2007. (in Russian)
- Mahfuz S., Shang Q., Piao X. Phenolic compounds as natural feed additives in poultry and swine diets: a review. *J. Animal Sci. Biotechnol.* 2021;12:48. DOI 10.1186/s40104-021-00565-3.
- Mandrone M., Marincich L., Chiochio I., Zannini P., Guarino R., Poli F. Metabolomic study of *Dactylis glomerata* growing on Aeolian archipelago (Italy). *Metabolites.* 2022;12(6):533. DOI 10.3390/metabo12060533.
- Maslennikov P.V., Chupakhina G.N., Skrypnik L.N. Content of phenolic compounds in medicinal plants of the Botanical garden. *Izvestiya RAN Biologicheskaya seria = Proceedings of Russian Academy of Sciences. Biological Series.* 2013;5:551-557. DOI 10.7868/s000233291305010x. (in Russian)
- Maslennikov P.V., Chupakhina G.N., Skrypnik L.N., Mal'tseva E.Yu., Poltavskaya R.L. The content of low-molecular antioxidants in medicinal plants of the Kaliningrad region. *Khimiya Rastitel'nogo Syr'ya = Chemistry of Plant Stock.* 2012;3:127-133. (in Russian)
- Matveeva T.V., Sokornova S.V. Biological traits of naturally transgenic plants and their evolutionary roles. *Fiziologiya Rastenii = Russian Journal of Plant Physiology.* 2017;64(5):323-336. DOI 10.1134/S1021443717050089. (in Russian)
- Misin V.M., Sazhina N.N., Zav'yalov A.Yu. Seasonal dynamics of changes in the content of phenolic antioxidants in plantain and dandelion leaves. *Khimiya Rastitel'nogo Syr'ya = Chemistry of Plant Stock.* 2010;3:103-106. (in Russian)
- Noiraud N., Mauroussat L., Lemoine R. Transport of polyols in higher plants. *Plant Physiol. Biochem.* 2001;39(9):717-728. DOI 10.1016/S0981-9428(01)01292-X.
- Novikov N.N. Biochemistry of plants. Moscow: KolosS Publ., 2012. (in Russian)
- Pandyan M., Subramanian R.K., Subramani P., Narayanan M., Wilson A., Shunmuga K.P., Manikandan R. Global analysis of threonine metabolism genes unravel key players in rice to improving the abiotic stress tolerance. *Sci. Rep.* 2018;8:9270. DOI 10.1038/s41598-018-27703-8.
- Rasmussen S., Parsons A.J., Jones C.S. Metabolomics of forage plants: a review. *Ann. Bot.* 2012;110(6):1281-12890. DOI 10.1093/aob/mcs023.
- Sazhina N.N., Misin V.M. Measurement of the total content of phenolic compounds in various parts of medicinal plants. *Khimiya Rastitel'nogo Syr'ya = Chemistry of Plant Stock.* 2011;3:149-152. (in Russian)
- Sampieva K.T., Oganova G.M., Ivashov M.N., Chuklin R.E., Guseinov A.K. Study of the effects of certain amino acids in hypoxic hypoxia. *Biomeditsina = Biomedicine.* 2010;4:122-123. (in Russian)
- Shelenga T.V., Solov'eva A.E., Shevarda A.L., Konarev A.V. Research of the VIR collection cultures metabolome. In: Abstracts of the international scientific conference 'Plant genetic resources – the basis for food security and improving the quality of life' dedicated to the 120th anniversary of VIR. October 6–8, 2014. St. Petersburg: VIR, 2014;98. (in Russian)
- Shilova I.V., Baranovskaya N.V., Syrchina A.I., Baranova O.V., Dudko V.V., Semenov A.A., Suslov N.I. Amino acid and elemental composition of the active fraction of the Siberian clematis. *Voprosy Biologicheskoy, Medicinskoy i Farmatsevticheskoy Khimii = Issue of Biological, Medical and Pharmaceutical Chemistry.* 2008;3:34-37. (in Russian)
- Shkrobot'ko P.Yu., Popov D.M., Fursa N.S. Amino acid composition of underground organs of Fory's valerian and elder leaf valerian. *Farmatsiya = Farmacia.* 2009;7:19-23. (in Russian)
- Shurson G.C., Kerr B.J., Hanson A.R. Evaluating the quality of feed fats and oils and their effects on pig growth performance. *J. Animal Sci. Biotechnol.* 2015;6:10. DOI 10.1186/s40104-015-0005-4.
- Solovyeva A.E., Shelenga T.V., Shavarda A.L., Burlayaeva M.O. Comparative analysis of wild and cultivated *Lathyrus* L. spp. according to their primary and secondary metabolite contents. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2019;23(6):667-674. DOI 10.18699/VJ19.539.
- Solovyeva A.E., Shelenga T.V., Shavarda A.L., Burlayaeva M.O. Comparative analysis of wild and cultivated *Lathyrus* L. species to assess their content of sugars, polyols, free fatty acids, and phytosterols. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2020;24(7):730-737. DOI 10.18699/VJ20.667.
- Stewart A.V., Ellison N.W. *Dactylis*. In: Kole C. (Ed.). Wild crop relatives: Genomic and breeding resources. Millets and grasses. Heidelberg: Springer: 2011;73-87. DOI 10.1007/978-3-642-14255-0_5.
- Tulinov A.G., Kosolapova T.V., Mikhailova E.A. Results of the evaluation of collection accessions of *Dactylis glomerata* L. in conditions of republic Komi. *Zemledelie = Husbandry.* 2019;3:41-43. DOI 10.24411/0044-3913-2019-10311. (in Russian)
- Yakovleva L.V., Kalashnik M.V., Zaika I.B., Gracheva L.S., Fesenko M.A., Evdokimova Z.Z., Gadzhiev N.M., Bekish L.P., Ivanova N.V., Pozdnyakov V.A., Andruschenko A.V. Catalogue of forage crop varieties breeding in Leningrad Research Agriculture Institute Branch of Russian Potato Research Centre. St-Petersburg: Leningrad Research Agriculture Institute Publ., 2015. (in Russian)

ORCID ID

N.Yu. Malysheva orcid.org/0000-0002-5688-6694
T.V. Shelenga orcid.org/0000-0003-3992-5353
A.E. Solovyeva orcid.org/0000-0002-6201-4294
T.B. Nagiev orcid.org/0000-0003-4594-5799
N.V. Kovaleva orcid.org/0000-0002-9020-8336
L.L. Malyshev orcid.org/0000-0002-8595-1336

Acknowledgements. The article was funded by the Ministry of Science and Higher Education of the Russian Federation under agreement No. 075-15-2021-1050 (September 28, 2021). The research was performed on the material from the collection of cocksfoot genetic resources held by VIR.

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

Received May 4, 2022. Revised November 28, 2022. Accepted November 30, 2022.

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

Alkaloid content variability in the seeds of narrow-leafed lupine accessions from the VIR collection under the conditions of the Russian Northwest

M.A. Vishnyakova , A.V. Salikova, T.V. Shelenga, G.P. Egorova, L.Yu. Novikova

Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia

 m.vishnyakova.vir@gmail.com

Abstract. Alkaloid content was assessed in the seeds of 59 narrow-leafed lupine (*Lupinus angustifolius* L.) accessions from the VIR collection in the environments of Leningrad Province. The selected set included accessions of different statuses (wild forms, landraces, and advanced cultivars) and different years of introduction to the collection. Alkaloids were analyzed using gas-liquid chromatography coupled with mass spectrometry. Concentrations of main alkaloids: lupanine, 13-hydroxylupanine, sparteine, angustifoline and isolupanine, and their total content were measured. The total alkaloid content variability identified in the seeds of the studied set of accessions was 0.0015 to 2.017 %. In most cases, the value of the character corresponded to the accession's status: modern improved cultivars, with the exception of green manure ones, entered the group with the range of 0.0015–0.052 %, while landraces and wild forms showed values from 0.057 to 2.17 %. It is meaningful that the second group mainly included accessions that came to the collection before the 1950s, i. e., before the times when low-alkaloid cultivars were intensively developed. Strong variability of the character across the years was observed in the accessions grown under the same soil and climate conditions in both years. In 2019, the average content of alkaloids in the sampled set was 1.9 times higher than in 2020. An analysis of weather conditions suggested that the decrease in alkaloid content occurred due to a significant increase in total rainfall in 2020. Searching for links between the content of alkaloids and the type of pod (spontaneously non-dehiscent, or cultivated, spontaneously dehiscent, or wild, and intermediate) showed a tendency towards higher (approximately twofold in both years of research) total alkaloid content in the accessions with the wild pod type and the nearest intermediate one compared to those with the pod non-dehiscent without threshing. The correlation between the average total alkaloid content and seed color, reduced to three categories (dark, or wild, light, or cultivated, and intermediate), was significantly stronger in the group with dark seeds (5.2 times in 2019, and 3.7 times in 2020). There were no significant differences in the percentage of individual alkaloids within the total amount either between the years of research or among the groups with different pod types or the groups with different seed coat colors.

Key words: narrow-leafed lupine; alkaloids; domestication traits; spontaneously dehiscent pods; nondehiscent pods; seed color.

For citation: Vishnyakova M.A., Salikova A.V., Shelenga T.V., Egorova G.P., Novikova L.Yu. Alkaloid content variability in the seeds of narrow-leafed lupine accessions from the VIR collection under the conditions of the Russian Northwest. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):119-128. DOI 10.18699/VJGB-23-17

Изменчивость содержания алкалоидов в семенах люпина узколистного у образцов коллекции ВИР в условиях Северо-Запада Российской Федерации

M.A. Вишнякова , A.V. Саликова, Т.В. Шеленга, Г.П. Егорова, Л.Ю. Новикова

Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия

 m.vishnyakova.vir@gmail.com

Аннотация. Изучали содержание алкалоидов в семенах люпина узколистного (*Lupinus angustifolius* L.) у 59 образцов из коллекции ВИР в условиях Ленинградской области. В выборку были включены образцы разного статуса (дикие формы, староместные сорта, сорта научной селекции) и различных лет поступления в коллекцию. Алкалоиды определяли методом газожидкостной хроматографии, сопряженной с масс-спектрометрией. Определены концентрации основных алкалоидов в семенах: люпанина, 13-гидроксилюпанина, спартеина, ангустифолина, изолюпанина и их суммарное содержание. Выявленная изменчивость суммарного содержания алкалоидов в семенах изучаемой выборки составляла 0.0015–2.017 %. В большинстве случаев значение признака соответствует статусу образца: сорта современной селекции, за исключением сидеральных, входят в группу с показателями 0.0015–0.052 %, в то время как старые, местные сорта и дикие формы имеют значения 0.057–2.17 %. Характер-

но, что ко второй группе относятся преимущественно образцы, поступавшие в коллекцию до 1950-х гг., т. е. до периода активной селекции низкоалкалоидных сортов. Отмечена сильная межгодовая изменчивость признака у образцов, выращиваемых в одних и тех же почвенно-климатических условиях в течение двух лет. В 2019 г. в среднем по выборке содержание алкалоидов было в 1.9 раза выше, чем в 2020 г. Анализ погодных условий вегетации позволяет предположить, что снижение содержания алкалоидов произошло за счет значительного увеличения суммы осадков в 2020 г. При поиске связей содержания алкалоидов с типом боба (не вскрывающийся без обмолота – культурный, спонтанно вскрывающийся – дикий и промежуточный) наблюдается тенденция к более высокому (примерно в 2 раза в оба года исследования) суммарному содержанию алкалоидов у образцов с диким типом боба и приближенным к нему промежуточным по сравнению с не вскрывающимся без обмолота бобом. Связь среднего суммарного содержания алкалоидов с окраской семени, сведенной к трем категориям (темная – дикая, светлая – культурная и промежуточная), была достоверно выше у группы с темными семенами: в 5.2 раза в 2019 г. и в 3.7 раза в 2020 г. Не обнаружено достоверных различий процентного содержания отдельных алкалоидов в общей сумме алкалоидов как между годами исследования, так и между группами с различным типом боба и с разной окраской семени.

Ключевые слова: люпин узколистный; алкалоиды; признаки доместикиции; спонтанно вскрывающийся боб; невскрывающийся боб; окраска семени.

Introduction

Narrow-leaved lupine (*Lupinus angustifolius* L., Fabaceae) is a species that has been cultivated as a crop for feed and food for less than 100 years. It was exploited for centuries as a green manure crop. Feeding the seeds of this high-protein plant to animals was possible only after soaking them in water with repeated water changes to extract antinutritional compounds – a complex of quinolizidine alkaloids. It was this feature that limited the use of the plant in fodder production, since alkaloids added bitterness to the feed and in high concentrations were toxic to animals and humans.

The development of fodder cultivars was triggered by the discovery of low-alkaloid mutants (Sengbusch, 1931, 1942) and identification of the recessive mutations determining this trait: *iucundus*, *esculentus*, and *depressus* (Hackbarth, Troll, 1956). This event genetically underpinned the development of low-alkaloid forms and was regarded as the beginning of the species' domestication (Gladstones, 1970). Nowadays, many fodder cultivars have been released for animal feed purposes and the possibility to use narrow-leaved lupine seeds in food production emerged (Vishnyakova et al., 2020).

The polymorphism of alkaloid content observed before the discovery of said mutants among wild forms of narrow-leaved lupine was 0.4–3.0 % dry weight (DW) for seeds and 0.3–0.5 % DW for herbage (Święcicki W., Święcicki W.K., 1995; Brummund, Święcicki, 2011). After the release of numerous cultivars based predominantly on one *iuc* mutation, this polymorphism significantly increased. In a recent study by Polish scientists, who analyzed 329 lupine accessions, the character's variability was recorded within the range from 0.0005 to 2.8752 % (Kamel et al., 2016). Currently, the threshold value for the content of alkaloids in seeds of food or feed lupine cultivars in a number of European countries and Australia is no more than 0.02 % DW (Frick, 2017). In Russia, the permissible level of alkaloid content ranges from 0.1 to 0.3 % DW for seeds of fodder lupine (State Standard R 54632-2011, 2013) and 0.04 % for food lupine seeds (according to the existing technical specifications developed by the Research Institute of Lupine (Specification No. 9716-004-0068502-2008).

In routine practice, the content of alkaloids in seeds at the level of 0.05 % is considered the boundary value to distin-

guish between high-alkaloid (bitter) and low-alkaloid (sweet) lupines (Lee et al., 2007).

The content of alkaloids is very responsive to the impact of environmental factors, such as droughts, air temperature, geographic location, insolation level, agricultural practices, and the presence of pathogens (Christiansen et al., 1997; Cowling, Tarr, 2004; Ageeva et al., 2020). Moreover, the concentration of alkaloids in seeds of the same genotype under different growing conditions can show at least twofold variation, reaching even a tenfold increase, thus exceeding the required alkaloid content threshold and turning lupine genotypes traditionally classified as sweet into bitter ones (Cowling, Tarr, 2004; Reinhard et al., 2006; Romanchuk, Anokhina, 2018).

Along with a radical reduction of seed alkaloid content, the crop's breeding improvement includes elimination of spontaneous pod dehiscence (opening) determined by the *le* (*lentus*) and *ta* (*tardus*) alleles, introgression of the genes responsible for early flowering and the absence of the need for vernalization (*Jul* and *Ku*) into the genotypes of cultivars as well as the genes controlling seed coat permeability (*moll* – *mollis*), and white color of flowers and seeds (*leuc* – *leucospermus*) (Taylor et al., 2020).

The narrow-leaved lupine collection held by VIR includes 887 accessions from 26 countries. There are 261 cultivars developed by scientific breeding, 370 genotypes representing breeding material, 142 landraces and local varieties, 55 wild forms, and 50 accessions with an unclear status (Vishnyakova et al., 2021). The diversity of breeding statuses and the presence of wild relatives provide a rather motley picture of the presence/absence of domestication traits in the collection's accessions. Many accessions have pods spontaneously dehiscent to various degrees, and all seed colors known for this species are present. Such versatility makes it possible to trace whether there are links among domestication traits in the accessions. Therefore, the objective of this study was to identify the degree of variability in the concentration of alkaloids in narrow-leaved lupine seeds under the impact of growing conditions during two years of research and analyze correlations of this character with seed color and the degree of spontaneous pod dehiscence in a set of accessions from the VIR collection.

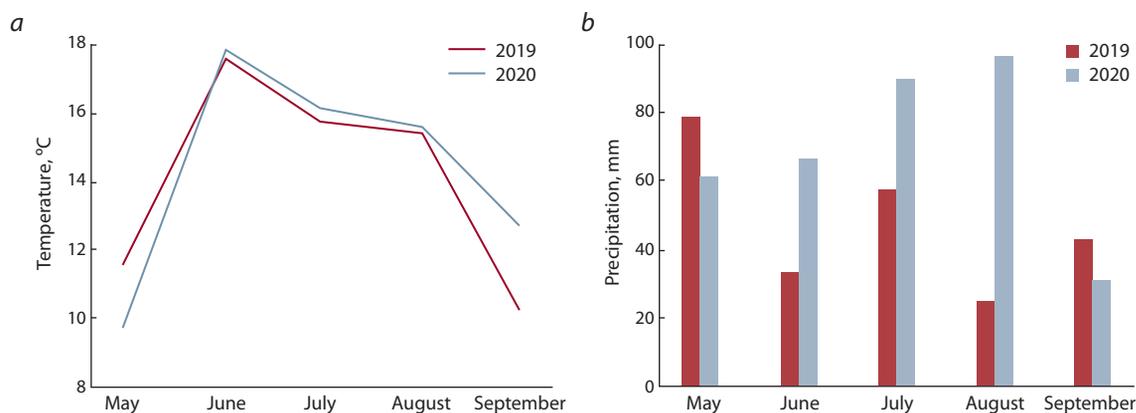


Fig. 1. Weather conditions during the experiment: *a* – mean monthly air temperature; *b* – monthly precipitation amount.

Materials and methods

Material. A set of 59 narrow-leaved lupine accessions from the VIR collection (Supplementary Material)¹, grown in the experimental fields of VIR (Pushkin, St. Petersburg) for two field seasons (2019–2020), served as the material for this study. The set consisted of accessions from 20 countries included in the collection in different years and having different breeding statuses: scientifically improved cultivars, local varieties, breeding lines, and wild forms.

Weather conditions during the experiment. The sums of active temperatures amounted to 1966 °C in 2019, and 2052 °C in 2020. Precipitation amounts for the period with temperatures above 10 °C were 175 mm in 2019, and 293 mm in 2020. Mean values for the last 30 years (1992–2021) were 2209 °C and 306 mm, respectively. Thus, the years of research were cooler and drier than the long-term average. The precipitation amount during the active growing season in 2019 was lower by 118 mm, or 1.7 times, than in 2020, with a comparable heat supply. Differences between the years in the precipitation amounts were particularly significant during the pod ripening period: 58 mm vs. 91 mm in July, and 25 mm vs. 97 mm in August, respectively. Air temperatures and precipitation amounts by months are shown in Fig. 1.

Alkaloid content measurement in seeds. Each accession selected for the study was represented by 8 plants. An average sample (30 g) was taken from the mixture of seeds. The seeds were ground to flour (50–100 μm) in a Lab Mill 1 QC-114 (Hungary). The qualitative and quantitative compositions of alkaloids in narrow-leaved lupine seeds were assessed according to a previously published protocol (Kushnareva et al., 2020).

Ethyl acetate (8 mL) and 15 % NaOH solution (2 mL) were added to a 500 mg sample of flour and incubated at +6 °C for 18 hours. The resulting extract, containing alkaloids in the form of bases, was separated from the precipitate by filtration through a paper filter. A solution of caffeine in ethyl acetate (1 mg/mL) was used as an internal standard. The composition of alkaloids was analyzed using gas-liquid chromatography coupled with mass spectrometry on an Agilent6850 A chro-

matograph (Agilent Technologies, Santa-Clara, CA, USA). The mixture was separated on an AgilentHP-5MS capillary column (5 % phenyl, 95 % methylpolysiloxane; 25 μm). Heating program: +170 °C to +320 °C, heating rate: 4 °C/min. Temperature of the mass spectrometer detector: +250 °C, injector temperature: +300 °C, injected sample volume: 1.2 μL, carrier gas (helium) flow rate: 1.5 mL/min. Chromatogram recording started after 4 min, which was necessary for the solvent to exit, and continued for 38 min. The analysis was performed in three analytical replicates.

Compounds were identified using the AMDIS program (Automated Mass Spectral Deconvolution and Identification System, National Institute of Standards and Technology, USA, Version 2.69, <http://www.amdis.net>). The NIST 2010 library (National Institute of Standards and Technology, USA, <http://www.nist.gov>) was employed for the analysis.

Alkaloid content was calculated according to the internal standard (caffeine, concentration: 1 μg/μL) using the UniChrom 5.0.19 program. The results of alkaloid content (absolute values) in narrow-leaved lupine seeds are given in mg/100 g DW. The percentage (%) of alkaloids (relative values) was calculated taking into account the proportion of an individual compound in the total alkaloid content, the latter being the sum of alkaloid values in an accession (mg/100 g DW). Mean values were calculated taking into account the resulting data of analytical replicates for each accession (see Supplementary Material).

The presence/absence of spontaneous pod dehiscence was assessed. It is better to describe this character shortly after harvesting, before the pods have reached the air-dry state, which provokes dehiscence even in such pods that were closed at the time of harvesting. Under our conditions, however, dry pods were assessed. On the one hand, it helped to reliably identify the type of pods nondehiscent without threshing; on the other hand, it hampered unambiguous identification of the pod opening time: whether the dehiscence of pods happened at harvesting or after complete drying. Therefore, this character was ranked according to the nature of the valves. The wild type (spontaneously dehiscent pods) had twisted valves (type 1). The cultivated type (nondehiscent pods) had flat valves, completely closed or slightly open (type 3). The remaining pods

¹ Supplementary Material is available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Vishnyakova_Engl_27_2.pdf

were open-valve, but flat or with some tendency to curl: they were classified into the intermediate type (type 2). A certain conventionality of the latter type and its closeness to the spontaneously dehiscent pod type should be recognized.

The seed coat color was also divided into three categories: dark (1), intermediate (2), and light (3).

Statistical processing. MS Excel programs and the Statistica 13.3 package (TIBCO Software Inc., USA) were used for data visualization. Statistical analysis was made in the Statistica 13.3 package.

Statistical significance of differences in alkaloid content in 2019 and 2020 was studied using Student's *t*-test for dependent (paired) samples (Dospekhov, 1973; Khalafyan, 2010). The difference in the characteristics of an accession in two versions of the experiment was calculated (in our case, between the years of research) and the significance of the mean difference of the accessions from zero was assessed using the *t*-test. Such criterion is more precise than a comparison of the differences between the means of independent samples, as it does not depend on the nature of the indicator's distribution within the sample.

The average alkaloid contents in three groups of accessions with different pod types were compared using the analysis of variance; the same approach was applied for the groups with different seed colors. Correlation coefficients were calculated for alkaloid content separately in 2019 and 2020. The strength of correlations was assessed according to B.A. Dospekhov (1973): if the correlation coefficient is higher than 0.7 in its absolute value, it is strong; from 0.3 to 0.7, it is medium; less than 0.3, it is weak. The significance level of 5 % was adopted for this study.

Results

Previously, the authors tested extraction techniques on alkaloids from leaves of the green manure cultivar Oligarkh (k-3814) reproduced in 2018 (Pushkin) and clarified the

qualitative composition of its alkaloid complex. The cultivar's leaves contained five alkaloids: lupanine (L), 13-hydroxylupanine (H), angustifoline (A), sparteine (S), and isolupanine (I), plus traces of their derivatives or unidentifiable alkaloids numbering up to 120 in narrow-leaved lupine (Frick et al., 2017). The qualitative composition of main (detectable) alkaloids in seeds identified in the present study corresponded to our previous findings for vegetative organs. Their content in seeds varied as follows: 70.0–85.4 % for L, 6.4–17.2 % for H, 0.7–2.0 % for A, 4.0–12.6 % for S, and 0.5–1.4 % for I. The variability of the total alkaloid content was 0.0015–2.017 % (Table 1, see Supplementary Material).

The mean alkaloid level in the accessions was 501.7 mg/100 g DW in 2019, which was significantly (by 90.5 %) higher than the same value for the seeds reproduced in 2020 – 263.6 mg/100 g DW (statistical significance of differences according to Student's *t*-test for dependent samples was $p = 0.009$). In 2020, a decrease in the mean alkaloid content values was observed: L was 389.7 in 2019 vs. 203.4 in 2020 ($p = 0.008$); H: 59.0 vs. 31.3 ($p = 0.014$); S: 41.9 vs. 23.0 ($p = 0.017$); A: 6.8 vs. 3.6 ($p = 0.014$); I: 4.4 vs. 2.4 ($p = 0.023$) (see Table 1 and Supplementary Material).

However, the amount of alkaloids in eight accessions (k-3172, 3457, 3947, 3607, 3526, 1546, 2856, and 3062) increased in 2020 compared to 2019.

L content increased in 6 accessions, H in 13, S in 12, A in 9, and I in 10. It should be mentioned that, according to the International COMECON list of descriptors for the genus *Lupinus* L. (Stepanova et al., 1985), five accessions from this group were classified as having “very low” alkaloid content (its amount in seeds was less than 25 mg/100 g), two as “medium” (from 100 to 300 mg/100 g), and only one accession (cv. Oligarkh, k-3814) had “very high” content (more than 300 mg/100 g). Characteristically, the accessions with very low or medium alkaloid content manifested insignificant differences across the two years: for example, k-2856 had

Table 1. Mean alkaloid content in the set of 59 narrow-leaved lupine accessions for two years of research (Pushkin, 2019–2020)

Concentration of alkaloids in seeds		2019			2020		
		Mean	Min	Max	Mean	Min	Max
Total alkaloids	mg/100 g DW	501.7±80.7	4.0	2017.4	263.6±38.6	1.7	898.8
Lupanine	mg/100 g DW	389.7±62.5	3.1	1573.0	203.4±29.8	1.4	729.0
	%	78.1±0.4	73.1	85.4	77.8±0.5	70.0	84.3
13-Hydroxylupanine	mg/100 g DW	59.0±10.1	0.3	273.2	31.3±4.8	0.1	115.5
	%	11.5±0.4	6.4	17.2	11.7±0.4	6.7	17.2
Sparteine	mg/100 g DW	41.9±6.9	0.2	173.4	23.0±3.6	0.1	85.2
	%	8.2±0.3	4.4	12.0	8.3±0.2	4.0	12.6
Angustifoline	mg/100 g DW	6.8±1.2	0.1	35.4	3.6±0.6	0.0	15.0
	%	1.3±0.1	0.7	2.0	1.3±0.1	0.7	1.9
Isolupanine	mg/100 g DW	4.4±0.8	0.0	23.8	2.4±0.4	0.0	9.3
	%	0.9±0.0	0.5	1.4	0.9±0.0	0.5	1.3

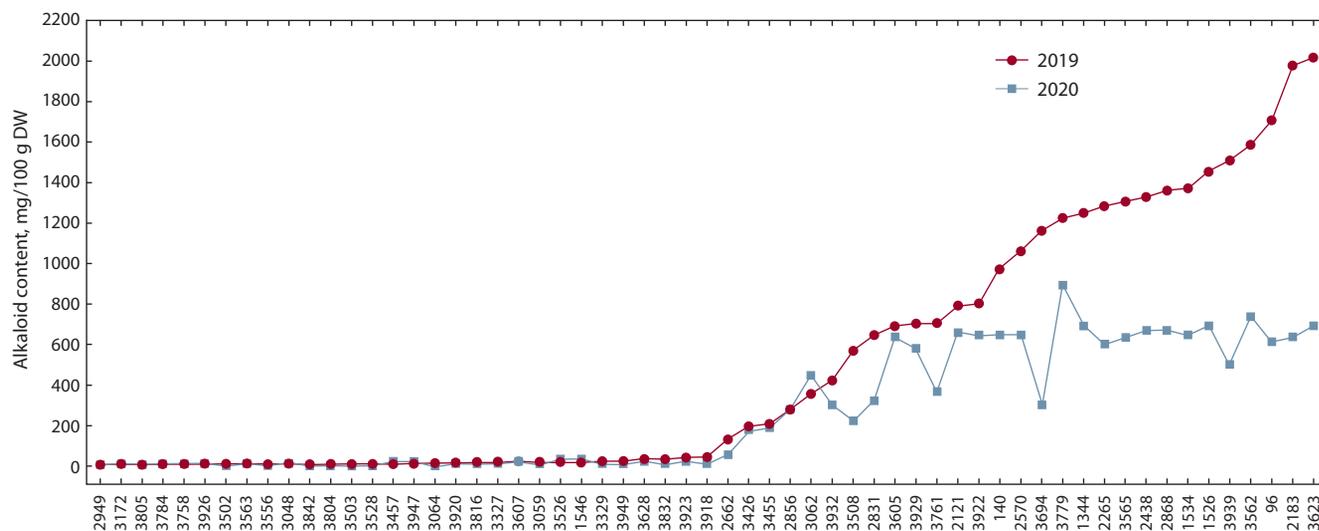


Fig. 2. Alkaloid content in seeds of 59 narrow-leaved lupine accessions in the two years of research.

279.1 and 280.9 mg/100 g DW (0.7 %); k-3607, 21.2 and 23.2 mg/100 g DW (1.1 %); k-3172, 5.3 and 6.8 mg/100 g DW (1.3 %), respectively (Fig. 2).

The character's low variability in low-alkaloid lupine forms, as observed above, was characteristic of all low-alkaloid accessions in the tested set, regardless of the increase or decrease in alkaloid concentrations across the years of research (see Fig. 2). Meanwhile, alkaloid content values remained within the established range: none of the sweet accessions with alkaloid content below 50 mg/100 g exceeded those values and did not shift into the bitter category. For example, the amount of alkaloids in cv. Yan (k-3832) was 7 mg/100 g in 2020 and 4.7 times higher (34 mg/100 g) in 2019.

Cv. Gerkules (k-3923) had 20 mg/100 g in 2020, and a more than twice higher amount in 2019 (47 mg/100 g), but in both cases the cultivar remained in the sweet category.

On average for the studied set of accessions, the proportion of individual alkaloids varied insignificantly over the years. The relative content of L in 2019 and 2020 was 78.1 and 77.8 %; H: 11.5 and 11.7 %, and S: 8.2 and 8.3 %, respectively. The shares of A (1.3 %) and I (0.9 %) in the composition of alkaloids did not change during the period of research. Thus, the relative content of individual alkaloids may be recognized as a fairly constant indicator.

A strong correlation was observed between the absolute content values (mg/100 g) of individual alkaloids: 0.89–0.96 in 2019 and 0.88–0.95 in 2020. The correlation between the amounts of individual compounds and total alkaloids was 0.94–0.999. The strongest relationship was observed between the total alkaloid content and L values: 0.999 in 2019 and 0.998 in 2020.

Pairwise correlations between the percentage (relative) contents of the studied alkaloids were mostly insignificant; no systematic shifts in the alkaloid composition structure were observed over the years. A *t*-test for dependent samples showed a significance level of differences in the percentage of individual alkaloids: from $p = 0.063$ to $p = 0.082$. Variations

of alkaloid composition in individual accessions were induced by changes in the representation of two main alkaloids, L and G: an increase in the proportion of one led to a decrease in the proportion of the other. Meanwhile, lupanine remained dominant in the composition of alkaloids in narrow-leaved lupine.

The types of pod dehiscence and seed coat color were analyzed for 45 accessions from the studied set. Twelve accessions were classified according to their pod characteristics into type 1 (wild, with twisted valves), 16 into type 2 (intermediate), and 17 into type 3 (nondehiscent without threshing) (Fig. 3 and Supplementary Material). There were no significant differences among the groups of accessions with different pod types in either the absolute or relative alkaloid content (Table 2, Fig. 4). With this in view, in both years of research, the highest alkaloid content was recorded for type 1 (693.7 mg/100 g DW in 2019, and 345.3 mg/100 g DW in 2020), and the lowest, for type 3 (320.3 mg/100 g DW in 2019, and 200.1 mg/100 g DW in 2020). Medium values were shown for type 2 (612.1 mg/100 g DW in 2019 and 300.7 mg/100 g DW in 2020).

Consequently, higher total alkaloids were characteristic of the accessions with the wild pod type. They exceeded the accessions with the nondehiscent pod type 2.3 (2019) and 1.8 times (2020). However, taking into account the high variability in the absolute and relative content of individual alkaloids and their total amount, no significant differences were observed among the groups (see Table 2, Fig. 4). The contributions of individual alkaloids to the total content did not depend on the pod type.

According to the color of the seed coat, 15 accessions were characterized as dark-seeded, 19 were of the intermediate type (between dark and light), and 11 were light-seeded.

Differences in the studied traits among the above-mentioned groups were assessed as statistically significant at a 10 % significance level. In 2019, all three groups significantly differed from each other in the following parameters: L ($p = 0.063$),



Fig. 3. The types of narrow-leaved lupine pods according to their ability to dehisce spontaneously. Pod type designations: 1 – wild; 2 – intermediate; 3 – cultivated (nondehiscent without threshing).

Table 2. Total content of alkaloids in the groups of lupine accessions with different pod types

Pod type	Number of accessions	2019			2020		
		Mean	Min	Max	Mean	Min	Max
Wild	12	693.7 ± 180.9	5.7	1508.7	345.3 ± 90.6	2.9	692.9
Intermediate	16	612.1 ± 183.2	5.3	2017.4	300.7 ± 78.6	4.3	736.3
Cultivated	17	320.3 ± 136.3	6.0	1976.8	200.1 ± 74.1	4.4	898.8
Total	45	523.6 ± 96.9	5.3	2017.4	274.6 ± 46.2	2.9	898.8

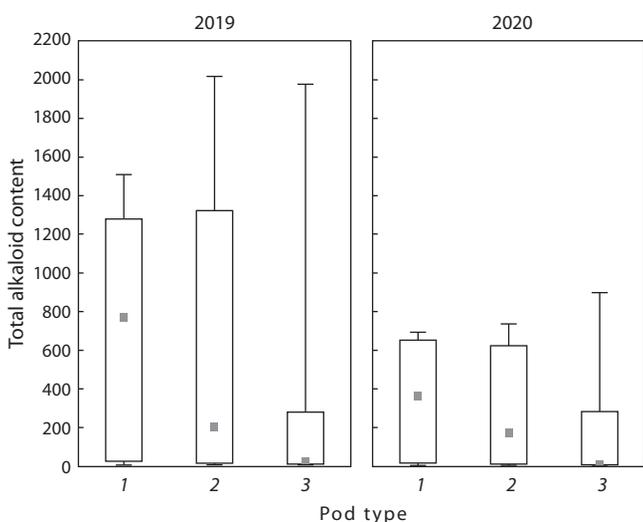


Fig. 4. Total alkaloid content in lupine groups with different pod types in 2019 and 2020.

Pod type designations: 1 – wild; 2 – intermediate; 3 – cultivated.

H ($p = 0.066$), S ($p = 0.070$), I ($p = 0.075$), and total alkaloids ($p = 0.062$) (Table 3, Fig. 5); in 2020, in L ($p = 0.083$), H ($p = 0.055$), S ($p = 0.060$), and total alkaloids ($p = 0.074$). For A and I, significance levels of differences were 0.108 and 0.130, respectively. The highest values of the total alkaloid content were recorded for the dark-seeded group (660.4 mg/100 g DW in 2019 and 334.4 mg/100 g DW in 2020), while the lowest values were found in the light-seeded group (125.9 mg/100 g DW in 2019 and 90.6 mg/100 g DW in 2020) (see Table 3, Fig. 5).

Significant differences at a 5 % level were observed between the 1st and 3rd contrasting groups of accessions: in 2019 for all indicators (L, H, S, A, I, and total alkaloids), and in 2020 for L, H, S, and total alkaloids. The dark-seeded group exceeded the light-seeded one in the mean value of total alkaloids 5.2 times in 2019 ($p = 0.023$) and 3.7 times in 2020 ($p = 0.030$).

Differences among the three color groups in the percent contribution of individual compounds to the total content of alkaloids were statistically insignificant ($p > 0.462$). No

Table 3. Total alkaloid content in the groups of lupine accessions with different seed colors

Seed color	Number of accessions	2019			2020		
		Mean	Min	Max	Mean	Min	Max
Dark	15	660.4 ± 176.1	8.8	1976.8	334.4 ± 78.2	5.2	658.2
Intermediate	19	646.0 ± 161.5	6.0	2017.4	333.9 ± 77.2	4.3	898.8
Light	11	125.9 ± 88.3	5.3	975.2	90.6 ± 61.0	2.9	648.4
Total	45	523.6 ± 96.9	5.3	2017.4	274.6 ± 46.2	2.9	898.8

significant differences were recorded between the groups contrasting in seed color (dark and light) during the period of research ($p > 0.237$). Thus, lupine accessions with different seed coat colors significantly differ from each other in the total content of alkaloids, while their alkaloid composition can be recognized as constant.

Discussion

The analysis of alkaloid content in narrow-leaved lupine seeds from the VIR collection disclosed high variability of this character. In the selected set of 59 accessions, there were genotypes with minimum (0.0015 %) and maximum (2.017 %) values of this indicator: Cv. Danko from Belarus (k-2949) and an Australian breeding line (k-3623), respectively. The accessions were divided into two groups according to their alkaloid content: low-alkaloid (alkaloid content in seeds was below 0.05 %), and high-alkaloid (alkaloid content was equal to or above 0.051 %). The first group included 28 accessions, representing breeding lines and cultivars that entered the collection after 1950, mainly from Russia, Belarus, and Australia. The second group consisted of accessions received before 1950: local varieties, improved cultivars and lines from Germany, United Kingdom, Poland, and Latvia. In addition, it included wild genotypes from Greece and Spain.

It should be mentioned that in the high-alkaloid group there were several improved cultivars of contemporary breeding, for example, cv. Oligarkh (k-3814, Leningrad Research Agriculture Institute) grown for green manure, and the Australian breeding line (k-3623) also, apparently, developed for green manure purposes. Cultivars intended to be used as green manure have high vegetative weight and, as a rule, low seed productivity. These features were observed in cv. Oligarkh, characterized by rapid initial growth, early maturation, and good leafiness, which ensures high yields of green biomass and readiness for plowing 50–60 days after sowing (Lysenko, 2020). Such cultivars are usually developed without any regard to the content of alkaloids, and they may appear non-uniform in this indicator. Sporadic low-alkaloid genotypes, in their turn, occur among wild lupine forms, for example, in accessions k-3607 (Spain) and k-3457 (Greece), the alkaloid content of which did not exceed 0.021 %, thus classifying them as sweet forms of narrow-leaved lupine.

Our data are quite in agreement with the results obtained by Polish scientists who screened the collection of narrow-leaved lupine maintained at the Polish genebank and found low-alkaloid genotypes in the group of wild accessions where the range of this character was 0.0163–2.8752 % DW. Contrari-

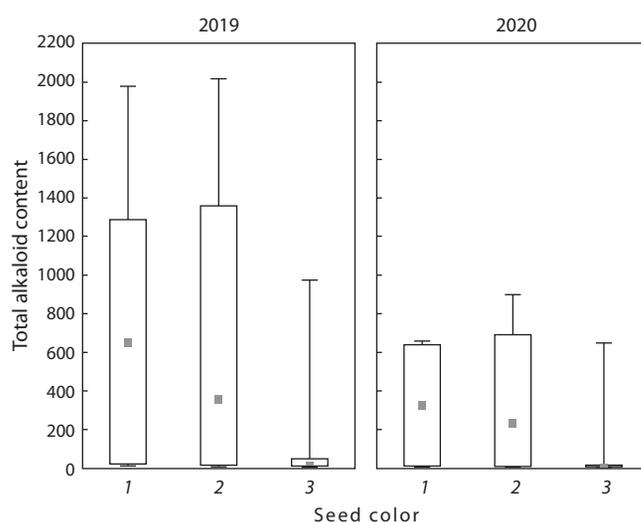


Fig. 5. Total alkaloid content in lupine groups with different seed colors in 2019 and 2020.

Seed color group designations: 1 – dark-seeded; 2 – intermediate; 3 – light-seeded.

wise, high-alkaloid accessions were identified among cultivars developed by scientific breeding, for example, cv. Karo (1.165–1.3011 %). The character's range in this group was 0.0022–2.1562 % DW (Kamel et al., 2016).

The variation of this character between the years of our experiment manifested itself in the fact that in 2020 the average alkaloid content in the studied set was 1.9 times lower than in 2019.

The high susceptibility of alkaloid content levels in lupine seeds to environmental factors has not yet been explained. The mechanisms of their impact are even called unpredictable (Frick et al., 2017). As mentioned previously, alkaloid content variability in the same genotype can be affected by a variety of environmental factors: temperature, humidity, soil characteristics and mineral composition, geographic location, etc. The amplitude of the character's variability also depends on the genotype: some cultivars are more variable under environmental impacts than others (Gremigni et al., 2001; Cowling, Tarr, 2004; Jansen et al., 2009).

Plants in this experiment were grown for two years in the same location, on a field relatively homogeneous in soil composition, and the same agricultural practices were used. Therefore, we consider weather conditions to be the main factor that could affect the content of alkaloids. The most significant

meteorological differences across the two years of research manifested themselves in an increase in the precipitation amount in 2020 compared to 2019 (see Fig. 1). A particularly noticeable rainfall deficit was observed in July and August of 2019 (58 mm and 25 mm, respectively). These are the months when seeds are swelling and ripening and alkaloids from vegetative organs accumulate in them (Vishnyakova, Krylova, 2022). Apparently, it was this factor that led to a decrease in total alkaloids on average for the studied set of 59 accessions from 501.7 mg/100 g DW in 2019 to 263.6 in 2020.

Droughts are believed to increase the content of alkaloids in lupine, but it is important at what stage of plant development the drought occurs (Christiansen et al., 1997). An increase in the level of alkaloids under drought conditions was observed in a number of plant species: *Nicotiana*, *Papaver somniferum*, and *Catharanthus roseus* (Waller, Nowacki, 1978; Szabó et al., 2003; Jaleel et al., 2007; Amirjani, 2013). Stresses are presumed to increase the synthesis of secondary metabolites, such as isoprenoids, phenols, and alkaloids (Selmar, Kleinwächter, 2013). In view of this, temporary exposure to drought is recommended to intensify the synthesis and increase the yield of alkaloids in medicinal and spicy herbs (Kleinwächter et al., 2015; Kleinwächter, Selmar, 2015).

High air temperatures from the start of flowering to pod maturation are also considered to be a factor raising alkaloid concentration in narrow-leafed lupine (Jansen et al., 2009). Under the conditions of our experiment, the year 2019, when the accumulation of alkaloids was at its peak, was on the whole much colder than either 2020 or the long-term mean value, but the temperatures during the growing season in both years were comparable. Therefore, we consider precipitation to be the decisive factor in the variation of this character across the two years of research.

No significant differences in the percentage of individual alkaloids within their total amount were found between the years of research. Their average contribution was as follows: 77.9 % of lupanine, 11.6 % of 13-hydroxylupanine, 8.3 % of sparteine, 1.3 % of angustifoline, and 0.9 % of isolupanine.

The almost twofold increase in alkaloid concentration, averaged for the studied set of accessions, was typical only for high-alkaloid and medium-alkaloid genotypes. For accessions with alkaloid content less than 0.05 %, this indicator changed relatively little in both years. It is quite possible that these accessions reached the lowest alkaloid accumulation threshold for narrow-leafed lupine seeds. In any case, these very low-alkaloid genotypes can be regarded as stable in the manifestation of the trait.

Spontaneous pod dehiscence is one of the key features differentiating wild species from cultivated ones in legumes. During pod maturation and drying, the valves of wild genotypes suddenly spontaneously open along the dorsal and ventral sutures and rapidly twist along their axis spirally in opposite directions, giving the opened pods a typical V-shaped appearance (Maysuryan, Atabekova, 1974). Wild species use this mechanism to disperse seeds, while in cultivated plants it is a highly undesirable trait that leads to yield loss.

Contemporary breeders, along with the efforts to develop alkaloid-free narrow-leafed lupine cultivars, make attempts to introgress as many other domestication genes into their ge-

nomes as possible, specifically those responsible for the absence of spontaneous pod dehiscence. This trait is known to be controlled by two recessive alleles: *ta* (*tardus*), which determines the fusion of pod valves by forming a solid strand of sclerenchyma cells along the pod's perimeter (Hackbarth, Troll, 1959), and *le* (*lentus*), which changes the orientation of endocarp cells and reduces the thickness of the parchment layer (Gladstones, 1970). Only the combination of both alleles can ensure complete absence of spontaneous pod dehiscence (Anokhina et al., 2012). It is quite possible that the accessions classified by us into the intermediate pod type according to their pod dehiscence nature possess only one of these two alleles. This study pinpointed a quite obvious tendency towards higher alkaloid content in the accessions with wild-type pods and intermediate ones that were close to the wild type, compared to cultivated genotypes with nondehiscent pods. They demonstrated an almost twofold difference in both years of research.

A similar relationship was observed between the content of alkaloids and the color of seeds (seed coat). There are up to 8 grades of seed coat color recognized in narrow-leafed lupine: (1) variegated, gray, with indistinct maculation; (2) almost black, with small white speckles and spots; (3) gray with white spots; (4) white with occasional brown and gray spots; (5) beige (nut-brown), with brown spots; (6) white, dull at the hilum, without a triangular spot or a stripe; (7) white, with sporadic brown spots; and (8) pure white, glossy. (Kurlovich, 2002). Similarly to the pod dehiscence pattern, we reduced these grades to three types: (1) dark, or wild type included seeds of the 1st and 2nd seed color grades, (2) intermediate type, with the 3rd and 5th seed color grades, and (3) light, or cultivated type, incorporating the 4th, 6th, 7th and 8th color grades.

It is known that wild forms of narrow-leafed lupine have blue flowers and dark seeds. Breeders, in their efforts to improve this crop, selected plants with the *leucospermus* locus, responsible for the white color of flowers and light-colored seeds (Nelson et al., 2006; Berger et al., 2012). The same pattern was also observed in the domestication of other legume species (Ku et al., 2020). In our study, the group with dark seeds significantly exceeded the group with light seeds in the average total content of alkaloids (5.2 times in 2019, and 3.7 times in 2020). Meanwhile, no such differences were found in the percentage content of individual alkaloids either among the groups of accessions with different seed coat colors or those with different pod types.

Thus, low alkaloid content in narrow-leafed lupine seeds, acquired by a part of the crop's gene pool as a result of domestication and breeding improvement, is associated with the absence of spontaneous pod dehiscence and the light color of seeds. We regard this phenomenon as the evidence of the joint introgression of domestication genes into modern narrow-leafed lupine cultivars.

Conclusion

Development of low-alkaloid narrow-leafed lupine forms, i. e., reducing the concentration of alkaloids in lupine seeds to a level below 0.05 %, has been a priority trend in the species' improvement in the process of domestication and breeding.

The VIR collection contains cultivars for feed and food uses with the content of alkaloids no higher than 0.0015 % DW. It is this minimum value that we found while screening the set of 59 accessions. Susceptibility of this trait to the impact of environmental conditions was seen in the fact that the synthesis of alkaloids in 2019 was 1.9 times more intensive than in 2020 on average for the studied set of accessions. A significant precipitation deficit was recorded in July and August of 2019, with all other growing conditions being comparable. This stressor, apparently, was the decisive factor that provoked an abrupt increase in the synthesis of alkaloids compared to 2020.

A distinctive feature of alkaloid content variability in narrow-leafed lupine seeds under the impact of environmental conditions is relatively low variation of this character in low-alkaloid genotypes.

The observed tendency towards higher (almost twofold) alkaloid content in the accessions with spontaneously dehiscent pods than in those with pods nondehiscent without threshing in both years of research and significantly higher content of alkaloids in seeds with dark seed coat color (wild) attest to the joint introgression of these domestication traits into modern cultivars.

References

- Ageeva P.A., Pochutina N.A., Matyukhina M.V. Blue lupine – source of valuable nutrients in forage production. *Kormoproizvodstvo = Fodder Production*. 2020;10:29-33. (in Russian)
- Amirjani M. Effects of drought stress on the alkaloid contents and growth parameters of *Catharanthus roseus*. *J. Agric. Biol. Sci.* 2013; 8(11):745-750.
- Anokhina V.S., Debely G.A., Konorev P.M. Lupine: Breeding, Genetics, Evolution. Minsk. (in Russian)
- Berger J., Buirchell B., Luckett D., Nelson M. Domestication bottlenecks limit genetic diversity and constrain adaptation in narrow-leafed lupin (*Lupinus angustifolius* L.). *Theor. Appl. Genet.* 2012; 124(4):637-652. DOI 10.1007/s00122-011-1736-z.
- Brummund M., Świącicki W. The recent history of lupin in agriculture. In: Naganowska B., Kachlicki P., Wolko B. (Eds.). *Lupin Crops – an Opportunity for Today, a Promise for the Future: Proc. of the 13th Intern. Lupin Conference*. Poznań, Poland. 6-10 June, 2011. Poznań: Institute of Plant Genetics Polish Academy of Sciences, 2011;15-23.
- Christiansen J.L., Jørnsgård B., Buskov S., Olsen C.E. Effect of drought stress on content and composition of seed alkaloids in narrow-leafed lupin, *Lupinus angustifolius* L. *Eur. J. Agron.* 1997;7(4):307-314. DOI 10.1016/S1161-0301(97)00017-8.
- Cowling W.A., Tarr A. Effect of genotype and environment on seed quality in sweet narrow-leafed lupin (*Lupinus angustifolius* L.). *Aust. J. Agric. Res.* 2004;55(7):745-751. DOI 10.1071/AR03223.
- Dospekhov B.A. *Methodology of Field Experiments*. Moscow: Kolos Publ., 1973. (in Russian)
- Gladstones J. Lupins as crop plants. *Field Crop Abstr.* 1970;23(2): 123-148.
- Gremigni P., Wong M., Edwards L.K., Harris D.J., Hambiln J. Potassium nutrition effects on seed alkaloid concentrations, yield and mineral content of lupins (*Lupinus angustifolius*). *Plant Soil.* 2001; 234:131-142. DOI 10.1023/A:1010576702139.
- Hackbarth J., Troll H.J. Lupinen als Körnerleguminosen und Futterpflanzen. In: *Handbuch der Pflanzenzüchtung*. Band IV. Züchtung der Futterpflanzen. Berlin: Paul Parey, 1959;1-51.
- Frick K.M., Kamphuis L.G., Siddique K.H.M., Singh K.B., Foley R.C. Quinolizidine alkaloid biosynthesis in lupins and prospects for grain quality improvement. *Front. Plant Sci.* 2017;8:87. DOI 10.3389/fpls.2017.00087.
- Jaleel C.A., Manivannan P., Kishorekumar A., Sankar B., Gopi R., Somasundaram R. Alterations in osmoregulation, antioxidantenzymes and indole alkaloid levels in *Catharanthus roseus* exposed to water deficit. *Colloids Surf. B Biointerfaces.* 2007;59(2):150-157. DOI 10.1016/j.colsurfb.2007.05.001.
- Jansen G., Jürgens H.U., Ordon F. Effects of temperature on the alkaloid content of seeds of *Lupinus angustifolius* cultivars. *J. Agron. Crop Sci.* 2009;195(3):172-177. DOI 10.1111/j.1439-037X.2008.00356.x.
- Kamel K.A., Świącicki W., Kaczmarek Z., Barzyk P. Quantitative and qualitative content of alkaloids in seeds of a narrow-leafed lupin (*Lupinus angustifolius* L.) collection. *Genet. Resour. Crop Evol.* 2016;63:711-719. DOI 10.1007/s10722-015-0278-7.
- Khalafyan A.A. *Statistica 6. Statistical Data Analysis*. Moscow: Binom Publ., 2010. (in Russian)
- Kleinwächter M., Paulsen J., Bloem E., Schnug E., Selmar D. Moderate drought and signal transducer induced biosynthesis of relevant secondary metabolites in thyme (*Thymus vulgaris*), greater celandine (*Chelidonium majus*) and parsley (*Petroselinum crispum*). *Ind. Crops Prod.* 2015;64:158-166. DOI 10.1016/j.indcrop.2014.10.062.
- Kleinwächter M., Selmar D. New insights explain that drought stress enhances the quality of spice and medicinal plants: Potential applications. *Agron. Sustain. Dev.* 2015;35:121-131. DOI 10.1007/s13593-014-0260-3.
- Ku Y.-S., Contador C.A., Ng M.-S., Yu J., Chung G., Lam H.-M. The effects of domestication on secondary metabolite composition in legumes. *Front. Genet.* 2020;11:581357. DOI 10.3389/fgene.2020.581357.
- Kurlovich B. *Lupins: Geography, Classification, Genetic Resources and Breeding*. St. Petersburg: Intan, 2002.
- Kushnareva A.V., Shelenga T.V., Perchuk I.N., Egorova G.P., Malyshchev L.L., Kerv Yu.A., Shavarda A.L., Vishnyakova M.A. Selection of an optimal method for screening the collection of narrow-leafed lupine held by the Vavilov Institute for the qualitative and quantitative composition of seed alkaloids. *Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding.* 2020;24(8): 829-835. DOI 10.18699/VJ20.680.
- Lee M.J., Pate J.S., Harris D.J., Atkins C.A. Synthesis, transport and accumulation of quinolizidine alkaloids in *Lupinus albus* L. and *L. angustifolius* L. *J. Exp. Bot.* 2007;58(5):935-946. DOI 10.1093/jxb/erl254.
- Lysenko O.G. The value of narrow-leafed lupine. *Sel'skokhozyajstvennye Vesti = Agricultural News.* 2020;1:30. (in Russian)
- Maysuryan N.A., Atabekova A.I. *Lupin*. Moscow: Kolos Publ., 1974. (in Russian)
- Nelson M.N., Phan H., Ellwood S., Moolhuijzen P., Hane J., Williams A., Clare E., Fosu-Nyarko J., Scobie M., Cakir M. The first gene-based map of *Lupinus angustifolius* L. – location of domestication genes and conserved synteny with *Medicago truncatula*. *Theor. Appl. Genet.* 2006;113(2):225-238. DOI 10.1007/s00122-006-0288-0.
- Reinhard H., Rupp H., Sager F., Streule M., Zoller O. Quinolizidine alkaloids and phomopsins in lupin seeds and lupin containing food. *J. Chromatogr. A.* 2006;1112(1-2):353-360. DOI 10.1016/j.chroma.2005.11.079.
- Romanchuk I.Yu., Anokhina V.S. Lupine alkaloids: structure, biosynthesis, genetics. *Molekulyarnaya i Prikladnaya Genetika = Molecular and Applied Genetics.* 2018;25:108-123. (in Russian)
- Selmar D., Kleinwächter M. Influencing the product quality by deliberately applying drought stress during the cultivation of medicinal plants. *Ind. Crop Prod.* 2013;42:558-566. DOI 10.1016/j.indcrop.2013.04.019.
- Sengbusch R. Bitterstoffarme Lupinen II. *Züchter.* 1931;4:93-109.
- Sengbusch R. Susslupinen und Ollupinen. Die Entstehungsgeschichte einiger neuen Kulturpflanzen. *Landw. Jb.* 1942;91:719-880.
- Specification No. 9716-004-00668502-2008. Food Lupine. Available at: <https://e-ecolog.ru/crc/57.01.01.000.%D0%A2.000230.05.08?>

- ysclid = 14sa0dvtbn325210024 (Accessed June 24, 2022). (in Russian)
- State Standard R 54632-2011. Fodder Lupine. Specification. 2013. Available at: <https://docs.cntd.ru/document/1200093158?ysclid=14s80m228h216628534> (Accessed June 24, 2022). (in Russian)
- Stepanova S., Nazarova N., Korneichuk V., Lehmann C., Mikolai-chik Y. The international COMECON list of descriptors for the genus *Lupinus* L. Leningrad: VIR Publ., 1985. (in Russian)
- Szabó B., Tyihák E., Szabó G., Botz L. Mycotoxin and drought stress induced change of alkaloid content of *Papaver somniferum* plantlets. *Acta Bot. Hung.* 2003;45(3):409-417. DOI 10.1556/ABot. 45. 2003.3-4.15.
- Święcicki W., Święcicki W.K. Domestication and breeding improvement of narrow-leaved lupin (*L. angustifolius* L.). *J. Appl. Genet.* 1995;36(2):155-167.
- Taylor C.M., Kamphuis L.G., Cowling W.A., Nelson M.N., Berger J.D. Ecophysiology and Phenology: Genetic Resources for Genetic/Genomic Improvement of Narrow-Leafed Lupin. In: Singh K., Kamphuis L., Nelson M. (Eds.). *The Lupin Genome. Compendium of Plant Genomes*. Cham: Springer, 2020;19-30. DOI 10.1007/978-3-030-21270-4_2.
- Vishnyakova M.A., Krylova E.A. Prospects for obtaining low-alkaloid and adaptive forms of narrow-leaved lupine based on the genome and transcriptome resources of the species. *Biotekhnologiya i Selekttsiya Rastenij = Plant Biotechnology and Breeding*. 2022;5(2):5-14. DOI 10.30901/2658-6266-2022-2-o1. (in Russian)
- Vishnyakova M.A., Kushnareva A.V., Shelenga T.V., Egorova G.P. Alkaloids of narrow-leaved lupine as a factor determining alternative ways of the crop's utilization and breeding. *Vavilovskii Zhurnal Genetiki i Selekttsii = Vavilov Journal of Genetics and Breeding*. 2020; 24(6):625-635. DOI 10.18699/VJ20.656.
- Vishnyakova M.A., Vlasova E.V., Egorova G.P. Genetic resources of narrow-leaved lupine (*Lupinus angustifolius* L.) and their role in its domestication and breeding. *Vavilovskii Zhurnal Genetiki i Selekttsii = Vavilov Journal of Genetics and Breeding*. 2021;25(6):620-630. DOI 10.18699/VJ21.070.
- Waller G.R., Nowacki E.K. *Alkaloid Biology and Metabolism in Plants*. New York: Plenum Press, 1978;129-133.

ORCID ID

M.A. Vishnyakova orcid.org/0000-0003-2808-7745
A.V. Salikova (Kushnareva) orcid.org/0000-0002-5709-7961
G.P. Egorova orcid.org/0000-0002-8645-3072
T.V. Shelenga orcid.org/0000-0003-3992-5353
L.Yu. Novikova orcid.org/0000-0003-4051-3671

Acknowledgements. The work was carried out with the support from the Russian Foundation for Basic Research (Project No. 20-016-00072-A).

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

Received July 5, 2022. Revised September 13, 2022. Accepted September 16, 2022.

DNA marker identification of downy mildew resistance locus *Rpv10* in grapevine genotypes

E.T. Ilnitskaya¹✉, M.V. Makarkina¹, S.V. Tokmakov¹, L.G. Naumova²

¹ North-Caucasian Federal Scientific Center of Horticulture, Viticulture, Winemaking, Krasnodar, Russia

² Ya.I. Potapenko All-Russian Research Institute of Viticulture and Winemaking – branch of Federal Rostov Agricultural Research Center, Novocherkassk, Russia
✉ ilnitskaya79@mail.ru

Abstract. One of the most common and harmful diseases of grapevine is downy mildew, caused by *Plasmopara viticola*. Cultivars of *Vitis vinifera*, the basis of high-quality viticulture, are mainly not resistant to downy mildew. Varieties with natural resistance to downy mildew belong to the vine species of North America and Asia (*V. aestivalis*, *V. berlandieri*, *V. cinerea*, *V. labrusca*, *V. amurensis*, etc.), as well as *Muscadinia rotundifolia*. The breeding of resistant cultivars is based on interspecific crossing. Currently, molecular genetic methods are increasingly used in pre-selection work and directly in breeding. One of the major loci of downy mildew resistance, *Rpv10*, was first identified in the variety Solaris and was originally inherited from wild *V. amurensis*. DNA markers that allow detecting *Rpv10* in grapevine genotypes are known. We used PCR analysis to search for donors of resistance locus among 30 grape cultivars that, according to their pedigrees, could carry *Rpv10*. The work was performed using an automatic genetic analyzer, which allows obtaining high-precision data. *Rpv10* locus allele, which determines resistance to the downy mildew pathogen, has been detected in 10 genotypes. Fingerprinting of grape cultivars with detected *Rpv10* was performed at 6 reference SSR loci. DNA marker analysis revealed the presence of a resistance allele in the cultivar Korinka russkaya, which, according to publicly available data, is the offspring of the cultivar Zarya Severa and cannot carry *Rpv10*. Using the microsatellite loci polymorphism analysis and the data from VIVC database, it was found that Korinka russkaya is the progeny of the cultivar Severnyi, which is the donor of the resistance locus *Rpv10*. The pedigree of the grapevine cultivar Korinka russkaya was also clarified.

Key words: *Vitis* sp.; target alleles; *Plasmopara viticola*; DNA fingerprinting.

For citation: Ilnitskaya E.T., Makarkina M.V., Tokmakov S.V., Naumova L.G. DNA marker identification of downy mildew resistance locus *Rpv10* in grapevine genotypes. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):129-134. DOI 10.18699/VJGB-23-18

ДНК-маркерная идентификация локуса устойчивости к милдью *Rpv10* в генотипах винограда

Е.Т. Ильницкая¹✉, М.В. Макаркина¹, С.В. Токмаков¹, Л.Г. Наумова²

¹ Северо-Кавказский федеральный научный центр садоводства, виноградарства, виноделия, Краснодар, Россия

² Всероссийский научно-исследовательский институт виноградарства и виноделия имени Я.И. Потапенко – филиал Федерального Ростовского аграрного научного центра, Новочеркасск, Россия

✉ ilnitskaya79@mail.ru

Аннотация. Милдью – одно из наиболее распространенных и вредоносных заболеваний виноградной лозы, возбудителем которого считают *Plasmopara viticola*. Сорты *Vitis vinifera*, выступая основой высококачественного виноградарства, практически не обладают генетической устойчивостью к милдью. Генотипы, имеющие природную устойчивость к поражению *P. viticola*, принадлежат видам винограда Северной Америки и Азии (*V. aestivalis*, *V. berlandieri*, *V. cinerea*, *V. labrusca*, *V. amurensis* и др.), а также к *Muscadinia rotundifolia*. По этой причине создание сортов винограда с повышенной устойчивостью к патогену основано на межвидовой гибридизации. В настоящее время молекулярно-генетические методы анализа все активнее используют на этапах предселекционной работы и непосредственно в селекции. Один из крупных локусов устойчивости к милдью – ген *Rpv10* – впервые идентифицирован в сорте межвидового происхождения Солярис и изначально происходит от дикого амурского винограда. Известны ДНК-маркеры данного гена, позволяющие детектировать наличие *Rpv10* в генотипах винограда. Методом ПЦР-анализа выполнен поиск доноров гена устойчивости среди генотипов 30 сортов винограда, которые, согласно родословным, могли бы нести ген *Rpv10*. Работа выполнена с использованием автоматического генетического анализатора, что позволяет получать высокоточные данные. По результатам ДНК-маркерного анализа в 10 генотипах винограда выявлено наличие аллели гена *Rpv10*, определяющей устойчивость к возбудителю милдью. Выполнено генотипирование сортов винограда, в которых обнаружен *Rpv10*, с помощью шести стандартных для ДНК-профилирования винограда SSR-маркеров. ДНК-маркерный анализ показал наличие аллели устойчивости у сорта Коринка русская, который, по общедоступным данным, является потомком сорта Заря Севера, не обладающим геном устойчивости *Rpv10*. С использованием анализа полиморфизма микросателлитных локусов и базы данных VIVC уточнена родословная сорта винограда Коринка русская. Установлено, что Коринка русская происходит от сорта Северный – донора локуса устойчивости *Rpv10*.

Ключевые слова: *Vitis* sp.; целевые аллели; *Plasmopara viticola*; ДНК-профилирование.

Introduction

The Eurasian grapevine (*Vitis vinifera* L.) is the most widely cultivated and economically important fruit crop in the world (De Mattia et al., 2008). Grapevines are grown both for direct food consumption and for the production of wine. The issue of creating pathogen-resistant genotypes is relevant in the breeding of table and wine cultivars. Downy mildew is one of the most common and harmful diseases of grapevine, caused by biotrophic oomycete *Plasmopara viticola* Berl. et de Toni. The pathogen has a narrow specialization and affects only grapevines: it develops on all green organs of the plant – leaves, shoots, inflorescences, berries, tendrils. The greatest damage is caused to vineyards in warm periods with high humidity. The creation of new grapevine forms is based on the use of the genetic diversity. The searching and identification of genotypes – donors of resistance, is an important task both for studying the diversity of the existing gene pool and for the purposes of breeding new resistant cultivars.

The *V. vinifera* genotypes, being the basis of high-quality viticulture, are mainly not resistant to *P. viticola*. The breeding of resistant cultivars is based on interspecific crossing. Genotypes with natural resistance to downy mildew belong to the vine species of North America (*V. riparia*, *V. aestivalis*, *V. berlandieri*, *V. cinerea*, *V. labrusca*) and East Asia (*V. amurensis*, *V. piasezkii*), as well as *Muscadinia rotundifolia* (Alleweldt, Possingham, 1988; Wan et al., 2007). It is generally accepted that resistance in American species developed simultaneously with the pathogen, which is endemic to North America. Resistance to *P. viticola* in some forms of *V. amurensis* could have developed through evolution from resistance to *P. cissii* and *P. amurensis*, these microorganisms are endemic to Asia (Riaz et al., 2011).

Molecular genetic analysis methods are successfully used now to identify and map loci of resistance to downy mildew. Both major loci with large influence in phenotypic variation and minor loci with smaller effects were identified (Bellin et al., 2009; Di Gaspero et al., 2012; Schwander et al., 2012; Venuti et al., 2013; Ochssner et al., 2016; Divilov et al., 2018; Lin et al., 2019; Sapkota et al., 2019; Bhattarai et al., 2020; Sargolzaei et al., 2020; Fu et al., 2020). The results of many such studies are successfully used for DNA marker selection to create quality grape cultivars with pyramided resistance genes (Eibach et al., 2007; Zini et al., 2019; Possamai et al., 2020; Ruiz-García et al., 2021).

Thus, a major locus of resistance inherited from wild *V. amurensis* was identified in the genotype of interspecific cultivar Solaris, it was named *Rpv10* (Schwander et al., 2012). The identified locus explained up to 50 % of observed phenotypic variance in the studied mapping hybrid population. Analysis of Solaris cultivar pedigree revealed that the allele that determines resistance to downy mildew was inherited from Severnyi (*V. amurensis* × Seyanets Malengra) cultivar. At the same time, studies have shown that in the genotype of Zarya Severa cultivar, which was selected from the same hybrid population as Severnyi (*V. amurensis* × Seyanets Malengra), the resistance allele is absent (Schwander et al., 2012). In the course of this study, flanking DNA markers of *Rpv10* locus were identified, which make it possible to search for genotypes – donors of the downy mildew resistance locus *Rpv10* in grapevine collections (Marker-Assisted Parental

Selection) and in the breeding process to identify hybrid samples carrying the target allele (Marker-Assisted Seedling Selection) according to DNA analysis data.

The aim of the work was to determine *Rpv10* locus in grape cultivar's genotypes using flanking DNA markers.

Material and methods

Grapevine accessions and DNA extraction. We included in the study grape cultivars that could have the resistance locus *Rpv10*, according to analysis of their well-known pedigree: cultivars-descendants of Severnyi cultivar or bred using wild *V. amurensis* (original gene donor). In total, 30 genotypes were analyzed: Amurets, Avgusta, Buytur, Cabernet severnyi, Cvetochnyi, Denisovskiy, Dimatskun, Druzhba, Dushystiy, Fioletovyi ranniy, Golubok, Grushevskiy belyi, Korinka russkaya, Kostyukovskiy, Kristall, Kunleany, Kurchanskiy, Lusakert, Morozko, Murometc, Muscat donskoi, Pamyati Dombkovskoy, Saperavi severnyi, Skromnyi, Stanichniy, Stepnyak, Sverkhanni volgodonskiy, Vostorg, Vydvizhenets, Zolotoy Don cultivars. Plant material was collected from the Anapa ampelographic collection (North-Caucasian Federal Scientific Center of Horticulture, Viticulture, Winemaking) and the collection of Ya.I. Potapenko All-Russian Research Institute of Viticulture and Winemaking – branch of Federal Rostov Agricultural Research Center. Genomic DNA samples were isolated from young tops of plant shoots. DNA extraction was carried out by the method based on the use of CTAB (Rogers, Bendich, 1985).

DNA analysis. Three DNA markers were used to identify the allelic status of *Rpv10* locus (GF09-44, GF09-46, GF09-47). The sequence of primer oligonucleotides was synthesized according to information from the literature (Schwander et al., 2012). Polymerase chain reaction (PCR) was carried out in total volume of 25 µl containing about 50 ng of genomic DNA, 1.5 units of Taq-polymerase (SibEnzyme, Russia), 1X Taq-polymerase buffer (SibEnzyme, Russia), 2 µM of MgCl₂ (SibEnzyme, Russia), 0.2 µM of each dNTP (SibEnzyme, Russia) and 200 µM of forward and reverse primers (Syntol, Russia). Amplification was carried out on a BioRad Thermo cycler T100 (USA). The following PCR conditions were used: initial denaturation for 5 min at 95 °C, 40 cycles of 30 s denaturation at 95 °C, annealing at 60 °C for 30 s and extension at 72 °C for 40 s, final step – 5 min extension at 72 °C. DNA of Solaris grape cultivar, which carries *Rpv10* resistance allele, was used as a control to identify target alleles and correct the size of the detected PCR fragments.

A standard set of SSR markers for DNA profiling of grapevine genotypes (VVS2, VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VrZAG62 and VrZAG79) was used for DNA fingerprinting of cultivars (This et al., 2004; This, 2007). Forward primers were labeled as follow: FAM (VVS2, VDMD27, VrZAG62), TAMRA (VVMD5, VVMD25, VVMD28, VVMD32), R6G (VVMD7, VrZAG79). The sequence of primer oligonucleotides was synthesized by Syntol (Russia). The following PCR conditions were applied: initial denaturation for 5 min at 95 °C; 34 cycles of 20 s denaturation at 95 °C, 30 s annealing at T_m (55 °C – VVS2, VVMD5, VVMD7, VVMD27; 58 °C – VrZAG62, VrZAG79; 60 °C – VVMD25, VVMD28, VVMD32) and 40 s extension at 72 °C; final extension of 3 min at 72 °C. To clarify the size

of the detected alleles, we used the DNA of reference cultivars Cabernet Sauvignon and Pinot noir. Fragment analysis was carried out using an ABI Prism 3130 genetic analyzer. Molecular genetic studies were carried out using the instrument park of the center for collective use of technological equipment of the North Caucasian Federal Scientific Center for Horticulture, Viticulture, Winemaking.

Results and discussion

Rpv10 locus detection

At the first stage of the work, 30 grapevine accessions were analyzed using the GF09-46 marker, this microsatellite locus was identified as a closely linked DNA marker, correlating with the presence of *Rpv10* locus, according to the studies of Schwander et al. (2012) (Schwander et al., 2012). The authors found that the PCR product of 416 base pairs size detected by the GF09-46 marker corresponds to the presence of *Rpv10* locus allele which determines downy mildew resistance in the grapevine genotype. The target fragment was identified in ten cultivars on the 30 analyzed accessions: Augusta, Golubok, Denisovskiy, Dimatskun, Korinka russkaya, Morozko, Saperavi severnyi, Stanichnyi, Fioletovyi ranniy, Cvetochnyi (Table 1). Some of the results were published earlier (Ilitskaya et al., 2019). At the second stage of the study, it was decided to analyze these ten cultivars with DNA markers GF09-44 and GF09-47, flanking the region of the chromosome where *Rpv10* locus is localized, which makes it possible to make sure that there is no crossing-over at this locus in the studied genotypes (Schwander et al., 2012).

Thus, according to the results of DNA marker analysis, target alleles at loci GF09-44 and GF09-47, correlating with the presence of a resistant allele in the *Rpv10* locus, according to published data, were detected in all ten samples (see Table 1).

It has been determined that there was no crossing-over at the analyzed part of the chromosome in the studied genotypes, thus, according to the DNA marker analysis, the presence of the downy mildew resistance locus *Rpv10* in grape cultivars

Augusta, Golubok, Denisovskiy, Dimatskun, Korinka russkaya, Morozko, Saperavi severnyi, Stanichnyi, Cvetochnyi and Fioletovyi ranniy is confirmed.

An analysis of the pedigree of these cultivars suggests that the locus is inherited directly from Severnyi cultivar (Saperavi severnyi, Denisovskiy, Golubok, Fioletovyi ranniy, Cvetochnyi) and from the descendants of this cultivar (Avgusta, Morozko, Stanichnyi) (Table 2).

In the genotype of Dimatskun, according to the origin of this cultivar, the *Rpv10* could be inherited from both the paternal and the maternal genotype, since the resistance donor is wild *V. amurensis*, which is present in the pedigrees of both parents of this cultivar. Of interest is the fact that *Rpv10* resistance locus is present in Korinka russkaya cultivar. So, the parents of this genotype are considered to be cultivars Zarya Severa and Kishmish chernyi (<http://www.vivc.de>). However, Zarya Severa genotype lacks *Rpv10* allele that determines resistance, according to the published data (Schwander et al., 2012) and our research. Thus, the reliability of information about Korinka russkaya cultivar pedigree is questionable.

According to the average long-term data of observations, the greatest field resistance to downy mildew among these cultivars is shown by Stanichnyi cultivar (5–25 % damage). Most likely, Stanichnyi genotype also contains downy mildew resistance genes inherited from North American grape species, this cultivar has a complex interspecific origin (see Table 2).

Fingerprinting

We carried out genotyping of Korinka russkaya and Zarya Severa by nine SSR loci used for DNA fingerprinting and identification of grapevine cultivars (This et al., 2004; This, 2007). The obtained data confirm the assumption that Zarya Severa cannot be the maternal parent of Korinka russkaya cultivar (Table 3).

If Korinka russkaya was bred from Zarya Severa cultivar, then, according to the codominant type of inheritance of SSR loci alleles, one of the alleles of Zarya Severa of each analyzed microsatellite loci would be found in the corresponding lo-

Table 1. The results of grape genotypes analysis with DNA markers linked to downy mildew resistance locus *Rpv10*

Cultivar	The sizes of identified alleles, base pair					
	GF09-44		GF09-46		GF09-47	
Avgusta	230	242	416	423	296	299
Golubok	230	–	416	–	296	299
Denisovskiy	230	243	395	416	296	299
Dimatskun	230	242	416	423	296	299
Korinka russkaya	230	242	416	423	296	299
Morozko	230	242	416	423	296	299
Saperavi severnyi	230	243	395	416	296	299
Stanichnyi	230	–	407	416	290	299
Fioletovyi ranniy	230	244	416	–	296	299
Cvetochnyi	230	236	394	416	296	299

Note. Target fragments that correlated with resistance are shown in bold.

Table 2. Pedigree of the analyzed grape genotypes

Cultivar	Pedigree	Originator (institution, country)
Avgusta	SV 12-309 × Kazachka (Kazachka-1 × Fioletovyi ranniy (Severnii × Muscat Gamburg))	Ya.I. Potapenko All-Russian Research Institute of Viticulture and Wine-making – branch of Federal Rostov Agricultural Research Center, Russia
Golubok	Severnii × a mix of pollen cultivars 40 let Oktyabrya, Odesskiy ranniy and No. 1-17-54 (Alicante Bouschet × Cabernet Sauvignon)	V.Ye. Tairov Institute of Viticulture and Winemaking of the National Academy of Agrarian Sciences of Ukraine, Ukraine
Denisovskiy	Severnii × pollen mix of muscat cultivars (Muscat a petits grains blancs, Muscat fleur d'oranger, Muscat of Alexandria)	Ya.I. Potapenko All-Russian Research Institute of Viticulture and Wine-making – branch of Federal Rostov Agricultural Research Center, Russia
Dimatskun	Karmrayut (Adisi × (<i>V. amurensis</i> × Cherniy sladkiy) × Seedling 1563/1 + 21 (Madeleine Angevine × <i>V. amurensis</i>) × Seyanets Malengra 65/16 (open pollination of Seedling Malengra cultivar))	Armenian Academy of Viticulture, Wine-Making and Fruit-Growing, Armenia
Korinka russkaya	Zarya Severa Severnii × Kishmish cherniy	I.V. Michurin Federal Scientific Center, Russia
Morozko	Mitsar × Saperavi severnii (Severnii × Saperavi)	North Caucasian Federal Scientific Center for Horticulture, Viticulture, Winemaking, Russia
Saperavi severnii	Severnii × Saperavi	Ya.I. Potapenko All-Russian Research Institute of Viticulture and Wine-making – branch of Federal Rostov Agricultural Research Center, Russia
Stanichnyi	Cvetochnyi (Severnii × pollen mix of muscat cultivars) × Zala Gyoengye	Ya.I. Potapenko All-Russian Research Institute of Viticulture and Wine-making – branch of Federal Rostov Agricultural Research Center, Russia
Fioletovyi ranniy	Severnii × Muscat Gamburg	Ya.I. Potapenko All-Russian Research Institute of Viticulture and Wine-making – branch of Federal Rostov Agricultural Research Center, Russia
Cvetochnyi	Severnii × pollen mix of muscat cultivars (Muscat a petits grains blancs, Muscat fleur d'oranger, Muscat of Alexandria)	Ya.I. Potapenko All-Russian Research Institute of Viticulture and Wine-making – branch of Federal Rostov Agricultural Research Center, Russia

Table 3. DNA profiles of grape cultivars Korinka russkaya, Zarya Severa and Severnii by nine SSRs

Cultivar	Alleles of SSR loci, base pairs								
	VVS2	VVMD5	VVMD7	VVMD25	VVMD27	VVMD28	VVMD32	VrZAG62	VrZAG79
Cabernet Sauvignon	139	234	239	239	176	234	240	188	247
	151	242	239	249	190	236	240	194	247
Zarya Severa	135	238	247	255	195	236	242	200	249
	139	240	249	271	195	244	272	204	251
Korinka russkaya	135	236	241	245	182	218	240	184	247
	155	238	253	255	184	244	250	188	255
Severnii (VIVC)	129	238	241	237	182	244	237	184	255
	135	238	247	255	184	252	240	204	259

Note. The alleles inherited by Korinka russkaya from Severnii genotype are shown in bold.

cus of Korinka russkaya cultivar. However, in five (VVMD7, VVMD27, VrZAG62, VrZAG79, VVMD32) out of nine studied SSR loci, these cultivars do not have common alleles (see Table 3).

Most likely, *Rpv10* locus in Korinka russkaya is inherited from Severnii cultivar, according to the analysis of the history of Korinka russkaya genotype origin. In addition, the information that Severnii cultivar is the parent of Korinka russkaya was found by us in a literary source describing the northern grape cultivars of Russia (Abuzov, 2009). Using data from the DNA profile database of Vitis International Variety Catalogue

(<http://www.vivc.de>), we performed the DNA profiles comparison between Korinka russkaya and Severnii. The allele from Severnii cultivar was identified in each analyzed locus of Korinka russkaya, accordingly (see Table 3). So Severnii is the parent of Korinka russkaya, Zarya Severa is not in the pedigree of Korinka russkaya.

We performed genotyping on VVS2, VVMD5, VVMD7, VVMD27, VrZAG62 and VrZAG79 SSR loci of cultivars, in which *Rpv10* resistance locus was identified (Table 4). The DNA profiles can then be used for the trueness-to-type analysis of accessions. Genotypes Avgusta, Golubok, Denisovskiy,

Table 4. DNA profiles of grape cultivars with detected *Rpv10* locus

Cultivar	Alleles of SSR loci, base pair					
	VVS2	VVMD5	VVMD7	VVMD27	VrZAG62	VrZAG79
Cabernet Sauvignon	139	234	239	176	188	247
	151	242	239	190	194	247
Pinot noir	137	230	239	186	188	239
	151	240	243	190	194	245
Avgusta	133	228	237	180	180	255
	133	238	249	190	186	261
Golubok	129	238	239	182	184	255
	133	248	241	184	188	259
Denisovskiy	135	236	245	184	202	243
	137	238	247	190	204	255
Dimatskun	133	238	241	182	188	243
	135	240	249	184	196	247
Morozko	129	226	247	182	202	247
	143	238	249	190	204	255
Saperavi severnyi	129	226	239	182	200	243
	133	238	247	192	204	255
Stanichnyi	135	238	233	184	188	255
	139	266	243	188	196	259
Fioletovyi ranniy	129	234	241	180	184	255
	149	238	249	182	186	259
Cvetochnyi	133	230	233	180	196	255
	135	238	247	184	204	255

Dimatskun, Korinka russkaya, Morozko, Saperavi severnyi, Stanichnyi, Cvetochnyi and Fioletovyi ranniy can be used in breeding as donors of *Rpv10*. Also, all these cultivars have increased frost resistance.

Conclusion

Using the DNA markers GF09-44, GF09-46 and GF09-47 linked to downy mildew resistance locus *Rpv10*, we analyzed 30 genotypes of grapes that could inherit this *R*-loci, according to their pedigrees. *Rpv10* locus was detected in the DNA of cultivars Avgusta, Golubok, Denisovskiy, Dimatskun, Korinka russkaya, Morozko, Saperavi severnyi, Stanichnyi, Cvetochnyi and Fioletovyi ranniy. All these cultivars were genetically characterized with the standard set of six SSRs for identification of grape cultivars. It was also shown by the results of SSR analysis of Korinka russkaya and Zarya Severa genotypes that cultivar Zarya Severa is not the parent of Korinka russkaya. The presence of *Rpv10* locus in Korinka russkaya genotype also confirms these data, since Zarya Severa does not carry *Rpv10*. Comparison of Korinka russkaya and Severnyi DNA profiles confirmed the assumption that Severnyi is the parent of Korinka russkaya cultivar. Thus, the pedigree of Korinka russkaya grape cultivar has been clarified.

References

Abuzov M. Atlas of Northern Grapes. Smolensk: KFH Pitomnik Publ., 2009. (in Russian)
 Alleweldt G., Possingham J.V. Progress in grapevine breeding. *Theor. Appl. Genet.* 1988;75:669-673. DOI 10.1007/BF00265585.

Bellin D., Peressotti E., Merdinoglu D., Wiedemann-Merdinoglu S., Adam-Blondon A.F., Cipriani G., Morgante M., Testolin R., Di Gaspero G. Resistance to *Plasmopara viticola* in grapevine ‘Bianca’ is controlled by a major dominant gene causing localised necrosis at the infection site. *Theor. Appl. Genet.* 2009;120:163-176. DOI 10.1007/s00122-009-1167-2.
 Bhattarai G., Fennell A., Londo J.P., Coleman C., Kovacs L.G. A novel grape downy mildew resistance locus from *Vitis rupestris*. *Am. J. Enol. Vitic.* 2020;2:12-20. DOI 10.5344/ajev.2020.20030.
 De Mattia F., Imazio S., Grassi F., Baneh H.D., Scienza A., Labra M. Study of genetic relationships between wild and domesticated grapevine distributed from middle east regions to European countries. *Rend. Lincei.* 2008;19:223-240. DOI 10.1007/s12210-008-0016-6.
 Di Gaspero G., Copetti D., Coleman C., Castellarin S.D., Eibach R., Kozma P., Lacombe T., Gambetta G., Zvyagin A., Cindrić P., Kovács L., Morgante M., Testolin R. Selective sweep at the *Rpv3* locus during grapevine breeding for downy mildew resistance. *Theor. Appl. Genet.* 2012;124(2):277-286. DOI 10.1007/s00122-011-1703-8.
 Divilov K., Barba P., Cadle-Davidson L., Reisch B.I. Single and multiple phenotype QTL analyses of downy mildew resistance in interspecific grapevines. *Theor. Appl. Genet.* 2018;131(5):1133-1143. DOI 10.1007/s00122-018-3065-y.
 Eibach R., Zyprian E., Welter L., Töpfer R. The use of molecular markers for pyramiding resistance genes in grapevine breeding. *Vitis.* 2007;46(3):120-124. DOI 10.5073/vitis.2007.46.120-124.
 Fu P., Wu W., Lai G., Li R., Peng Y., Yang B., Wang B., Yin L., Qu J., Song Sh., Lu J. Identifying *Plasmopara viticola* resistance Loci in grapevine (*Vitis amurensis*) via genotyping-by-sequencing-based QTL mapping. *Plant Physiol. Biochem.* 2020;154:75-84. DOI 10.1016/j.plaphy.2020.05.016.
 Ilnitskaya E., Tokmakov S., Makarkina M., Suprun I. Identification of downy mildew resistance genes *Rpv10* and *Rpv3* by DNA-marker

- analysis in a Russian grapevine germplasm collection (Conference Paper). *Acta Hort.* 2019;1248:129-134. DOI 10.17660/ActaHortic.2019.1248.19.
- Lin H., Leng H., Guo Y., Kondo S., Zhao Y., Shi G., Guo X. QTLs and candidate genes for downy mildew resistance conferred by interspecific grape (*V. vinifera* L. × *V. amurensis* Rupr.) crossing. *Sci. Hort.* 2019;244:200-207. DOI 10.1016/j.scienta.2018.09.045.
- Ochssner I., Hausmann L., Töpfer R. *Rpv14*, a new genetic source for *Plasmopara viticola* resistance conferred by *Vitis cinerea*. *Vitis*. 2016;55:79-81. DOI 10.5073/vitis.2016.55.79-81.
- Possamai T., Migliaro D., Gardiman M., Velasco R., De Nardi B. *Rpv* mediated defense responses in grapevine offspring resistant to *Plasmopara viticola*. *Plants*. 2020;9(6):781. DOI 10.3390/plants9060781.
- Riaz S., Tenschler A.C., Ramming D.W., Walker M.A. Using a limited mapping strategy to identify major QTLs for resistance to grapevine powdery mildew (*Erysiphe necator*) and their use in marker-assisted breeding. *Theor. Appl. Genet.* 2011;122:1059-1073. DOI 10.1007/s00122-010-1511-6.
- Rogers S.O., Bendich A.J. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 1985;19:69-76. DOI 10.1007/BF00020088.
- Ruiz-García L., Gago P., Martínez-Mora C., Santiago J.L., Fernández-López D.J., Martínez M.D.C., Boso S. Evaluation and pre-selection of new grapevine genotypes resistant to downy and powdery mildew, obtained by cross-breeding programs in Spain. *Front. Plant Sci.* 2021;12:674510. DOI 10.3389/fpls.2021.674510.
- Sapkota S., Chen L.L., Yang S., Hyma K.E., Cadle-Davidson L., Hwang C.F. Construction of a high-density linkage map and QTL detection of downy mildew resistance in *Vitis aestivalis*-derived "Norton". *Theor. Appl. Genet.* 2019;132:137-147. DOI 10.1007/s00122-018-3203-6.
- Sargolzaei M., Maddalena G., Bitsadze N., Maghradze D., Bianco P.A., Failla O., Toffolatti S.L., De Lorenzis G. *Rpv29*, *Rpv30* and *Rpv31*: three novel genomic loci associated with resistance to *Plasmopara viticola* in *Vitis vinifera*. *Front. Plant Sci.* 2020;11:1537. DOI 10.3389/fpls.2020.562432.
- Schwander F., Eibach R., Fechter I., Hausmann L., Zyprian E., Töpfer R. *Rpv10*: a new locus from the Asian *Vitis* gene pool for pyramiding downy mildew resistance loci in grapevine. *Theor. Appl. Genet.* 2012;124:163-176. DOI 10.1007/s00122-011-1695-4.
- This P., Jung A., Boccacci P., Borrego J., Botta R., Costantini L., Crespan M., Dangl G.S., Eisenheld C., Ferreira-Monteiro F., Grando S., Ibáñez J., Lacombe T., Laucou V., Magalhães R., Meredith C.P., Milani N., Peterlunger E., Regner F., Zulini L., Maul E. Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theor. Appl. Genet.* 2004;109:1448-1458. DOI 10.1007/s00122-004-1760-3.
- This P. Microsatellite markers analysis. In: Minutes of the First Grape Gen06 Work-shop March 22nd and 23rd, INRA, Versailles (France). 2007;3-42.
- Venuti S., Copetti D., Foria S., Falginella L., Hoffmann S., Bellin D., Cindrić P., Kozma P., Scalabrin S., Morgante M., Testolin R., Di Gaspero G. Historical introgression of the downy mildew resistance gene *Rpv12* from the Asian species *Vitis amurensis* into grapevine varieties. *PLoS One*. 2013;8:e61228. DOI 10.1371/journal.pone.0061228.
- Wan Y., Schwaninger H., He P., Wang Y. Comparison of resistance to powdery mildew and downy mildew in Chinese wild grapes. *Vitis*. 2007;46:132-136. DOI 10.5073/vitis.2007.46.132-136.
- Zini E., Dolzani C., Stefanini M., Gratl V., Bettinelli P., Nicolini D., Betta G., Dorigatti C., Velasco R., Letschka T., Vezzulli S. R-loci arrangement versus downy and powdery mildew resistance level: a *Vitis* hybrid survey. *Int. J. Mol. Sci.* 2019;20(14):3526. DOI 10.3390/ijms20143526.

ORCID ID

E.T. Ilnitskaya orcid.org/0000-0002-2446-0971
M.V. Makarkina orcid.org/0000-0002-3397-0666
S.V. Tokmakov orcid.org/0000-0002-2092-7757
L.G. Naumova orcid.org/0000-0002-5051-2616

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

Received June 17, 2022. Revised August 30, 2022. Accepted August 30, 2022.

Stability analysis for seed yield of chickpea (*Cicer arietinum* L.) genotypes by experimental and biological approaches

R. Karimizadeh¹, P. Pezeshkpour², A. Mirzaee³, M. Barzali⁴, P. Sharifi⁵✉, M.R. Safari Motlagh⁶

¹ Kohgiluyeh and Boyer-Ahmad Agricultural and Natural Resources Research and Education Center, Dryland Agricultural Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Gachsaran, Iran

² Lorestan Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Khorramabad, Iran

³ Ilam Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Ilam, Iran

⁴ Golestan Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Gonbad, Iran

⁵ Department of Agronomy and Plant Breeding, Rasht Branch, Islamic Azad University, Rasht, Iran

⁶ Department of Plant Protection, Rasht Branch, Islamic Azad University, Rasht, Iran

✉ Peyman.sharifi@gmail.com and sharifi@iaurasht.ac.ir

Abstract. A range of environmental factors restricts the production of chickpea; therefore, introducing compatible cultivars to a range of environments is an important goal in breeding programs. This research aims to find high-yielding and stable chickpea genotypes to rainfed condition. Fourteen advanced chickpea genotypes with two control cultivars were cultivated in a randomized complete block design in four regions of Iran during 2017–2020 growing seasons. The first two principal components of AMMI explained 84.6 and 10.0 % of genotype by environment interactions, respectively. Superior genotypes based on simultaneous selection index of ASV (ssiASV), ssiZA, ssiDi and ssiWAAS were G14, G5, G9 and G10; those based on ssiEV and ssiSIPC were G14, G5, G10 and G15 and those based on ssiMASD were G14, G5, G10 and G15. The AMMI1 biplot identified G5, G12, G10 and G9 as stable and high-yielding genotypes. Genotypes G6, G5, G10, G15, G14, G9 and G3 were the most stable genotypes in the AMMI2 biplot. Based on the harmonic mean and relative performance of genotypic values, G11, G14, G9 and G13 were the top four superior genotypes. Factorial regression indicated that rainfall is very important at the beginning and end of the growing seasons. Genotype G14, in many environments and all analytical and experimental approaches, has good performance and stability. Partial least squares regression identified genotype G5 as a suitable genotype for moisture and temperature stresses conditions. Therefore, G14 and G5 could be candidates for introduction of new cultivars.

Key words: AMMI; HMRPGV; factorial regression (FR); mixed models; partial least squares regression (PLSR); simultaneous selection index (ssi).

For citation: Karimizadeh R., Pezeshkpour P., Mirzaee A., Barzali M., Sharifi P., Safari Motlagh M.R. Stability analysis for seed yield of chickpea (*Cicer arietinum* L.) genotypes by experimental and biological approaches. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):135-145. DOI 10.18699/VJGB-23-19

Анализ стабильности урожайности семян генотипов нута (*Cicer arietinum* L.) с помощью экспериментальных и биологических подходов

Р. Каримизаде¹, П. Пезешкпур², А. Мирзаеи³, М. Барзали⁴, П. Шарифи⁵✉, М.Р. Сафари Мотлах⁶

¹ Центр исследований и образования в области сельского хозяйства и природных ресурсов Кохгилуе и Бойерахмад, Институт сельскохозяйственных исследований засушливых земель, Организация сельскохозяйственных исследований, образования и распространения знаний (AREEO), Гачсаран, Иран

² Лорестанский центр исследований и образования в области сельского хозяйства и природных ресурсов, Организация сельскохозяйственных исследований, образования и распространения знаний (AREEO), Хоррамабад, Иран

³ Иламский центр исследований и образования в области сельского хозяйства и природных ресурсов, Организация сельскохозяйственных исследований, образования и распространения знаний (AREEO), Илам, Иран

⁴ Голестанский центр исследований и образования в области сельского хозяйства и природных ресурсов, Организация сельскохозяйственных исследований, образования и распространения знаний (AREEO), Гонбад, Иран

⁵ Кафедра агрономии и растениеводства Раштского филиала Исламского университета Азад, Рашт, Иран

⁶ Кафедра защиты растений сельскохозяйственного факультета Раштского филиала Исламского университета Азад, Рашт, Иран

✉ Peyman.sharifi@gmail.com and sharifi@iaurasht.ac.ir

Аннотация. Ряд факторов окружающей среды ограничивают производство нута, поэтому исследование сортов в различных средах является важной составляющей селекционных программ. Цель представленного исследования заключалась в поиске высокоурожайных и устойчивых к условиям богарного земледелия генотипов нута. Четырнадцать перспективных генотипов нута с двумя контрольными сортами выращены в рандомизированном полном факторном эксперименте в четырех регионах Ирана в вегетационные периоды 2017–2020 гг. Первые два главных компонента объясняют 84.6 и 10.0 % генотипа взаимодействиями с окружающей средой (GEI)

соответственно. Лучшими генотипами на основе индексов одновременной селекции ASV (ssiASV), ssiZA, ssiDi и ssiWAAS стали G14, G5, G9 и G10, на основе ssiEV и ssiSIPC – G14, G5, G10 и G15, на основе ssiMASV – G14, G5, G10 и G15. Результаты AMMI1-биplot-анализа позволили идентифицировать G5, G12, G10 и G9 как стабильные и высокоурожайные генотипы. По данным модели AMMI2-биplot наиболее стабильными определены генотипы G6, G5, G10, G15, G14, G9 и G3. На основе гармонического среднего и относительной эффективности генотипических значений (HMRPGV) G11, G14, G9 и G13 отмечены как четыре лучших генотипа. Факторная регрессия показала, что количество осадков крайне важно в начале и конце вегетационного периода. Генотип G14 продемонстрировал хорошую урожайность и стабильность во многих различных условиях среды и при использовании всех аналитических и экспериментальных подходов. Методом частичной регрессии наименьших квадратов генотип G5 был идентифицирован как наиболее устойчивый к неблагоприятным условиям – как по влажности, так и по температуре. Следовательно, G14 и G5 могут быть кандидатами для интродукции новых сортов. Ключевые слова: AMMI; HMRPGV; факторная регрессия (FR); смешанные модели; частичная регрессия наименьших квадратов (PLSR); индекс одновременной селекции (ssi).

Introduction

Chickpea (*Cicer arietinum* L.) is one of the most important pulse crops that are well adapted to arid and semiarid conditions. Pulse crops are important sources of protein in human food and are suitable for animal feeds (Gaur et al., 2010). Chickpea is the fifth most rainfed crop in Iran and its harvested area is about 561,000 ha in Iran, which is mostly (98.59 %) cultivated in dryland areas (FAO, 2020)¹. Chickpea is a cool season legume and sensitive to heat stress (Devasirvatham et al., 2012) grown mainly in semi-arid and arid regions, where its production is restricted by a range of environmental factors such as high (or very low) temperature, lack (or excess) of soil moisture availability and day length (Richards M.F. et al., 2020).

Introducing compatible cultivars to a range of environments is the important goal in breeding programs (Karimizadeh, Mohammadi, 2010). Awareness of genotype by environment interaction (GEI) helps breeders to check genotypes more accurately and select the best genotypes. Because of exhibition of various phenotypic expressions of a specified genotype to different environments and unknown responses of some of the genotypes to a specified environment, investigation of GEI depends on the phenotypic stability and adaptation of genotypes (Yan et al., 2000). In other words, GEI created a hard situation for breeders and growers to choose high-yielding and stable varieties to different environments and decreased the efficiency in selection of superior genotypes and cultivar introduction (Yan, Kang, 2003). Because a stable variety is adapted to environmental variation, plant breeders are interested in the analysis of yield stability as a worthwhile characteristic of a genotype (Annicchiarico, 2002). Therefore, for evaluation of yield stability and performance, it is very important to use a variety to wide range of environments (Yan et al., 2011). Developing broadly adapted genotypes with a high level of phenotypic stability and yield potential is a tool to overcome the genotype by environment interaction (Kanouni et al., 2015). However, since it is difficult or impossible to find such a variety, specific adaptation of varieties permits plant breeders to manage GEI and develop suitable genotypes for different environments (Gauch, Zobel, 1997).

Statistical models, which incorporate environmental and genotypic variables into the multi-environmental trial (MET) analysis, have been used to study and explain GEI. Two main statistical methods for analyzing GEI are experimental (or

empirical) and analytical (or biological) approaches (van Eeuwijk et al., 1996). The empirical approaches focus on performance-based selection, whereas the analytical approaches refer to the integration of some agronomic/climatic variables that determine the response variable (such as grain yield) (Richards R.A., 1982). Factorial regression (FR) (van Eeuwijk et al., 1996) and partial least squares regression (PLSR) (Vargas et al., 1998), which directly incorporate environmental variables and/or external varieties, can be considered as a predictive strategy for recommendation purposes (Basford, Cooper, 1998).

There are several methods for stability analysis in experimental approaches, including multivariate and univariate models. Additive main effect and multiplicative interaction (AMMI) (Gauch, Zobel, 1988), as a multivariate model in experimental approach, is postdictive, because it has to handle the problem of repeatability of GEI (Basford, Cooper, 1998). All models attempt to provide a biological interpretation of GEI using information on external environmental and/or external genotypic variables. An alternative method of experimental approach for stability analysis is the harmonic mean, and the relative performance of genotypic values (HMRPGV) based on mixed models (Resende, 2007). This method provides information on stability, adaptability and yield performance of genotypes in the same unit and scale. In this method, selection of the genotypes with the highest values of harmonic mean of genotypic values (HMGV), relative performance of genotypic values (RPGV) and HMRPGV allows a simultaneous selection for yield performance and stability. This methodology is used in evaluation of stability of yield performance in rice (Colombari-Filho et al., 2013), wheat (Coan et al., 2018; Verma, Singh, 2020) and corn (Rodvalho et al., 2015). M.A. Rodvalho et al. (2015) compared HMRPGV, Lin and Binns's and Annicchiarico's methods for stability of maize hybrids and indicates high agreement between these methodologies, however, the HMRPGV method enables breeders to directly assess the breeding values for the yield, genotypic stability and adaptability simultaneously. The FR has been used successfully to interpret GEI in maize (Romay et al., 2010), wheat (Campbell et al., 2004; Voltas et al., 2005; Joshi et al., 2010), durum wheat (Mohammadi et al., 2020a, b) and barley (Ahakpaz et al., 2021). PLS regression to interpret the GE interaction has also been applied in wheat (Vargas et al., 1999; Kondić-Špika et al., 2019), maize (Stojaković et al., 2015), sorghum (Das et al., 2012) and barley (Hilmars-son et al., 2021). Although many researchers have evaluated

¹ FAO. Statistics of Food and Agriculture Organization. 2020. <https://www.fao.org/statistics/en/>

stability of chickpea genotypes by stability methods such as AMMI (Farshadfar et al., 2011, 2013; Zali et al., 2012; Funga et al., 2017; Pouresmael et al., 2018; Azam et al., 2020), there have been no reports of analytical approaches in the case of this crop.

This study was carried out to get high-yielding and adaptable genotypes to rainfed condition of Iran, to compare empirical methods and to assess the role of climatic factors in GEI.

Materials and methods

Experimental conditions and plant material

Fourteen advanced chickpea genotypes with two control varieties (Adel and Azad) (Table 1) were cultivated in randomized complete block design in four regions of Iran, including Gachsaran, Gonbad, Khorramabad and Ilam (Table 2), during 2017–2020 growing years. The experiment was performed in Gonbad in every three cropping years, in Gachsaran and Ilam in the first two cropping years and in Khorramabad only in the second cropping year. Fifty seeds per m² were grown in plots with six m length and one m wide. Chemical fertilizer at the rate of 100 kg ha⁻¹ of ammonium phosphate and 35 kg ha⁻¹ of urea was evenly mixed with the soil. After harvest, seed yield was weighed and statistical analyzes were performed on the data.

Statistical analysis

Experimental approaches. The AMMI model was used to analyze the genotype (G) × environment (E) interactions. AMMI constitutes a model family, with AMMI0 having no interaction principal component (IPC), AMMI1 having 1 IPC, AMMI2 having 2 IPC, and so on up to AMMIF (residual discarded). The AMMI model equation is:

$$Y_{ij} = \mu + \alpha_j + \beta_i + \sum_n \lambda_n \gamma_{in} \delta_{jn} + \rho_{ij}$$

where Y_{ij} is the yield of genotype i in environment j ; μ is the grand mean; α_i is the genotype deviation from the grand mean; β_j is the environment deviation; λ_n is the singular value for IPC n and correspondingly λ_n^2 is its eigenvalue; γ_{in} is the eigenvector value for genotype i and component n ; δ_{jn} is the eigenvector value for environment j and component n , with both eigenvectors scaled as unit vectors; and ρ_{ij} is the residual.

Simple and combined analysis of variance and stability analysis performed by METAN R packages (Olivoto, DalCol Lucio, 2020). The agricolae R package (Mendiburu, 2019) was also used for calculation of some of AMMI indices. Stability indices were calculated using the equations in Table 3.

The SSI1/SSI2 ratio in equation 1 is the weight assigned to the first interaction principal component (IPC1), which is the product of dividing the sum of squares of first IPC by the sum of squares of the second IPC. In equation 2,

Table 1. Code and name of studied chickpea genotypes

No.	Origin	Name/Pedigree
1	ICARDA	TDS-Maragheh90-92/Gn-PR-93-15/Gn-PR-94-8
2	ICARDA	TDS-Maragheh90-137/Gn-PR-93-18/Gn-PR-94-10
3	ICARDA	TDS-Maragheh90-150/Gn-PR-93-23/Gn-PR-94-14
4	ICARDA	TDS-Maragheh90-162/Gn-PR-93-27/Gn-PR-94-17
5	ICARDA	TDS-Maragheh90-239/Gn-PR-93-49/Gn-PR-94-35
6	ICARDA	TDS-Maragheh90-292/Gn-PR-93-66/Gn-PR-94-45
7	ICARDA	TDS-Maragheh90-300/Gn-PR-93-67/Gn-PR-94-46
8	ICARDA	TDS-Maragheh90-423/Gn-PR-93-97/Gn-PR-94-65
9	ICARDA	FLIP09-53C-X04TH175/FLIP95-51XFLIP97-165
10	ICARDA	FLIP09-178C-X06TH46/FLIP02-3XFLIP00-14
11	ICARDA	FLIP09-228C-S00794(30 KR)-2/
12	ICARDA	FLIP09-249C-S00794(30 KR)-6/
13	ICARDA	FLIP09-441C-X04TH61/X03TH-129XFLIP96-154
14	ICARDA	FLIP09-350C-X06TH44/FLIP00-50XFLIP01-60
15	IRAN	ADEL
16	IRAN	AZAD

Table 2. Geographic characteristics of trials area

Location	Altitude, m	Longitude	Latitude	Average rainfall, mm
Gachsaran	710	50° 50' E	30° 17' N	455
Gonbad	45	55° 12' E	37° 16' N	548
Ilam	975	46° 36' E	33° 47' N	362
Khorramabad	1147	48° 18' E	33° 29' N	445

Table 3. Equations for calculation the stability analysis indices

No.	Index	Formula	Reference
1	AMMI stability value	$ASV = \sqrt{\left[\frac{SSIPC1}{SSIPC2} (IPC1) \right]^2 + (IPC2)^2}$	Purchase et al., 2000
2	Sum of IPCs scores	$SIPC_i = \sum_{n=1}^N \lambda_n^{0.5} \gamma_{in}$	Sneller et al., 1997
3	Eigenvalue stability parameter of AMMI	$EV_i = \sum_{n=1}^N \gamma_{in}^2 / N'$	Zobel et al., 1988
4	Absolute value of the relative contribution of IPCs to the interaction	$Z\alpha_i = \sum_{n=1}^N \theta_n \gamma_{in} $	Zali et al., 2012
5	Modified AMMI stability value	$MASV = \sqrt{\sum_{n=1}^{N-1} \left[\frac{SSIPC_n}{SSIPC_{n+1}} (IPC_n) \right]^2 + (IPC_{n+1})^2}$	Adugna, Labuschange, 2002
6	Distance coefficient	$D_i = \sqrt{\sum_{n=1}^N \gamma_{in}^2}$	Zhang et al., 1998
7	Weighted average of absolute scores	$WAAS_i = \frac{\sum_{k=1}^p IPCA_{ik} \times \theta_k }{\sum_{k=1}^p \theta_k}$	Olivoto et al., 2019
8	Simultaneous selection index	$SSI = R$ (AMMI stability indices) + RY	Farshadfar, 2008
9	Harmonic mean of genotypic values	$HMGV_i = \frac{l}{\sum_{j=1}^l \frac{1}{GV_{ij}}}$	Resende, 2007
10	Relative performance of genotypic values	$RPGV_i = \frac{1}{l} \left[\sum_{j=1}^l \frac{GV_{ij}}{\mu_j} \right]$	Resende, 2007
11	Harmonic mean of relative performance of genotypic values	$HMRPGV_i = \frac{l}{\sum_{j=1}^l \frac{1}{GV_{ij} / \mu_j}}$	Resende, 2007

λ_n is the root of the n^{th} IPC, which for SIPC1 and SIPC2 is one and the number of principal components remaining in the model, respectively. In equations 3 and 4, γ_{in} is the root of the n^{th} axis and N' is the number of significant principal components in the analysis of variance of AMMI by F -test. In equation 4, the percentage of the sum of squares explained by the n^{th} axis of IPC denotes by θ_n . In equation 5, SSIPC1, SSIPC2, ..., SSIPC $_n$ are the sum of squares of the 1st, 2nd, ..., and n^{th} IPC; and PC $_1$, PC $_2$, ..., PC $_n$ are the scores of 1st, 2nd, ..., and n^{th} IPC. In equation 6, the AMMI distance (D) calculated as distance of the interaction principal component from the origin. In equation 7, IPC $_{ik}$ is the score of i^{th} genotype on k^{th} IPC axis. θ_k is the explained variance of the k^{th} IPCA for $k = 1, 2, \dots, p$, considering p the number of significant PCAs. In these equations, the most stable genotypes have the lowest values of stability indices.

In equation 8, the simultaneous selection index (ssi) is the sum of the rankings of genotypes based on the AMMI [R (AMMI stability indices)] and the average rank of seed yield of genotypes in all environments (RY). AMMI1 (IPC1 vs. seed yield) and AMMI2 (IPC1 vs. IPC2) biplots were drawn using the standard method described by R.W. Zobel et al. (1988).

The BLUP model for MET trials, unlike the classical additive model, assumes the genotypic effects as random and uses a different computational procedure (Olivoto et al., 2019).

In BLUP, μ_j is the general mean for j^{th} environment; l is the number of environments; GV_{ij} : $u_j + g_i + ge_{ij}$ is the genotypic value of i^{th} genotype in j^{th} environment. u_j is the mean of the j^{th} environment, and g_i and ge_{ij} are the BLUP values of i^{th} genotype and the interaction between i^{th} genotype and j^{th} environment, respectively. Stability indices based on this mixed model are: HMGV, RPGV and HMRPGV were calculated by Equations 9–11, respectively (Table 3).

Analytical approaches. Seasonal rainfall and average temperature of autumn, winter and spring were used as environmental co-variables. Integration of external data into GEI analysis by PLSR and FR methods was carried out by GEA-R software (Pacheco et al., 2015).

Partial least squares regression

The PLSR model includes independent matrices X (rainfall and average temperature data) and a dependent matrix Y (seed yield) and the latent variables t as follows (Vargas et al., 1998):

$$X = t1p1' + t1p1' + \dots + E = TP' + E$$

$$Y = t1q1' + t1q1' + \dots + F = TQ' + F,$$

where, matrices T , P and Q contain X -scores, X -loadings and Y -loadings, respectively. F and E are the residuals of the unexplained variation. A biplot was built based on the first two PLSR factors to investigate the relationships among co-variables, genotypes and environments.

Factorial regression

The FR model is also as follows (van Eeuwijk et al., 1996):

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j + \sum_{k=1}^K \xi_{ik} Z_{jk},$$

where α_i represents the genotype main effect; Z_{jk} refers to the value of any environmental variable k for environment j ; and ξ_{ik} represents the sensitivity of genotype i to the explicit environmental variable k . The heterogeneity in the ξ_i 's for successive $z_1 \dots z_K$ variables accounts for the interaction, while the sum of multiplicative terms $\sum_{k=1}^K \xi_{ik} Z_{jk}$ approximates GE.

To facilitate the interpretation of genotype by environment, the external variables can be centered to mean zero. The parameter ξ_{ik} can be easily estimated by standard least squares techniques.

The Akaike's information criterion (AIC) (Akaike, 1974) was used to determine the number of covariates that are included in the model.

Results

Analysis of variance

Analysis of variance showed that the effects of environment, genotype and genotype by environment interaction were significant on seed yield at 1 % probability level and these three components explained 37.13, 16.90 and 31.30 % of phenotypic variation, respectively (Table 4).

Due to significance of GEI, it is possible to analyze the stability of these data. Therefore, the stability analysis was performed by the AMMI method and harmonic mean, and HMRPGV based on mixed models. AMMI analysis of variance showed only the first two principal components were significant and explained 84.6 and 10.0 % of genotype by environment interaction, respectively (see Table 4).

Experimental stability approaches

AMMI stability indices and ssi. The highest seed yield was observed in G11, followed by G14, G9, G16 and G13, which were higher than average yield of genotypes in all environments (1069.25 kg ha⁻¹). Stability of genotypes was evaluated across different environments by AMMI indices. The ASV, WAAS, Za and MASV stability indices identified genotypes G5, G14, G12, G1 and G10 as the most stable genotypes. The SIPC and EV indices indicated genotypes G14, G5, G6, G10 and G15 were the most stable genotypes. According to D index, genotypes G14, G5, G10, G6 and G1 were more stable than other genotypes (Table 5).

Based on ssi of AMMI stability value (ssiASV), ssiZA, ssiDi and ssiWAAS, genotypes G14, G5, G9, G10 and G12 were identified as superior genotypes; while based on ssiEV and ssiSIPC, genotypes G14, G5, G10, G15 and G9 were the superior genotypes. The ssiMASV index identified genotypes G14, G5, G10, G15 and G12 as superior genotypes (Table 6).

Biplot interpretation. The AMMI1 biplot indicated the score of the first principal component in genotypes G5, G12, G10, G9, and G14 was near zero and so these genotypes had low interaction with environment and were identified as stable genotypes. The yield of these genotypes was also higher than the average seed yield of all genotypes in all environments

(1069.25 kg ha⁻¹). Genotypes G11, G8, G16, and G4 at the farthest point from the biplot origin were unstable genotypes (Fig. 1).

The AMMI2 biplot showed that genotypes G4, G2, G1, G12, G13, G11, G16, G7 and G8 with the longest distance from the biplot origin had a high contribution in genotype by environment interaction and were unstable genotypes, but these genotypes adapted to their close environments (Fig. 2). Therefore, genotype G2 was the best genotype for E1; genotype G1 for E4 and E5; genotypes G12 and G13 for E3; genotype G11 for E7 and genotypes G7 and G8 for E2, E6 and E8. The other genotypes within polygon, including G6, G5, G10, G15, G14, G9 and G3, were the most stable genotypes. Other usefulness of this biplot, in addition to identifying adaptable genotypes with any environment and introducing genotypes with general stability, are identification of environments with the long vector that could be more effective in finding stable genotypes (Yan, Kang, 2003). Accordingly, all environments except for E3 could be used as discriminative and representative environments.

Determination of genotypic stability and adaptability using HMRPGV. The top four superior genotypes compared to control varieties (ADEL and AZAD), based on the measure of stability and adaptability (HMRPGV), were genotypes G11, G14, G9 and G13. The products of this index and the general mean (HMRPGV* $\bar{\mu}$) of these genotypes were 1570, 1287, 1231 and 1200 kg ha⁻¹, respectively (Table 7). The selection of these genotypes for seed yield increased to 20.63 % over the general mean (1069.25 kg ha⁻¹).

Analytical stability approaches

FR analysis. Since the climatic information of the third year in Gonbad was not available, analytical approaches (factorial and partial least squares regression) with environmental co-

Table 4. AMMI analysis of variance for seed yield of chickpea genotypes

SOV	Df	MS	Percent	Accumulate
Env	7	6153924**	37.13	
Rep(Env)	16	120111		
Gen	15	1307022**	16.90	
Env*Gen	105	345820**	31.3	
PC1	21	1463178**	84.6	84.6
PC2	19	190452**	10	94.6
PC3	17	69654	3.3	97.8
PC4	15	25555	1.1	98.9
PC5	13	20294	0.7	99.6
PC6	11	8973	0.3	99.9
PC7	9	3976	0.1	100
Residuals	240	62892	.	.
Total	383	302898	.	.

** $p < 0.01$.

Table 5. AMMI-based stability indices and ranks for stability indices

Gen	Seed yield, kg ha ⁻¹	rY	ASV	rASV	SIPC	rSIPC	EV	rEV	Za	rZa	WAAS	rWAAS	Di	rD	MASV	rMASV
G1	792	15	78.9	4	17.9	6	0.047	7	0.164	5	9.18	4	12.6	5	596	5
G2	814	13	113.0	9	21.5	10	0.058	9	0.223	10	12.70	9	15.6	9	797	10
G3	841	12	117.0	10	18.0	7	0.038	6	0.219	9	12.80	10	14.5	8	794	9
G4	812	14	162.0	14	19.6	9	0.057	8	0.287	13	17.20	14	19.1	12	1082	13
G5	1100	8	15.6	1	9.0	2	0.025	2	0.046	1	2.23	1	7.52	2	260	2
G6	789	16	104.0	7	15.1	3	0.027	3	0.192	6	11.30	7	12.6	4	700	6
G7	905	11	143.0	12	26.8	13	0.089	13	0.282	12	16.10	12	19.6	13	1008	12
G8	1069	10	166.0	15	29.0	14	0.100	14	0.320	15	18.40	15	21.7	14	1146	15
G9	1272	3	92.1	6	22.1	11	0.076	11	0.195	7	10.80	6	15.6	10	715	7
G10	1116	7	81.6	5	16.4	4	0.036	5	0.164	4	9.29	5	11.8	3	587	3
G11	1616	1	202.0	16	30.3	15	0.108	15	0.375	16	22.00	16	24.7	16	1363	16
G12	1098	9	70.7	3	19.3	8	0.067	10	0.157	3	8.53	3	13.8	7	594	4
G13	1221	5	132.0	11	25.2	12	0.080	12	0.261	11	14.90	11	18.3	11	933	11
G14	1280	2	35.7	2	6.6	1	0.005	1	0.070	2	4.01	2	4.83	1	250	1
G15	1122	6	113.0	8	17.1	5	0.034	4	0.210	8	12.30	8	13.8	6	760	8
G16	1261	4	146.0	13	32.3	16	0.151	16	0.302	14	16.90	13	22.9	15	1091	14

Note. ASV, AMMI stability value; SIPC, sum of IPCs scores; EV, eigenvalue stability parameter of AMMI; Za, absolute value of the relative contribution of IPCs to the interaction; WAAS, weighted average of absolute scores.

Table 6. Simultaneous selection indices (ssi) for each of genotypes

Gen	ssiASV	ssiSIPC	ssiEV	ssiZa	ssiWAAS
G1	19	21	22	20	19
G2	22	23	22	23	22
G3	22	19	18	21	22
G4	28	23	22	27	28
G5	9	10	10	9	9
G6	23	19	19	22	23
G7	23	24	24	23	23
G8	25	24	24	25	25
G9	9	14	14	10	9
G10	12	11	12	11	12
G11	17	16	16	17	17
G12	12	17	19	12	12
G13	16	17	17	16	16
G14	4	3	3	4	4
G15	14	11	10	14	14
G16	17	20	20	18	17

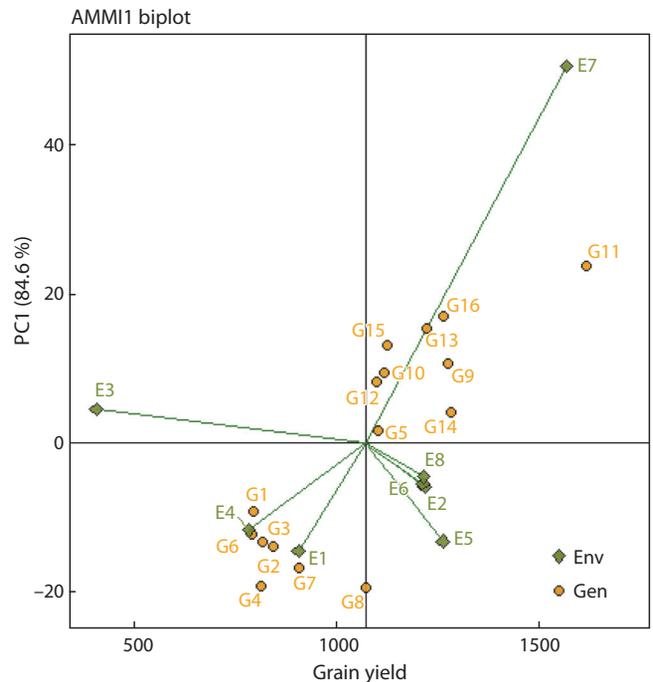


Fig. 1. AMMI1 biplot for identity of the superior lentil genotypes based on seed yield mean and PC1.

The naming of genotypes is similar to Table 9. E1, Gachsaran 2017-18; E2, Gonbad 2017-18; E3, Ilam 2017-18; E4, Gachsaran 2018-19; E5, Khorramabad 2018-19; E6, Gonbad 2018-19; E7, Ilam 2018-19; E8, Gonbad 2019-20.

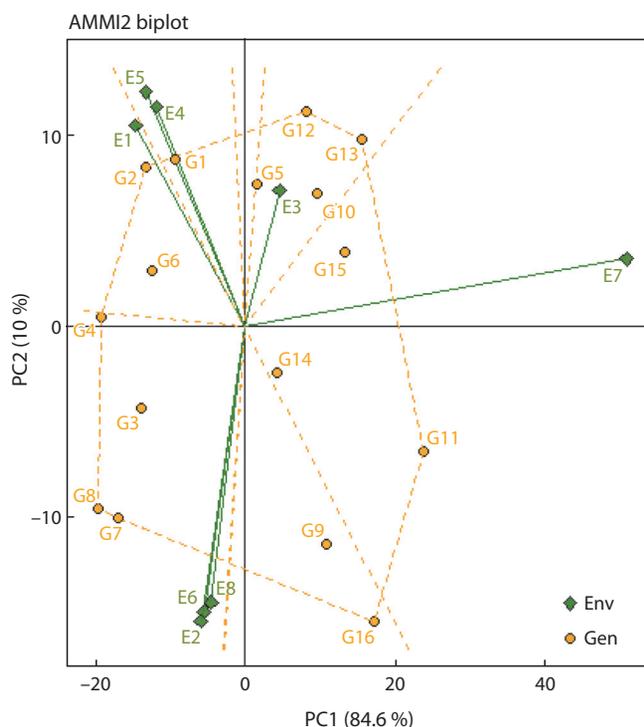


Fig. 2. The AMMI2 biplot for identity of the superior chickpea genotypes based on first PCs.

The naming of genotypes and environments is similar to Table 9 and Fig. 1, respectively.

variables were performed with seven environments. The steps for climatic variables in the FR model based on AIC indicated average temperature in autumn (FallT), average rainfall in spring (SpringR), average rainfall in autumn (FallR) and average temperature in spring (SpringT) were detected to be important contributors to GE interaction (Table 8).

Partial least squares regression analysis. The first and second PLSR factors based on environmental co-variables accounted for 73.12 and 9.16 % of the GE interaction sum of squares, respectively (Fig. 3). Environments located on the right-hand side of the biplot (E1, E2, E4 and E6) had high values for temperature co-variables (i. e., FallT, SpringT and WinterT), whereas the other environments (E3, E5 and E7) on the opposite side tended to have high rainfall (Table 9). These results indicate that some genotypes (G4, G11, G9, G14, G12, G8, G10, G15 and G13) performed better in high rainfall in winter and autumn seasons (see Fig. 3). The PLSR biplot displayed that high rainfall in the environments in the west of Iran (E3, Ilam 2017-18; E5, Khorramabad 2018-19 and E7, Ilam 2018-19) led to high performance in genotypes.

Discussion

The significant effect of genotype (16.90 %) and environment (37.13 %) is a sign of the comprehensive genetic background of experimental materials and diversity of experimental locations and cropping seasons. The significant effect of GEI shows different performance of genotypes in different environments. Other researchers also reported a greater contribution

Table 7. Ranking of the genotypes in all environments evaluated for adaptability parameters of genotypic values for the grain yield of chickpea genotypes evaluated in eight environments

Gen	Seed yield, kg ha ⁻¹	HMGV	HMGV_order	RPGV	RPGV_Y	RPGV_order	HMRPGV	HMRPGV_Y
G1	792	577	11	0.724	774	15	0.688	735
G2	814	568	12	0.755	807	13	0.688	736
G3	841	558	13	0.771	824	12	0.691	739
G4	812	418	16	0.741	792	14	0.564	603
G5	1100	970	8	1.05	1123	9	1.03	1106
G6	789	467	15	0.71	759	16	0.616	659
G7	905	535	14	0.83	888	11	0.689	737
G8	1069	705	10	0.986	1055	10	0.855	914
G9	1272	1070	4	1.19	1270	3	1.15	1231
G10	1116	991	6	1.06	1138	6	1.04	1111
G11	1616	1406	1	1.54	1647	1	1.47	1570
G12	1098	985	7	1.05	1126	8	1.03	1096
G13	1221	1076	3	1.16	1244	5	1.12	1200
G14	1280	1122	2	1.21	1292	2	1.2	1287
G15	1122	964	9	1.06	1129	7	1.02	1092
G16	1261	1059	5	1.19	1269	4	1.12	1193
Average	1069.25							

Note. RPGV, performance genetic value; HMGV, harmonic mean of genotypic values; HMRPGV, harmonic mean and relative performance of genotypic values.

Table 8. Stepwise factorial regression model for climatic variables based on Akaike's information criterion (AIC)

Effect name	Sum Sq	Df	F-value	AIC	Pr > F	TorF
Gen* FallT	8084394	15	3.745061	4992.872	4.81E-06	Effect entered
Gen* SpringR	8993428	15	5.055404	4939.702	8.60E-09	Effect entered
Gen* FallR	13256411	15	12.00947	4790.141	2.79E-22	Effect entered
Gen* SpringT	3053140	15	3.109111	4760.475	0.00012	Effect entered

Note. FallT, average temperature in autumn; SpringT, average temperature in spring; SpringR, average rainfall in spring; FallR, average rainfall in autumn.

Table 9. Seasonal rainfall and temperature in seven environments

Env	FallR	WinterR	SpringR	FallT	WinterT	SpringT
Gachsaran 1 (E1)	17.47	21.3	20.17	19.9	14.13	25.87
Gonbad 1 (E2)	27.03	59.17	28.5	19.73	12.7	22.73
Ilam 1 (E3)	18.37	82.47	64.27	16.03	9.7	19.83
Gachsaran 2 (E4)	110.9	85.33	58.57	19.53	11.7	24.53
Khorramabad 2 (E5)	99.63	113.03	105.03	14.87	6.67	18.83
Gonbad 2 (E6)	42.37	133.57	33	19.7	12.87	22.43
Ilam 2 (E7)	127.13	109.43	63.73	15.7	9.1	19.83
Average	63.27	86.33	53.32	17.92	10.98	22.01

Note. The naming of environments is similar to Fig. 1.

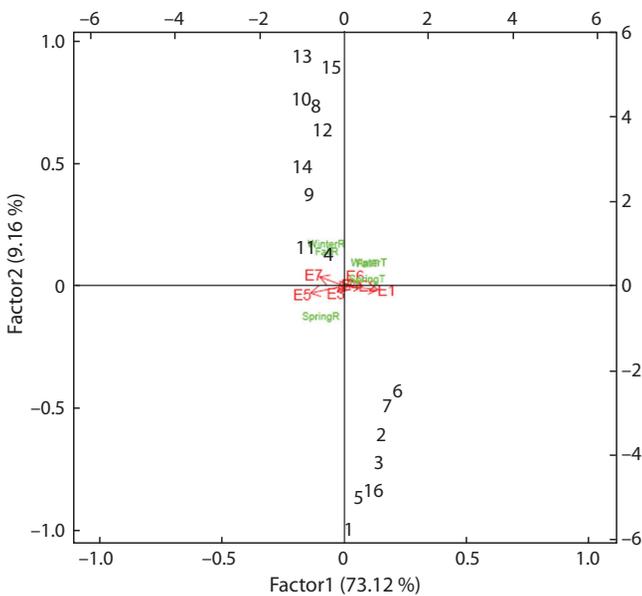


Fig. 3. The biplot based on PLSR method with rainfall seasons' covariates for seed yield of 16 chickpea genotypes in seven environments.

FallT, average temperature in autumn; winterT, average temperature in winter; SpringT, average temperature in spring; SpringR, average rainfall in spring; winterR, average rainfall in winter; FallR, average rainfall in autumn. The naming of genotypes and environments is similar to Table 9 and Fig. 1, respectively.

of environmental effect on total variation of chickpea seed yield (Farshadfar et al., 2011; Sayar, 2017; Pouresmael et al., 2018; Azam et al., 2020). Therefore, due to dependence of diversity of seed yield of chickpea genotypes on environment and genotype by environment interaction, further analysis needs to increase the selection efficiency of genotypes. In other words, the presence of significant GEI indicates the necessity

to find the yield potential and adaptability of genotypes based on evaluations at several locations and in cropping seasons (Annicchiarico, 2002). Since the genotype by environment interaction can reduce any improvement due to selection, therefore, in selection of cultivars, combination of stability with seed yield can lead to good results (Yan et al., 2001).

AMMI analysis of variance showed a high contribution of the first two principal components, especially PC1 (84.6 %) in GEI. Other researchers also indicated contribution of 52.5 and 21.95 % (Tilahun et al., 2015), 40.5 and 27.5 % (Farshadfar et al., 2013), 56.0 and 24.0 % (Farshadfar et al., 2011), 53.34 and 33.25 % (Azam et al., 2020) and 32.7 and 20.4 % (Funga et al., 2017) of the first two principal components in GEI of chickpea seed yield. In accordance with the results of present study, the other researchers were also identified stable chickpea genotypes by AMMI stability indices (Farshadfar et al., 2011, 2013; Zali et al., 2012; Funga et al., 2017; Pouresmael et al., 2018). Since the first two principal components had a high contribution on genotype by environment interaction, the stability indices including ASV, WAAS, Za and MASV had similar results in identifying the stable genotypes.

Identification of superior genotypes with AMMI indices was only based on genotype stability; so, genotypes G1 and G6 with a lower yield than average seed yield were identified as stable genotypes. Hence, ssi (Farshadfar, 2008) based on AMMI indices was used to find the superior genotypes. Since both aspects of stability and yield of a genotype were used in simultaneous selection index, the use of these indices prevents selection of stable genotypes with a low yield (Farshadfar, 2008). In accordance with these findings, A. Funga et al. (2017) also used ssi for yield performance and stability in chickpea to find stable and high-yielding genotypes. Use of simultaneous selection index for yield performance and stability can perform selection process with more confidence (Moghadam, 2003).

Based on the AMMI1 biplot, G5, G12, G10, G9, and G14 were the stable genotypes. Because the AMMI1 biplot uses both aspects of stability based on the first principal component and seed yield to identify stable genotypes, when the contribution of the first principal component in GEI is high (84.6 %), the results of the AMMI 1 biplot are very similar to the results of ssi based on AMMI indices. H.G. Gauch and R.W. Zobel (1988) stated that despite the high value of environment main effect, for evaluation of genotypes, only the effects of genotype (G) and GEI are appropriate and so it is necessary to remove the environment mean effect (E) and concentrate on G and GE.

The AMMI2 biplot identified G6, G5, G10, G15, G14, G9 and G3 as the most stable genotypes. This view of the biplot was also used for identifying the adapted genotypes to any of environments, so that the genotype placed at the top of each section is the best genotype for the environments in that section (Yan et al., 2000). Genotypes G7 and G8 were compatible with three environments E2, E6 and E8; they can be considered as genotypes with specific adaptability to these environments. Identification of environments with the long vector could be more effective in finding stable genotypes (Yan, Kang, 2003). The discrimination and representatives of all of the environments except E3 must be ascribed to the amount of rainfall and its proper distribution in different seasons. In agreement with the present finding, other researchers have identified stable chickpea genotypes using the AMMI2 biplot (Pouresmael et al., 2018; Funga et al., 2017; Farshadfar et al., 2013). Another remarkable point is that when the contribution of the first principal component is very high, identification of stable and high-yielding genotypes based on the AMMI1 biplot is better than on the AMMI2 biplot, so that G12, which was unstable in the AMMI2 biplot, in terms of ssiASV, ssiZA, ssiDi and ssiWAAS and AMMI1 was found to be the superior genotype.

The top four superior genotypes compared to control varieties based on HMRPGV were G11, G14, G9 and G13. In HMRPGV, the predicted genotypic values are declared as a proportion of the overall mean for each environment and then the mean value of this ratio is obtained across the environments (Rodvalho et al., 2015). The selection of genotypes in this method is based on stability, adaptability and yield performance; therefore, this method indicated a positive response of genotypes to environmental improvements and the stability of genotypes over the environments. M.D.V. Resende (2007) declared the HMRPGV method evaluated simultaneously seed yield, adaptability and stability, in a genotypic context. In this stability index, the genotypes can be simultaneously sorted by genotypic values and stability using the harmonic means of the BLUP (Rodvalho et al., 2015).

The analytical approach to analyzing GE interaction is important to enhance the value of MET and gain an understanding of the causes of GE interaction. These approaches have been demonstrated successfully in a range of crop species (van Eeuwijk et al., 1996; Mohammadi et al., 2020a, b). Factorial regression indicated rainfall to be very important at the beginning of the season to germination and establishing of seedlings and at the end of the season for its proper developmental and reproductive growth stages. In confirmation of this result, S.H. Sabaghpour et al. (2006) stated that chickpea needs the most water during flowering, podding and seed

filling and so, due to the lack of rainfall during these stages, terminal drought stress is a major abiotic stress for reducing chickpea productivity.

The rainfall was relatively high in environments E3, E5 and E7 that favored the positive GE interaction with G13, G15, G10, G8, G12, G14, G9 and G4. The best genotypes based on experimental methods (G11, G14, G9, G13, G5, G10, G15 and G12) were in the upper left quarter of the PLSR biplot (see Fig. 3). The seasonal rainfall of autumn and winter in environments E3, E5 and E7 in this quarter of the biplot, especially the last two environments, were higher than the average seasonal rainfall of all environments. The average seasonal temperature was also lower than the average temperature of all environments in these three environments. Thus, these environments can be considered as favorable environments in terms of these two climatic co-variables and the mentioned genotypes (G11, G14, G9, G13, G5, G10, G15 and G12) can be identified as superior genotypes in favorable conditions. The AMMI2 biplot also identified genotype G11 as a desirable genotype for environment E7.

On the other hand, the seasonal rainfall in environment E1 was much lower than the average seasonal rainfall in all environments, and its temperature was higher than the seasonal temperature in all environments. Therefore, this trial environment can be considered as an environment with drought and heat stresses for chickpea, which is a cold-loving crop. The PLSR biplot also demonstrates this hypothesis well, because the seasonal rainfall was on the opposite side of this environment and the average seasonal temperature was on its same side. Hence, genotypes located in the quadrant of this environment (right and bottom) can be considered genotypes tolerant to drought and heat stresses. The AMMI2 biplot also identified genotypes G2 and G6 as suitable for this environment. This environment had a high discriminating power due to the vector length in the AMMI2 biplot, so its results can be trusted and these results can be used properly. In the PLSR biplot, genotypes G6, G7, G2, G3, G16, G5 and G1 were in the same quarter along with environment E1. From these genotypes, G5 and G16 had a higher performance than the average yield of all genotypes and can be considered as tolerant genotypes to heat and drought stresses. Since genotype G16 was previously introduced as a cultivar, genotype G5 can be recommended as a suitable genotype for dryland and hot conditions. Such a conclusion is possible only from a combination of analytical and experimental approaches. If such analyzes were not performed here, we would not be able to achieve such results. It is happening on moisture stress towards the end of the cropping season with frequent events of heat stress in chickpea. Thus, the crop is exposed to stress conditions during the reproductive stage causing yield losses (Devasirvatham, Tan, 2018). A decrease in chickpea yields was observed with a 1 °C increase in seasonal temperature (Kalra et al., 2008; Upadhaya et al., 2011). Similarly, with every 0.1 °C temperature rise combined with 31 % reduction in seasonal rainfall, the yield of chickpea decreased (Dubey et al., 2011). This shows that high temperature and drought are the major factors that affect chickpea production. M.D. Kadiyala et al. (2016) have stated that unpredictable climate change is the main restriction for chickpea production as it increases the frequency of drought and temperature extremes, i.e. high

(> 30 °C) and low (< 15 °C) temperatures, which reduces seed yield considerably. Thus, high and stable yielding cultivars of chickpea during such stress conditions need to be developed (Chaturvedi, Nadarajan, 2010; Krishnamurthy et al., 2010; Devasirvatham et al., 2015; Devasirvatham, Tan, 2018).

Conclusion

Stability analysis was performed by analytical (FR and PLSR) and experimental (AMMI analysis and HMRPGV based on mixed models) approaches. Simultaneous selection index was superior to AMMI indices for identifying stable and high yielding genotypes. Comparison between HMRPGV method and AMMI indices shows that HMRPGV index relies more on seed yield performance than stability of the genotype, so that genotypes G11 and G13, which were not stable in any of the AMMI indices and had specific adaptation to environments E7 and E3, respectively, with HMRPGV stability index, have been identified as superior genotypes. Factorial regression indicated that rainfall is very important at the beginning and end of the growing seasons. The PLSR biplot indicated that E3, E5 and E7 can be considered as favorable environments in terms of seasonal rainfall and temperature and G11, G14, G9, G13, G5, G10, G15 and G12 can be identified as superior genotypes in favorable conditions. In general, based on different methods, genotype G14 had good performance and stability of seed yield in many environments and in all of the methods and could be a candidate for introduction of new cultivars. The PLSR biplot also identified genotype G5 as a suitable genotype for moisture and temperature stresses conditions.

References

- Adugna W., Labuschagne M.T. Genotype-environment interactions and phenotypic stability analyses of linseed in Ethiopia. *Plant Breed.* 2002;121(1):66-71.
- Ahakpaz F., Abdi H., Neyestani E., Hesami A., Mohammadi B., Nader Mahmoudi K., Abedi-Asl G., Jazayeri Noshabadi M.R., Ahakpaz F., Alipour H. Genotype-by-environment interaction analysis for grain yield of barley genotypes under dryland conditions and the role of monthly rainfall. *Agric. Water Manag.* 2021;245:106665. DOI 10.1016/j.agwat.2020.106665.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Aut. Cont.* 1974;19(6):716-723. DOI 10.1109/TAC.1974.1100705.
- Annicchiarico P. Genotype × environment interactions: challenges and opportunities for plant breeding and cultivar recommendations (No. 174). Rome: Food and Agriculture Organization, 2002.
- Azam M.G., Iqbal M.S., Hossain M.A., Hossain M.F. Stability investigation and genotype × environment association in chickpea genotypes utilizing AMMI and GGE biplot model. *Gen. Mol. Res.* 2020; 19(3):16039980.
- Basford K.E., Cooper M. Genotype by environment interactions and some considerations of their implications for wheat breeding in Australia. *Aust. J. Agric. Res.* 1998;49(2):153-174. DOI 10.1071/A97035.
- Campbell B.T., Baenziger P.S., Eskridge K.M., Budak H., Streck N.A., Weiss A., Gill K.S., Erayman M. Using environmental covariates to explain genotype × environment and QTL × environment interactions for agronomic traits in chromosome 3A of wheat. *Crop Sci.* 2004;44(2):620-627. DOI 10.2135/cropsci2004.6200.
- Chaturvedi S.K., Nadarajan N. Genetic enhancement for grain yield in chickpea—accomplishments and resetting research agenda. *Elec. J. Plant Breed.* 2010;1(4):611-615.
- Coan M.M.D., Marchioro V.S., Franco A., Pinto R.J.B., Scapim C.A., Baldissera J.N.C. Determination of genotypic stability and adaptability in wheat genotypes using mixed statistical models. *J. Agric. Sci. Tech.* 2018;20(7):1525-1540.
- Colombari-Filho J.M., Resende M.D.V., de Moraes O.P., Castro A.P., Guimaraes E.L., Pereira J.M., Utumi M.M., Bresghehlo F. Upland rice breeding in Brazil: A simultaneous genotypic evaluation of stability, adaptability and grain yield. *Euphytica.* 2013;192:117-129. DOI 10.1007/s10681-013-0922-2.
- Das R.R., Anil Kumar V., Rakshit S., Maraboina R., Panwar S., Sava-dia S., Rathore A. Interpreting genotype by environment interaction using weather covariates. *Stat. Appl.* 2012;10(1-2):45-62.
- Devasirvatham V., Gaur P.M., Mallikarjuna N., Tokachichu R.N., Trethowan R.M., Tan D.K.Y. Effect of high temperature on the reproductive development of chickpea genotypes under controlled environments. *Func. Plant Biol.* 2012;39:1009-1018. DOI 10.1071/FP12033.
- Devasirvatham V., Gaur P.M., Raju T.N., Trethowan R.M., Tan D.K.Y. Field response of chickpea (*Cicer arietinum* L.) to high temperature. *Field Crop. Res.* 2015;172:59-71. DOI 10.1016/j.fcr.2014.11.017.
- Devasirvatham V., Tan D. Impact of high temperature and drought stresses on chickpea production. *Agron. J.* 2018;8(8):145. DOI 10.3390/agronomy8080145.
- Dubey S.K., Sah U., Singh S.K. Impact of climate change in pulse productivity and adaptation strategies as practiced by the pulse growers of Bundelkhand region of Uttar Pradesh. *J. Food Leg.* 2011;24(3): 230-234.
- Farshadfar E. Incorporation of AMMI stability value and grain yield in a single non-parametric index (GSI) in bread wheat. *Pak. J. Biol. Sci.* 2008;11(14):1791-1796. DOI 10.3923/pjbs.2008.1791.1796.
- Farshadfar E., Rashidi M., Jowkar M.M., Zali H. GGE Biplot analysis of genotype × environment interaction in chickpea genotypes. *Europ. J. Experim. Biol.* 2013;3(1):417-423.
- Farshadfar E., Zali H., Mohammadi R. Evaluation of phenotypic stability in chickpea genotypes using GGE-Ballot. *Ann. Biol. Res.* 2011; 2(6):282-292.
- Funga A., Tadesse M., Eshete M., Fikre A., Korbu L., Girma N., Bekele D., Mohamed R., Bishaw Z., Rao G., Siambi M., Monyo E., Gaur P., Ojiewo C. Genotype by environment interaction on yield stability of desi type chickpea (*Cicer arietinum* L.) at major chickpea producing areas of Ethiopia. *Austr. J. Crop Sci.* 2017;11(2):212-219. DOI 10.21475/ajcs.17.11.02.p297.
- Gauch H.G., Zobel R.W. Identifying mega-environments and targeting genotypes. *Crop Sci.* 1997;37(2):311-326. DOI 10.2135/cropsci1997.0011183X003700020002x.
- Gauch H.G., Zobel R.W. Predictive and postdictive success of statistical analyses of yield trials. *Theor. Appl. Gen.* 1988;76(1):1-10. DOI 10.1007/BF00288824.
- Gaur P.M., Tripathi S., Gowda C.L.L., Ranga Rao G.V., Sharma H.C., Pande S., Sharma M. Chickpea seed production manual. Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics, 2010.
- Hilmarsson H.S., Rio S., Sánchez J.I. Genotype by environment interaction analysis of agronomic spring barley traits in Iceland using AMMI, factorial regression model and linear mixed model. *Agronomy.* 2021;11(3):499. DOI 10.3390/agronomy11030499.
- Joshi A.K., Crossa J., Arun B., Chand R., Trethowan R., Vargasf M., Ortiz-Monasterio I. Genotype × environment interaction for zinc and iron concentration of wheat grain in eastern Gangetic plains of India. *Field Crop. Res.* 2010;116(3):268-277. DOI 10.1016/j.fcr.2010.01.004.
- Kadiyala M.D., Kumara Charyulu M., Nedumaran D., Shyam S.D., Gumma M.M.K., Bantilan M.C.S. Agronomic management options for sustaining chickpea yield under climate change scenario. *J. Agromet.* 2016;18(1):41-47.
- Kalra N., Chakraborty D., Sharma A., Rai H.K., Jolly M., Chander S., Kumar P.R., Bhadraray S., Barman D., Mittal R.B. Effect of temperature on yield of some winter crops in northwest India. *Current Sci.* 2008;94:82-88.
- Kanouni H., Farayedi Y., Saeid A., Sabaghpour S.H. Stability analyses for seed yield of chickpea (*Cicer arietinum* L.) genotypes in the western cold zone of Iran. *J. Agric. Sci.* 2015;7(5):219-230. DOI 10.5539/jas.v7n5p219.

- Karimizadeh R., Mohammadi M. AMMI adjustment for rainfed lentil yield trials in Iran. *Bulg. J. Agric. Sci.* 2010;16:66-73.
- Kondić-Špika A., Mladenov N., Grahovac N., Zorić M., Mikić S., Trkulja D., Marjanović-Jeromela A., Miladinović D., Hristov N. Biometric analyses of yield, oil and protein contents of wheat (*Triticum aestivum* L.) genotypes in different environments. *Agronomy*. 2019;9(6):270. DOI 10.3390/agronomy9060270.
- Krishnamurthy L., Kashiwagi J., Gaur P.M., Upadhyaya H.D., Vadez V. Sources of tolerance to terminal drought in the chickpea (*Cicer arietinum* L.) minicore germplasm. *Field Crop. Res.* 2010;119:322-330. DOI 10.1016/j.fcr.2010.08.002.
- Mendiburu F. *Agricolae: Statistical Procedures for Agricultural Research*. R Package version 1.3-1. 2019. <https://CRAN.R-project.org/package=agricolae>
- Moghadam A. Simultaneous selection for yield and stability and its comparison with stability different statistics. *Seed Plant.* 2003;19(2):1-13.
- Mohammadi R., Sadeghzadeh B., Ahmadi M.M., Amri A. Biological interpretation of genotype × environment interaction in rainfed durum wheat. *Cereal Res. Commun.* 2020a;48(4):547-554. DOI 10.1007/s42976-020-00056-7.
- Mohammadi R., Sadeghzadeh B., Poursiahbidi M.M., Ahmadi M.M. Integrating univariate and multivariate statistical models to investigate genotype × environment interaction in durum wheat. *Ann. Appl. Biol.* 2020b;178(3):450-465. DOI 10.1111/aab.12648.
- Olivoto T., DalCol Lucio A. Metan: an R package for multi-environment trial analysis. *Meth. Ecol. Evol.* 2020;11(6):783-789. DOI 10.1111/2041-210X.13384.
- Olivoto T., Lucio A.D.C., da Silva J.A.G., Marchioro V.S., de Souza V.Q. Mean performance and stability in multi-environment trials I: Combining features of AMMI and BLUP techniques. *Agron. J.* 2019;111(6):2949-2960. DOI 10.2134/agronj2019.03.0220.
- Pacheco A., Vargas M., Alvarado G., Rodríguez F., Crossa J., Burgueño J. GEA-R (genotype × environment analysis with R for Windows), Version 2.0. Mexico: CIMMYT, 2016. <http://hdl.handle.net/11529/10203>. (Accessed 20 June 2016).
- Pouresmael M., Kanouni H., Hajihasani M., Astraki H., Mirakhorli A. Stability of chickpea (*Cicer arietinum* L.) landraces in national plant gene bank of Iran for drylands. *J. Agric. Sci. Technol.* 2018;20(2):387-400.
- Purchase J.L., Hatting H., Van Deventer C.S. Genotype × environment interaction of winter wheat (*Triticum aestivum* L.) in South Africa: II. Stability analysis of yield performance. *South Afric. J. Plant Soil.* 2000;17(3):101-107. DOI 10.1080/02571862.2000.10634878.
- Resende M.D.V. *Matemática e Estatística na Análise de Experimentos e no Melhoramento Genético*. Colombo: Embrapa Florestas, 2007.
- Richards M.F., Preston A.L., Napier T., Jenkins L., Maphosa, L. Sowing date affects the timing and duration of key chickpea (*Cicer arietinum* L.) growth phases. *Plant.* 2020;9(10):1257. DOI 10.3390/plants9101257.
- Richards R.A. Breeding and selection for drought resistant wheat. In: *Drought Resistance in Crops with Emphasis on Rice*. Manila: International Rice Research Institute, 1982;303-316.
- Rodvalho M.A., Coan M.M.D., Scapim C.A., Pinto R.J.B., Contreiras-Soto R.I. Comparison of HMRPGV, Lin and Binn's and Anni-chiarico's methods for maize hybrid selection for high and stable yield. *Maydica*. 2015;60(1):M10 ref.many.
- Romay C.M., Malvar R.A., Ramirez L.C., Alvarez A., Moreno-Gonzalez J., Ordas A., Revilla P. Climatic and genotypic effects for grain yield in maize under stress conditions. *Crop Sci.* 2010;50(1):51-58. DOI 10.2135/cropsci2008.12.0695.
- Sabaghpour S.H., Mahmodi A.A., Saeed A., Kamel M., Malhotra R.S. Study on chickpea drought tolerance lines under dryland condition of Iran. *Ind. J. Crop Sci.* 2006;1(1-2):70-73.
- Sayar M.S. Additive main effects and multiplicative interactions (AMMI) analysis for fresh forage yield in common vetch (*Vicia sativa* L.) genotypes. *Agric. Fores.* 2017;63:119-127. DOI 10.17707/AgriForest.63.1.14.
- Sneller C.H., Kilgore-norquest L., Dombek D. Repeatability of yield stability statistics in soybean. *Crop Sci.* 1997;37(2):383-390. DOI 10.2135/cropsci1997.0011183X003700020013x.
- Stojaković M., Mitrović B., Zorić M., Ivanović M., Stanisavljević D., Nastasić A., Dodig D. Grouping pattern of maize test locations and its impact on hybrid zoning. *Euphytica*. 2015;204:419-431. DOI 10.1007/s10681-015-1358-7.
- Tilahun G., Mekbib F., Fikre A., Eshete M. Genotype × environment interaction and stability analysis for yield and yield related traits of Kabuli-type Chickpea (*Cicer arietinum* L.) in Ethiopia. *Afr. J. Biotechnol.* 2015;14(18):1564-1575. DOI 10.5897/AJB2014.14320.
- Upadhaya H.D., Dronavalli N., Gowda C.L.L., Singh S. Identification and evaluation of chickpea germplasm for tolerance to heat stress. *Crop Sci.* 2011;51(5):2079-2094. DOI 10.2135/cropsci2011.01.0018.
- van Eeuwijk F.A., Denis J.B., Kang M.S. Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: Kang M.S., Gauch H.G. (Eds.). *Genotype-by-environment interaction*. Boca Raton, Florida: CRC Press, 1996;15-49.
- Vargas M., Crossa J., Sayre K., Reynolds M., Ramirez M., Talbot M. Interpreting genotype × environment interaction in wheat by partial least squares regression. *Crop Sci.* 1998;38(3):679-689. DOI 10.2135/cropsci1998.0011183X003800030010x.
- Vargas M., Crossa J., van Eeuwijk F.A., Ramirez M.E., Sayre K. Using partial least squares, factorial regression and AMMI models for interpreting genotype × environment interaction. *Crop Sci.* 1999;39(4):955-967. DOI 10.2135/cropsci1999.0011183X003900040002x.
- Verma A., Singh G.P. Simultaneous application of AMMI measures and yield for stability analysis of wheat genotypes evaluated under irrigated late sown conditions of Central Zone of India. *J. Appl. Nat. Sci.* 2020;12(4):541-549. DOI 10.31018/jans.v12i4.2391.
- Voltaš J., Lopez-Corcoles H., Borrás G. Use of biplot analysis and factorial regression for the investigation of superior genotypes in multi-environment trials. *Europ. J. Agron.* 2005;22(3):309-324. DOI 10.1016/j.eja.2004.04.005.
- Yan W., Cornelius P.L., Crossa J., Hunt L.A. Two types of GGE biplots for analyzing multi-environment trial data. *Crop Sci.* 2001;41(3):656-663. DOI 10.2135/cropsci2001.413656x.
- Yan W., Hunt L.A., Shen Q., Szlavnic Z. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 2000;40(3):597-605. DOI 10.2135/cropsci2000.403597x.
- Yan W., Kang M.S. GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists. Boca Raton, Florida: CRC Press LLC, 2003.
- Yan W., Pageau D., Fréreau-Reid J., Durand J. Assessing the representativeness and repeatability of test locations for genotype evaluation. *Crop Sci.* 2011;51(4):1603-1610. DOI 10.2135/cropsci2011.01.0016.
- Zali H., Farshadfar E., Sabaghpour S.H., Karimizadeh R. Evaluation of genotype × environment interaction in chickpea using measures of stability from AMMI model. *Ann. Biol. Res.* 2012;3(7):3126-3136.
- Zhang Z., Lu C., Xiang Z. Analysis of variety stability based on AMMI model. *Acta Agron. Sin.* 1998;24(3):304-309.
- Zobel R.W., Wright A.J., Gauch H.G. Statistical analysis of a yield trial. *Agron. J.* 1988;80(3):388-393. DOI 10.2134/agronj1988.0002196208000030002x.

ORCID ID

P. Sharifi orcid.org/0000-0002-1658-7366

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

Received March 27, 2022. Revised August 27, 2022. Accepted August 29, 2022.

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

The mitochondrial genome of *Dendrobaena tellermanica* Perel, 1966 (Annelida: Lumbricidae) and its phylogenetic position

S.V. Shekhovtsov^{1,2}✉, G.V. Vasiliev¹, R. Latif³, T.V. Poluboyarova¹, S.E. Peltek¹, I.B. Rapoport⁴

¹ Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Institute of Biological Problems of the North of the Far Eastern Branch of the Russian Academy of Sciences, Magadan, Russia

³ Semnan University, Semnan, Iran

⁴ Tembotov Institute of Ecology of Mountain Territories of Russian Academy of Sciences, Nalchik, Russia

✉ shekhovtsov@bionet.nsc.ru

Abstract. Earthworms are an important ecological group that has a significant impact on soil fauna as well as plant communities. Despite their importance, genetic diversity and phylogeny of earthworms are still insufficiently studied. Most studies on earthworm genetic diversity are currently based on a few mitochondrial and nuclear genes. Mitochondrial genomes are becoming a promising target for phylogeny reconstruction in earthworms. However, most studies on earthworm mitochondrial genomes were made on West European and East Asian species, with much less sampling from other regions. In this study, we performed sequencing, assembly, and analysis of the mitochondrial genome of *Dendrobaena tellermanica* Perel, 1966 from the Northern Caucasus. This species was earlier included into *D. schmidti* (Michaelsen, 1907), a polytypic species with many subspecies. The genome was assembled as a single contig 15,298 bp long which contained a typical gene set: 13 protein-coding genes (three subunits of cytochrome *c* oxidase, seven subunits of NADH dehydrogenase, two subunits of ATP synthetase, and cytochrome *b*), 12S and 16S ribosomal RNA genes, and 22 tRNA genes. All genes were located on one DNA strand. The assembled part of the control region, located between the *tRNA-Arg* and *tRNA-His* genes, was 727 bp long. The control region contained multiple hairpins, as well as tandem repeats of the AACGCTT monomer. Phylogenetic analysis based on the complete mitochondrial genomes indicated that the genus *Dendrobaena* occupied the basal position within Lumbricidae. *D. tellermanica* was a rather distant relative of the cosmopolitan *D. octaedra*, suggesting high genetic diversity in this genus. *D. schmidti* turned out to be paraphyletic with respect to *D. tellermanica*. Since *D. schmidti* is known to contain very high genetic diversity, these results may indicate that it may be split into several species.

Key words: earthworms; Lumbricidae; *Dendrobaena tellermanica*; mitochondrial genomes.

For citation: Shekhovtsov S.V., Vasiliev G.V., Latif R., Poluboyarova T.V., Peltek S.E., Rapoport I.B. The mitochondrial genome of *Dendrobaena tellermanica* Perel, 1966 (Annelida: Lumbricidae) and its phylogenetic position. *Vavilovskii Zhurnal Genetiki i Selekcii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):146-152. DOI 10.18699/VJGB-23-20

Митохондриальный геном *Dendrobaena tellermanica* Perel, 1966 (Annelida: Lumbricidae)

С.В. Шеховцов^{1,2}✉, Г.В. Васильев¹, Р. Латиф³, Т.В. Полубаярова¹, С.Е. Пельтек¹, И.Б. Рапопорт⁴

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Институт биологических проблем Севера Дальневосточного отделения Российской академии наук, Магадан, Россия

³ Университет Семнана, Семнан, Иран

⁴ Институт экологии горных территорий им. А.К. Темботова Российской академии наук, Нальчик, Россия

✉ shekhovtsov@bionet.nsc.ru

Аннотация. Дождевые черви – важная экологическая группа, которая оказывает значительное влияние как на состав почвенной фауны, так и на растительность. Генетическое разнообразие и филогения дождевых червей при этом остаются относительно слабо изученными. В настоящее время большинство работ по генетическому разнообразию дождевых червей основывается на единичных митохондриальных и ядерных генах. В связи с этим для реконструкции филогенетических отношений у дождевых червей перспективными становятся митохондриальные геномы. Почти все работы по этой теме посвящены видам из Западной Европы или Восточной Азии, другие регионы практически не затронуты. В настоящей работе мы провели секвенирование, сборку и анализ митохондриального генома *Dendrobaena tellermanica* Perel, 1966. Этот вид ранее входил в состав кавказского вида *D. schmidti* (Michaelsen, 1907) – политипического вида, в пределах которого выделяли множество таксонов. Геном был собран в виде одного контига длиной 15 298 п. н., содержащего типичный набор генов: 13 белок-кодирующих генов (три субъединицы цитохромоксидазы, семь субъединиц NADH дегидрогеназы, две субъединицы АТФ синтетазы, цитохром *b*), гены 12S и 16S рибосомальной РНК и 22 гена тРНК. Все гены были расположены на

одной цепи ДНК. На контрольный регион, находящийся между генами *tRNA-Arg* и *tRNA-His*, приходилось 727 п. н. Контрольный регион содержал множество шпилек, а также tandemные повторы мономера AACGCTT. Филогенетический анализ на основе полных митохондриальных геномов показал, что род *Dendrobaena* является базальным в семействе Lumbricidae. *D. tellermanica* оказалась довольно далеким родственником космополитного вида *D. octaedra*, что говорит о высоком генетическом разнообразии в этом роде. Вид *D. schmidtii* был парафилетичным по отношению к *D. tellermanica*. Поскольку для *D. schmidtii* характерна очень высокая генетическая изменчивость, можно рассматривать эти данные как свидетельство в пользу разделения *D. schmidtii* на несколько видов.

Ключевые слова: дождевые черви; Lumbricidae; *Dendrobaena tellermanica*; митохондриальные геномы.

Introduction

Earthworms are an important ecological group that accounts for the highest biomass among the soil fauna in many habitats (Hendrix et al., 2008). Its representatives process plant detritus to soil humus and return organic matter to the global cycles (Blouin et al., 2013). Earthworms also form soil structure, which has high impact on both soil fauna composition and vegetation (Lavelle et al., 2016). Therefore, this group defines ecosystem productivity in many respects.

Genetic diversity and phylogeny of earthworms remain insufficiently studied (Marchán et al., 2018). Currently, most works on earthworm genetic diversity are based on single mitochondrial and nuclear genes (Jamieson et al., 2002; Marchán et al., 2022). Construction of multigene nuclear datasets is impeded by frequent polyploidy characteristic for this group (Viktorov, 1997; Vsevolodova-Perel, Bulatova, 2008; Mezhzherin et al., 2018), which makes it hard to detect suitable orthologs and amplify them by PCR.

Mitochondrial genomes are thus a promising tool for reconstruction of phylogenetic relationships in earthworms. A lot of mitochondrial genomes were sequenced and published in recent years (Zhang L. et al., 2014–2016a, b; Wang et al., 2015; Conrado et al., 2017; Hong et al., 2017; Shekhovtsov, Peltek, 2019; Zhang Q. et al., 2019; Liu et al., 2020; Seto et al., 2021; Csuzdi et al., 2022; Kim, Hong, 2022), and studies on phylogenetic relationships of certain groups were also conducted (Shekhovtsov et al., 2020a; Liu et al., 2021). However, almost all of these studies were made on species from West Europe and East Asia, with almost no representatives of other regions.

In this study, we performed sequencing, assembly, and analysis of the mitochondrial genome of *Dendrobaena tellermanica* Perel, 1966. This species was earlier included into *D. schmidtii* (Michaelsen, 1907), a polytypic species that was considered to contain multiple subspecies (Michaelsen, 1907; Kvavadze, 1985). *D. tellermanica* was believed to be a parthenogenetic form of *D. schmidtii* (Perel, 1966). T.S. Vsevolodova-Perel (2003) demonstrated that many populations of *D. tellermanica* are amphimictic (sexual) and so isolated it into a separate species. *D. tellermanica* differs from *D. schmidtii* by the lack of pigmentation, different position of the clitellum and the form of tuberculae pubertatis (Vsevolodova-Perel, 2003).

Currently, there is only one complete mitochondrial genome of the genus *Dendrobaena* in GenBank belonging to the cosmopolitan *D. octaedra* (Savigny, 1826). The mitochondrial genome of *D. tellermanica* will be the first sequenced mitochondrial genome of a Caucasian earthworm and will be important for studying the phylogeny of lumbricids.

Materials and methods

Specimens of *D. tellermanica* were collected in the Karachay-Cherkess Republic (right bank of r. Uchkulan, road to the Chiper pass, 1483 a. s. l., 4–5 km from the Aktyube town, *Alchemilla* and *Geranium* meadow, N 43.410944, E 42.174538). Worms were fixed in ethanol. Morphological identification was performed according to the key of T.S. Vsevolodova-Perel (1997).

DNA was extracted using the standard phenol-chloroform method and sonicated on Covaris M220 to the target fragment length of 350 bp. The fragments were purified by 1.2 volume of AMPureXP (Beckman Coulter, USA) and quantified using fluorometry on a Qubit device. Genomic libraries were obtained from 100 ng of DNA using Roche KAPA Hyper Prep according to the manufacturer's protocol using KAPA UDI Adapter double barcodes. Quality and molarity of the obtained genomic library was assessed on a BA2100 bioanalyzer using the Agilent DNA High Sensitivity Kit and sequenced on an Illumina NextSeq550 with the Mid Output Kit v. 2.5 (300 Cycles) for 2×150 bp paired reads.

The obtained data were processed by TrimmomaticPE (Bolger et al., 2014) with the ILLUMINACLIP:TruSeq3-SE:2:30:10 SLIDINGWINDOW:4:15 MINLEN:36 options. SPAdes v. 3.14.1 was used for contig assembly (Bankevich et al., 2012) with the --isolate option. The assembled contigs were aligned with mitochondrial earthworm genomes from the NCBI database with blastn (<https://blast.ncbi.nlm.nih.gov>) in order to search for mitochondrial sequences.

Preliminary annotation was done by MITOS 2 (Bernt et al., 2013) with subsequent manual comparison with annotated earthworm genomes. The mitochondrial genome of *D. tellermanica* was deposited in GenBank under accession number ON960857. Map of the genome was constructed using Benchling (<https://www.benchling.com/>).

Secondary structures of tRNAs were visualized using MITOS 2 (Bernt et al., 2013); of the control region, using RNAfold Web Server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>) and forna (<http://rna.tbi.univie.ac.at/forna/forna.html>) (Gendron et al., 2001). Search for tandem repeats was done by Tandem Repeats Finder (Benson, 1999). For phylogenetic reconstructions, mitochondrial genomes were aligned with Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>); control regions were not included into the alignments. Ambiguously aligned regions were removed with gblocks 0.91b (Castresana, 2000). Earthworm mitochondrial genomes and sequences of the *COXI* gene of representatives of the *Dendrobaena* genus were extracted from GenBank. Phylogenetic trees were built using the Maximum Likelihood approach in RAxML v. 8.2.12 (Stamatakis,

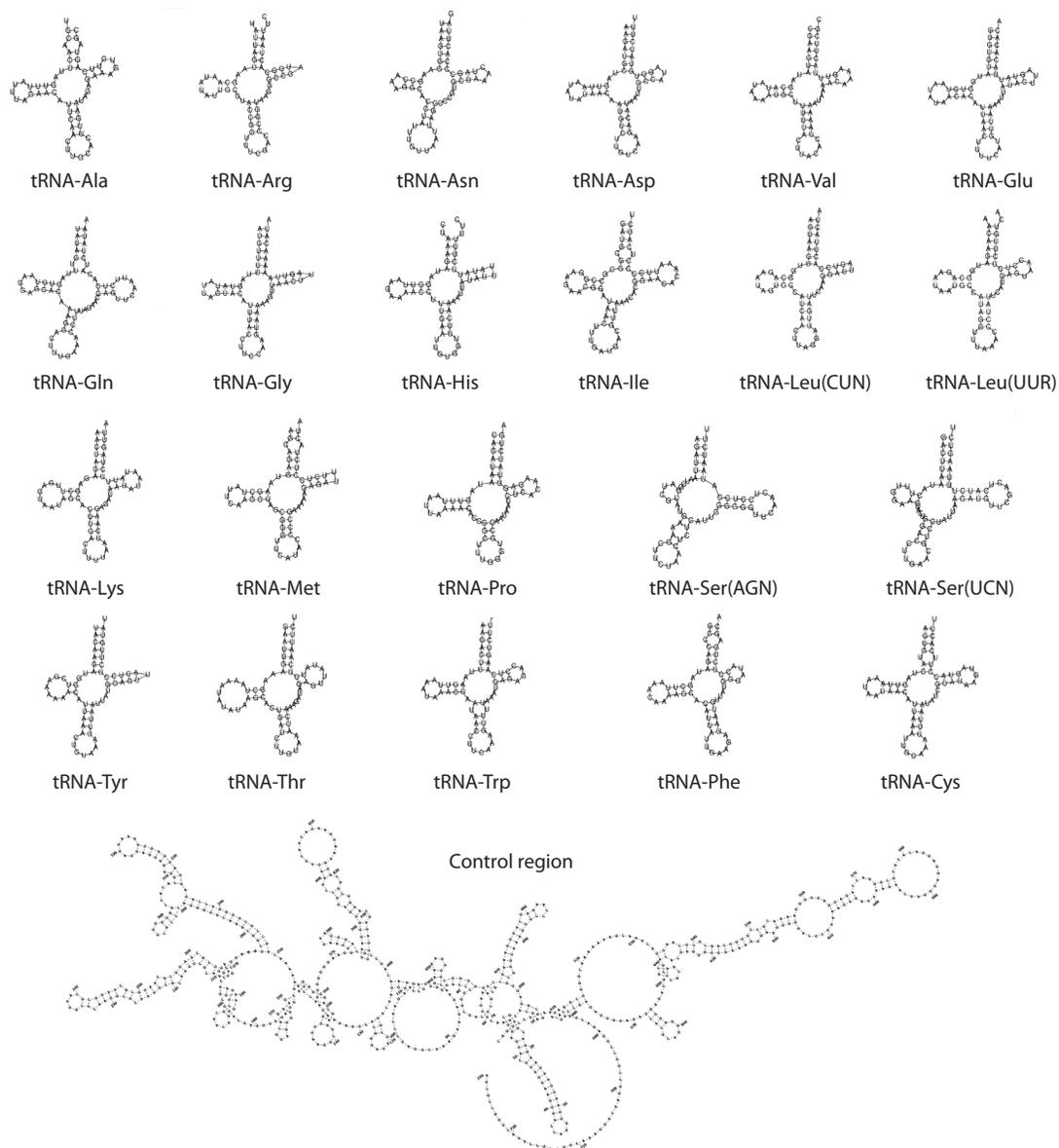


Fig. 2. Secondary structures of tRNAs and the control region of the mitochondrial genome of *D. tellermanica*.

Within the family Lumbricidae, which includes *D. tellermanica*, the genus *Dendrobaena* was the sister group to all other genera of the family. There are only two known mitochondrial sequences for the genus *Dendrobaena*, the cosmopolitan *D. octaedra* and *D. tellermanica* obtained in this study. Expectedly, *D. tellermanica* forms a clade with *D. octaedra*, but they are rather distantly related.

We can conclude that this work is a first step in the study of the basal branches of Lumbricidae. Representatives of the genus *Dendrobaena* from the Caucasus are particularly interesting in this respect, because they account for a large part of its species diversity.

Earlier we performed a genetic analysis of morphological forms of *D. schmidtii* (Shekhovtsov et al., 2020b), demonstrating that it represents at least two separate species. On the phylogenetic tree constructed using the *COXI* gene (Fig. 4), *D. tellermanica* was inside one of the branches of *D. schmidtii*.

We should note that single mitochondrial genes, including *COXI*, are unsuitable for phylogenetic reconstruction on the family level (Klarica et al., 2012; Shekhovtsov et al., 2016, 2020c), since they demonstrate poor resolution of the relationships between species and do not support the monophyly of most genera. *COXI* is however of much use in the search for closely related species or genetic lineages. Moreover, there are thousands of *COXI* sequences in the public databases and only a few mitochondrial genomes; e.g., mtDNA of *D. schmidtii* has not been sequenced yet. Therefore, the tree on Fig. 4 is given only to demonstrate the close relationship of *D. tellermanica* and *D. schmidtii*.

Conclusion

The obtained preliminary results indicate that *D. tellermanica* could be treated as a subspecies of *D. schmidtii*, as was believed earlier, or split *D. schmidtii* into several species. The latter

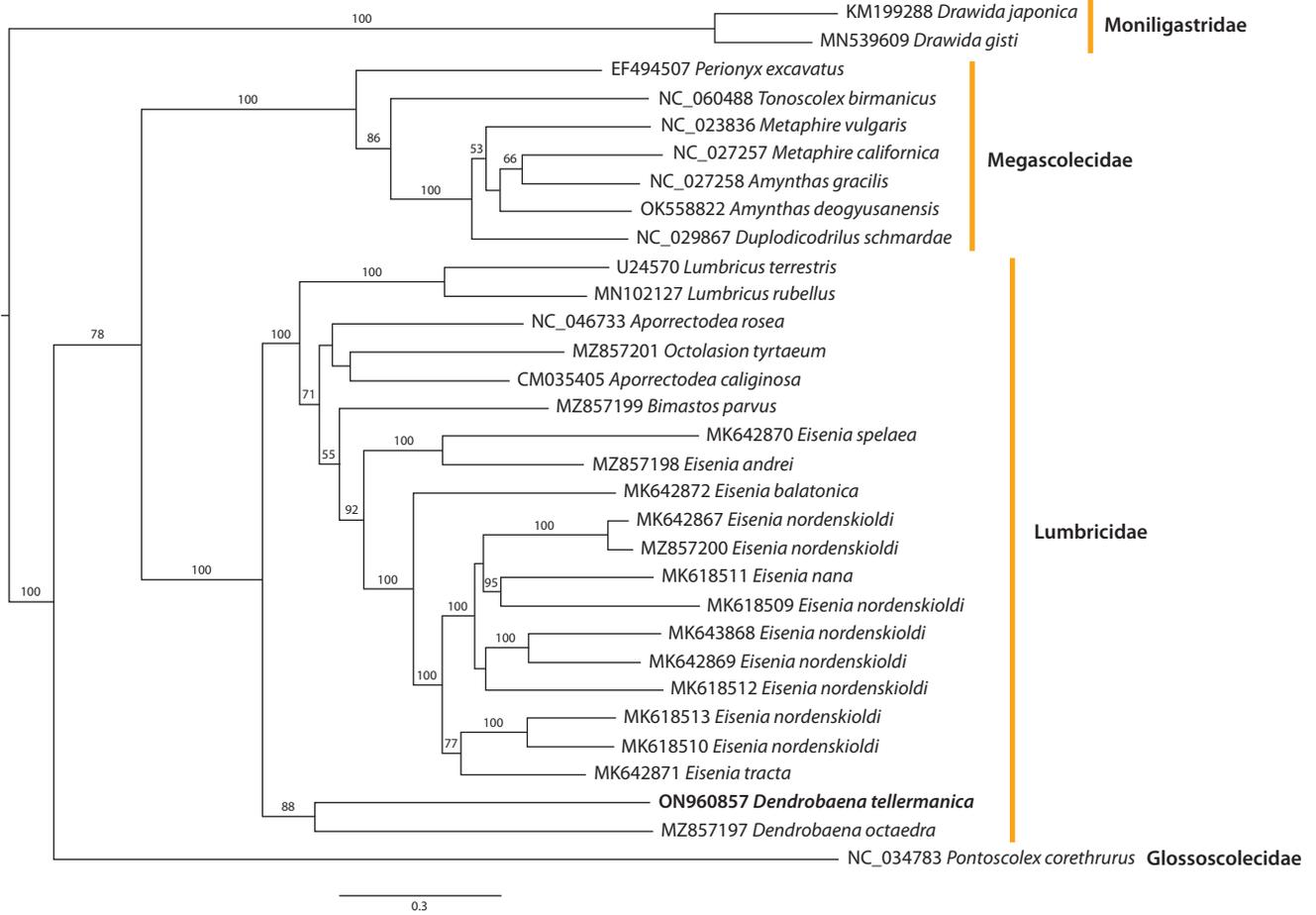


Fig. 3. Phylogenetic tree based on earthworm mitochondrial genomes using the Maximum Likelihood method. Here and in Fig. 4: Numbers near the branches indicate bootstrap support.

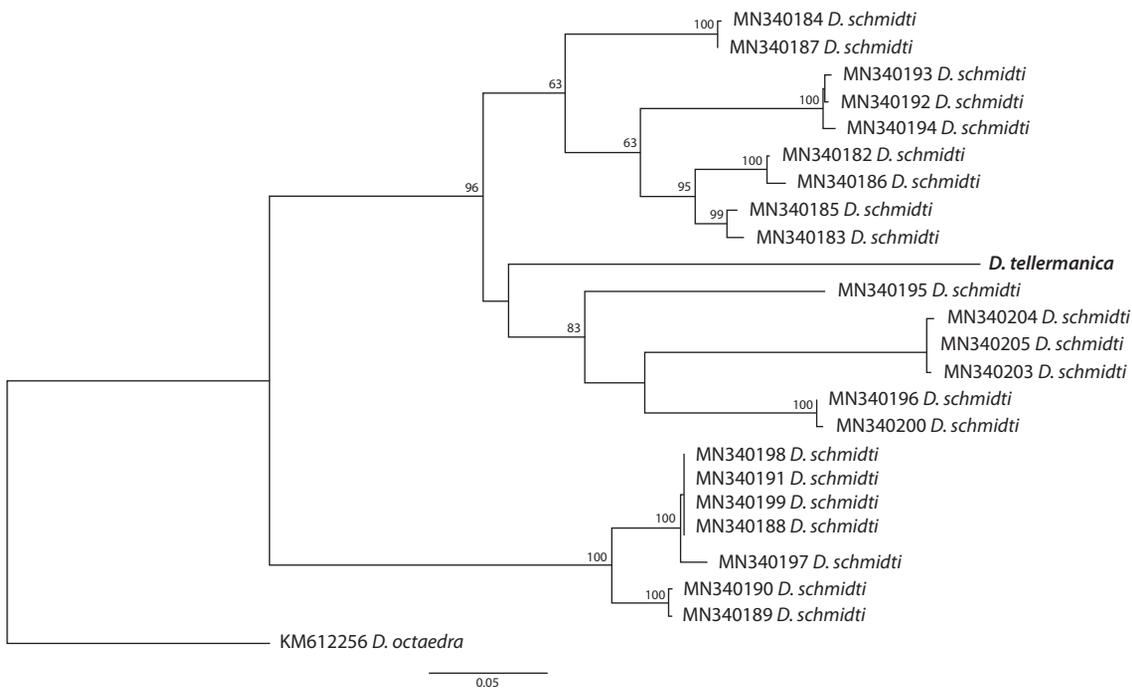


Fig. 4. Phylogenetic tree based on the COX1 gene for the genus *Dendrobaena* using the Maximum Likelihood method.

option is supported by the high genetic and morphological variation within this species. However, such conclusions would require an analysis based on several loci, including nuclear ones.

References

- Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., Nikolenko S.I., Pham S., Pribelski A.D., Pyskin 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.
- Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 1999;27(2):573-580. DOI 10.1093/nar/27.2.573.
- Bernt M., Donath A., Jühling F., Externbrink F., Florentz C., Fritsch G., Pütz J., Middendorf M., Stadler P.F. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 2013;69(2):313-319. DOI 10.1016/j.ympev.2012.08.023.
- Blouin M., Hodson M.E., Delgado E.A., Baker G., Brussaard L., Butt K.R., Dai J., Dendooven L., Peres G., Tondoh J.E., Cluzeau D., Brun J.-J. A review of earthworm impact on soil function and ecosystem services. *Eur. J. Soil Sci.* 2013;64(2):161-182. DOI 10.1111/ejss.12025.
- Bolger A.M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-2120. DOI 10.1093/bioinformatics/btu170.
- Boore J.L., Brown W.M. Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. *Genetics.* 1995;141(1):305-319. DOI 10.1093/genetics/141.1.305.
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 2000;17(4):540-552. DOI 10.1093/oxfordjournals.molbev.a026334.
- Clayton D.A. Transcription and replication of animal mitochondrial DNAs. *Int. Rev. Cytol.* 1992;141:217-232. DOI 10.1016/S0074-7696(08)62067-7.
- Conrado A.C., Arruda H., Stanton D.W.G., James S.W., Kille P., Brown G., Silva E., Dupont L., Taheri S., Morgan A.J., Simões N., Rodrigues A., Montiel R., Cunha L. The complete mitochondrial DNA sequence of the pantropical earthworm *Pontoscolex corethrurus* (Rhindrilidae, Clitellata): Mitogenome characterization and phylogenetic positioning. *ZooKeys.* 2017;688:1-13. DOI 10.3897/zookeys.688.13721.
- Csuzdi C., Koo J., Hong Y. The complete mitochondrial DNA sequences of two sibling species of lumbricid earthworms, *Eisenia fetida* (Savigny, 1826) and *Eisenia andrei* (Bouché, 1972) (Annelida, Crassidellata): comparison of mitogenomes and phylogenetic positioning. *ZooKeys.* 2022;1097:167-181. DOI 10.3897/zookeys.1097.80216.
- Gendron P., Lemieux S., Major F. Quantitative analysis of nucleic acid three-dimensional structures. *J. Mol. Biol.* 2001;308(5):919-936. DOI 10.1006/jmbi.2001.4626.
- Hendrix P.F., Callahan M.A., Drake J.M., Huang C.-Y., James S.W., Snyder B.A., Zhang W. Pandora's box contained bait: the global problem of introduced earthworms. *Annu. Rev. Ecol. Syst.* 2008;39(1):593-613. DOI 10.1146/annurev.ecolsys.39.110707.173426.
- Hong Y., Kim M.J., Wang A.R., Kim I. Complete mitochondrial genome of the earthworm. *Amyntas jiriensis* (Clitellata: Megascolecidae). *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 2017;28(2):163-164. DOI 10.3109/19401736.2015.1115491.
- Jamieson B.G.M., Tillier S., Tillier A., Justine J.-L., Ling E., James S., McDonald K., Hugall A.F. Phylogeny of the Megascolecidae and Crassidellata (Annelida, Oligochaeta): combined versus partitioned analysis using nuclear (28S) and mitochondrial (12S, 16S) rDNA. *Zoosystema.* 2002;24(4):707-734.
- Kim M.J., Hong Y. Complete mitochondrial genome of the earthworm *Amyntas seungpanensis* (Clitellata: Megascolecidae). *Mitochondrial DNA B Resour.* 2022;7(6):989-991. DOI 10.1080/23802359.2022.2080604.
- Klarica J., Kloss-Brandstätter A., Traugott M., Juen A. Comparing four mitochondrial genes in earthworms – Implications for identification, phylogenetics, and discovery of cryptic species. *Soil Biol. Biochem.* 2012;45:23-30. DOI 10.1016/j.soilbio.2011.09.018.
- Kvavadze E.Sh. The Earthworms (Lumbricidae) of the Caucasus. Tbilisi: Metsniereba Publ., 1985. (in Russian)
- Lavelle P., Spain A., Blouin M., Brown G., Decaëns T., Grimaldi M., Jiménez J.J., McKey D., Mathieu J., Velasquez E., Zangerlé A. Ecosystem engineers in a self-organized soil. *Soil Sci.* 2016;181(3/4):91-109. DOI 10.1097/SS.000000000000155.
- Liu H., Xu N., Zhang Q., Wang G., Xu H., Ruan H. Characterization of the complete mitochondrial genome of *Drawida gisti* (Metagnophora, Moniligastridae) and comparison with other Metagnophora species. *Genomics.* 2020;112(5):3056-3064. DOI 10.1016/j.ygeno.2020.05.020.
- Liu H., Zhang Y., Xu W., Fang Y., Ruan H. Characterization of five new earthworm mitogenomes (Oligochaeta: Lumbricidae): mitochondrial phylogeny of Lumbricidae. *Diversity.* 2021;13(11):580. DOI 10.3390/d13110580.
- Marchán D.F., Cosin D.J.D., Novo M. Why are we blind to cryptic species? Lessons from the eyeless. *Eur. J. Soil Biol.* 2018;86:49-51. DOI 10.1016/j.ejsobi.2018.03.004.
- Marchán D.F., Decaëns T., Domínguez J., Novo M. Perspectives in earthworm molecular phylogeny: recent advances in Lumbricoidea and standing questions. *Diversity.* 2022;14(1):30. DOI 10.3390/d14010030.
- Mezhzherin S.V., Garbar A.V., Vlasenko R.P., Onishchuk I.P., Kotsyuba I.Yu., Zhalai E.I. The Evolutionary Paradox of Parthenogenetic Earthworms. Kiev: Naukova Dumka Publ., 2018. (in Ukrainian)
- Michaelsen W. Die Lumbriciden des Kaukasischen Museums in Tiflis. *Mitteilungen des Kaukasischen Museums.* 1907;3:81-93.
- Perel T.S. Earthworms in forest soils of the Northwestern Caucasus. In: The Impact of Animals on the Productivity of Forest Cenoses. Moscow: Nauka Publ., 1966;146-165. (in Russian)
- Seto A., Endo H., Minamiya Y., Matsuda M. The complete mitochondrial genome sequences of Japanese earthworms *Metaphire hilgendorfi* and *Amyntas yunoshimensis* (Clitellata: Megascolecidae). *Mitochondrial DNA B. Resour.* 2021;6(3):965-967. DOI 10.1080/23802359.2020.1830728.
- Shekhovtsov S., Berman D.I., Bazarova N.E., Bulakhova N.A., Porco D., Peltek S.E. Cryptic genetic lineages in *Eisenia nordenskioldi pallida* (Oligochaeta, Lumbricidae). *Eur. J. Soil Biol.* 2016;75:151-156. DOI 10.1016/j.ejsobi.2016.06.004.
- Shekhovtsov S.V., Golovanova E.V., Ershov N.I., Poluboyarova T.V., Berman D.I., Bulakhova N.A., Szederjesi T., Peltek S.E. Phylogeny of the *Eisenia nordenskioldi* complex based on mitochondrial genomes. *Eur. J. Soil Biol.* 2020a;96:103137. DOI 10.1016/j.ejsobi.2019.103137.
- Shekhovtsov S.V., Peltek S.E. The complete mitochondrial genome of *Aporrectodea rosea* (Annelida: Lumbricidae). *Mitochondrial DNA Part B.* 2019;4(1):1752-1753. DOI 10.1080/23802359.2019.1610091.
- Shekhovtsov S.V., Rapoport I.B., Poluboyarova T.V., Geraskina A.P., Golovanova E.V., Peltek S.E. Morphotypes and genetic diversity of *Dendrobaena schmidtii* (Lumbricidae, Annelida). *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2020b;24(1):48-54. DOI 10.18699/VJ20.594. (in Russian)
- Shekhovtsov S.V., Shipova A.A., Poluboyarova T.V., Vasiliev G.V., Golovanova E.V., Geraskina A.P., Bulakhova N.A., Szederjesi T., Peltek S.E. Species delimitation of the *Eisenia nordenskioldi* complex (Oligochaeta, Lumbricidae) using transcriptomic data. *Front. Genet.* 2020c;11:1508. DOI 10.3389/fgene.2020.598196.

- Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9): 1312-1313. DOI 10.1093/bioinformatics/btu033.
- Vallès Y., Boore J.L. Lophotrochozoan mitochondrial genomes. *Integr. Comp. Biol.* 2006;46(4):544-557. DOI 10.1093/icb/ijc056.
- Viktorov A.G. Diversity of polyploid races in the family Lumbricidae. *Soil Biol. Biochem.* 1997;29(3-4):217-221. DOI 10.1016/S0038-0717(96)00086-7.
- Vsevolodova-Perel T.S. The Earthworms of the Russian Fauna: Cadaster and Key. Moscow: Nauka Publ., 1997. (in Russian)
- Vsevolodova-Perel T.S. Addition to the fauna of earthworms (Oligochaeta, Lumbricidae) of the Northern Caucasus. *Zoologicheskii Zhurnal = Zoological Journal*. 2003;82(2):275-280. (in Russian)
- Vsevolodova-Perel T.S., Bulatova N.Sh. Polyploid races of earthworms (Lumbricidae, Oligochaeta) in the East European Plain and Siberia. *Biol. Bull. Russ. Acad. Sci.* 2008;35(4):385-388. DOI 10.1134/S1062359008040092.
- Wang A.R., Hong Y., Win T.M., Kim I. Complete mitochondrial genome of the Burmese giant earthworm, *Tonoscolex birmanicus* (Clitellata: Megascolecidae). *Mitochondrial DNA*. 2015;26(3):467-468. DOI 10.3109/19401736.2013.830300.
- Weigert A., Golombek A., Gerth M., Schwarz F., Struck T.H., Bleidorn C. Evolution of mitochondrial gene order in Annelida. *Mol. Phylogenet. Evol.* 2016;94(Pt. A):196-206. DOI 10.1016/j.ympev.2015.08.008.
- Zhang L., Jiang J., Dong Y., Qiu J. Complete mitochondrial genome of an *Amyntas* earthworm, *Amyntas aspergillus* (Oligochaeta: Megascolecidae). *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 2014; 27(3):1-2. DOI 10.3109/19401736.2014.971267.
- Zhang L., Jiang J., Dong Y., Qiu J. Complete mitochondrial genome of four pheretimoid earthworms (Clitellata: Oligochaeta) and their phylogenetic reconstruction. *Gene*. 2015;574(2):308-316. DOI 10.1016/j.gene.2015.08.020.
- Zhang L., Jiang J., Dong Y., Qiu J. Complete mitochondrial genome of a Pheretimid earthworm *Metaphire vulgaris* (Oligochaeta: Megascolecidae). *Mitochondrial DNA*. 2016a;27(1):297-298. DOI 10.3109/19401736.2014.892085.
- Zhang L., Sechi P., Yuan M., Jiang J., Dong Y., Qiu J. Fifteen new earthworm mitogenomes shed new light on phylogeny within the *Pheretima* complex. *Sci. Rep.* 2016b;6:20096. DOI 10.1038/srep20096.
- Zhang Q., Liu H., Zhang Y., Ruan H. The complete mitochondrial genome of *Lumbricus rubellus* (Oligochaeta, Lumbricidae) and its phylogenetic analysis. *Mitochondrial DNA Part B*. 2019;4(2):2677-2678. DOI 10.1080/23802359.2019.1644242.
- Zhao H., Fan S., Aspe N.M., Feng L., Zhang Y. Characterization of four earthworm mitogenomes from Northeast China and phylogenetic implication (Oligochaeta: Lumbricidae, Moniligastridae). *Diversity*. 2022;14(9):714. DOI 10.3390/d14090714.

ORCID ID

S.V. Shekhovtsov orcid.org/0000-0001-5604-5601
R. Latif orcid.org/0000-0002-7201-5426
T.V. Poluboyarova orcid.org/0000-0002-5652-0553
S.E. Peltek orcid.org/0000-0002-3524-0456
I.B. Rapoport orcid.org/0000-0002-6766-1482

Acknowledgements. This study was supported by the grant of the Russian Foundation for Basic Research No. 20-54-5603_Iran_t and of the Iran National Science Foundation No. 99003929, as well as by the State Budget Projects No. AAAA-A18-122011900453-0 and FWNR-2022-0022.

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

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

The use of the primary structure of the ITS1–ITS2 region for species identification in some submerged aquatic macrophytes of the genus *Stuckenia*

A.V. Mglinets, O.E. Kosterin

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

Abstract. Applicability of ITS1–ITS2 primary structure for species attribution of representatives of the genus *Stuckenia* was experimentally tested. Analysis of the ITS1–ITS2 region sequences of *S. vaginata* and *S. pectinata* from public databases showed that they differed by insertions/deletions and single or double nucleotide substitutions. Besides, the ITS1–ITS2 region of *S. pectinata* was shown to be represented by two haplotype groups designated as *S. pectinata* type A and *S. pectinata* type B with good bootstrap support in phylogenetic reconstructions. In 28 samples identified as *S. pectinata*, *S. vaginata*, *S. macrocarpa* and *S. chakassiensis* on the basis of morphology, the ITS1–ITS2 region was sequenced in this study. Three groups of samples with good bootstrap support were revealed to be corresponding to *S. vaginata*, *S. pectinata* type A and *S. pectinata* type B. The *S. vaginata* group was formed by the samples identified on the basis of morphology as *S. vaginata*, and the *S. pectinata* type A group was formed by the samples identified on the basis of morphology as *S. pectinata*. The *S. pectinata* type B group was further divided into two subgroups, *S. pectinata* type B subgroup and *S. chakassiensis* subgroup. The *S. chakassiensis* subgroup included mainly the samples identified as such on the basis of morphology. The *S. pectinata* type B subgroup included samples identified on the basis of morphology as *S. pectinata*, *S. vaginata* and *S. macrocarpa*. We suppose that these samples were *S. pectinata* type B, *S. macrocarpa* and their hybrids.

Key words: Potamogetonaceae; *Stuckenia*; *S. chakassiensis*; *S. macrocarpa*; *S. pectinata*; *S. vaginata*; ITS1–ITS2 region; species identification.

For citation: Mglinets A.V., Kosterin O.E. The use of the primary structure of the ITS1–ITS2 region for species identification in some submerged aquatic macrophytes of the genus *Stuckenia*. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2023;27(2):153-161. DOI 10.18699/VJGB-23-21

Использование первичной структуры района ITS1–ITS2 для видовой идентификации у некоторых представителей водных макрофитов рода *Stuckenia*

А.В. Мглинец, О.Э. Костерин

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

Аннотация. Проведен биоинформационный анализ первичной структуры района ITS1–ITS2 образцов *S. vaginata* и *S. pectinata*, взятых из публичных баз данных, ссылки на которые приведены в опубликованных работах. Показано, что межвидовые различия *S. vaginata* и *S. pectinata* обусловлены делециями/вставками и одно- или двунуклеотидными заменами. Более того, вид *S. pectinata* по структуре района ITS1–ITS2 представлен двумя генотипами, которые обозначены как *S. pectinata* тип А и *S. pectinata* тип В, различия между которыми обусловлены одно- или двунуклеотидными заменами. Это демонстрирует возможность применения данного района для определения видовой принадлежности у представителей рода *Stuckenia*. Для экспериментальной проверки возможности использования данного района у 28 образцов, определенных на основании морфологических признаков как *S. pectinata*, *S. vaginata*, *S. macrocarpa* и *S. chakassiensis*, выполнено определение первичной структуры района ITS1–ITS2. Анализ полученных экспериментальных данных показал, что они распадаются на три группы, третья группа представлена двумя подгруппами. Эти группы соответствуют *S. vaginata*, *S. pectinata* тип А и *S. pectinata* тип В. В группу *S. vaginata* попали образцы, которые на основании морфологических признаков определены как *S. vaginata*. В группу *S. pectinata* тип А попали образцы, которые на основании морфологических признаков определены как *S. pectinata*. Группу *S. pectinata* тип В на основании первичной структуры района ITS1–ITS2 можно разделить на две подгруппы: *S. pectinata* тип В и *S. chakassiensis*. В подгруппу *S. pectinata* тип В вошли образцы, которые на основании морфологических признаков определены как *S. pectinata*, *S. vaginata* и *S. macrocarpa*. В подгруппу *P. chakassiensis* в основном вошли образцы, которые на основании морфологических

признаков определены как *S. chakassiensis*. На основании выравнивания последовательностей подгруппы *S. pectinata* тип В сделано предположение, согласно которому в данную подгруппу объединены последовательности, принадлежащие *S. pectinata* тип В, *S. macrocarpa* и их гибридам.

Ключевые слова: Potamogetonaceae; *Stuckenia*; *S. chakassiensis*; *S. macrocarpa*; *S. pectinata*; *S. vaginata*; район ITS1–ITS2; идентификация видов.

Introduction

Representatives of the genera *Potamogeton* L. and *Stuckenia*, formerly considered as a single genus *Potamogeton*, are aquatic plants present in all the continents except for the Antarctica. They inhabit both fresh and brackish standing and slow-moving waters. Both genera are characterized by high intraspecific morphologic variability causing difficulties for the systematics (Kaplan, Stepanek, 2003). Besides, there exist a lot of interspecies hybrids that are sometimes taken for individual species (Wiegleb, Kaplan, 1998). Another difficulty in the taxonomy of the genus is due to the existence of polyploids and aneuploids (Les, 1983; Hollingsworth et al., 1998; Kaplan, 2002; Fant et al., 2003; 2005; Lindqvist et al., 2006; Kaplan et al., 2009; Kaplan, 2010). According to literature data, the genus *Potamogeton* in the former broad sense counted about 1300 described species and interspecies hybrids, however, analysis of the herbarium samples allowed to identify only 69 to 90 species and 40 to 50 interspecies hybrids (Wiegleb, 1988; Wiegleb, Kaplan, 1998). In that sense, the genus was split in two subgenera *Potamogeton* L. and *Coleogeton* Rchb., species of the latter being distinguished by floating thickened leaves with long sheaths, hydrophilic (not anemophilic) inflorescences on long peduncles, commonly bearing widely separated whorls of flowers, as well as characteristic pollen structure (Sorsa, 1988). In the species of the subgenus *Potamogeton*, chromosome number varies from $2n = 14$ to $2n = 52$, while in the species of the subgenus *Coleogeton* it is $2n = 78$ (Les, 1983; Les, Haynes, 1996). The distinction of these two subgenera was supported by the complete absence of hybrids between species of these subgenera, while within both subgenera hybridisation is quite widespread (Tsvelev, 1996; Wiegleb, Kaplan, 1998). Due to a number of reasons, *Coleogeton* was proposed to be considered as a separate genus (Les, Haynes, 1996). Of the names suggested at the generic rank, *Coleogeton* and *Stuckenia* Börner, the latter is correct (Holub, 1997; Haynes et al., 1998a, b). At present, there exist two parallel versions of species names, for example, *Potamogeton pectinatus* L. is a synonym of *Stuckenia pectinata* (L.) Börner and so on. In this work we will consider the taxon in question at the generic level, as *Stuckenia*, and will use the specific names even if the cited authors used *Potamogeton*.

Studies of phylogenetic relationships in the family Potamogetonaceae, including representatives of the genera *Potamogeton* and *Stuckenia* using both plastid DNA markers (Iida et al., 2004) and 5S-NTS region of the nuclear genome (Lindqvist et al., 2006) showed that members of these genera cluster into two clearly distinguishable groups with high bootstrap support. This is in good accordance with conclusions made on the basis of morphologic characters. However, Q.D. Wang et al. (2007) found the latter region of *Potamogeton* and *Stuckenia* similar and did not support separation of the latter genus.

Taxonomy of the genera *Potamogeton* and *Stuckenia*, as well as other aquatic plants, is based mainly on anatomy and morphology of leaves, fruits, stems. The study of herbarium specimens showed that these characters are highly variable within a species. On the whole, all species of these two genera can be divided into three groups according to the degree of variability: (i) species with rather uniform morphological traits in spite of wide geographic range, their species attribution does not cause difficulty (*P. obtusifolius* Mert. et Koch, *P. praelongus* Wulf., *P. crispus* L.); (ii) species with a wide spectrum of variability within geographic range, so they can be sometimes misidentified as novel species or interspecies hybrids (*P. striatus* Ruiz. et Pav., *S. filiformis* Pers. and others); (iii) species with extremely high morphologic variability even in the same area so that their species attribution is always problematic (*S. pectinata* and others) (Wiegleb, 1988). Experimental cultivation of the clones of different species under controlled conditions (at different depths, different nutritional values of the substrate, different illumination) showed that morphological traits essential for the taxonomy vary with environmental changes and therefore cannot serve as reliable markers for species attribution (Kaplan, 2002). For example, herbarium samples collected in Central Russia, in the Caucasus, Middle Asia, Southern Siberia identified as *S. filiformis* upon re-examination turned out to be *S. pectinata* (Maemets, 1979). It has been noted that in the Arctic region of the European part of the former USSR, *S. filiformis* as well as interspecies hybrid *S. filiformis* Pers. × *S. vaginatus* Turcz. are often identified as *S. pectinata* (Maemets, 1979). Taxonomic revision of the *Stuckenia* species also revealed cases of erroneous species attribution in the group considered (Kaplan, 2008).

In the late 20th century biochemical markers, first of all, isozymes came into use for the study of the representatives of the genus *Potamogeton* and *Stuckenia*. These markers were used for the study of presumed interspecies hybrids considered as such on the basis of morphology (Hollingsworth et al., 1996; Kaplan, Stepanek, 2003). Later on, methods of molecular biology (RAPD, PCR RLFP, AFLP analyses) were employed for the study of interspecies hybrids (Whittall et al., 2004; Uehara et al., 2006; Kaplan et al., 2009).

The primary structure of the ITS1–ITS2 region of the nuclear genome is widely used for the study of phylogenetic relationships of a broad spectrum of organisms. At the same time, it can be utilized for species attribution of a given specimen when other approaches are inapplicable or complicated (Kress et al., 2005; Fazekas et al., 2012). This is the principle of the method of DNA barcoding of living organisms. However, this approach has certain limitations which should be considered in its practical applications and which are widely debated in literature (Shneyer, Rodionov, 2019). Since a large body of biodiversity remains poorly studied, primary structure

of the ITS1–ITS2 region allows to make a conjecture about existence of new species but not isolate and describe them (Desalle, 2006).

The aim of the present work was to study the applicability of the primary structure of the ITS1–ITS2 region for species attribution of a number of samples of the genus *Stuckenia*, classified on the basis of morphology as *S. pectinata*, *S. vaginata*, *S. macrocarpa* (Dobroch.) and *S. chakassiensis* (Kashina) Volobaev.

Material and methods

The present study is based on sequences of both reliably identified species present in Gene Bank at the moment of this study and original sequences from 28 plant specimens. Of these, 7 were identified as *S. pectinata*, 7 as *S. vaginata*, 9 as *S. macrocarpa* and 5 as *S. chakassiensis* on the basis of morphology (Suppl. Material¹). Plant material has been provided by L.M. Kipriyanova (Institute for Water and Environmental Problems of Siberian Branch of the Russian Academy of Sciences, Novosibirsk department, Russia). DNA was extracted from dry (herbarium) material or fixed and stored in ethanol. DNA extraction was performed with the use of 2x CTAB buffer as described by S.O. Rogers and A.J. Bendich with modifications (1985). Plant tissue (0.02–0.05 g of dry or 0.2–0.3 g of stored in ethanol) was thoroughly grinded in a mortar in the presence of 0.05 g aluminium oxide and 1 ml of extraction buffer freshly prepared before the extraction procedure dissolving 0.03 g polyethylene glycol 6000 and 0.05 g dithiothreitol in 1 ml 2x CTAB (2 % CTAB, 1.4M NaCl, 0.1M TRIS pH = 8.0, 20 mM EDTA). Homogenate was transferred to 2 ml tubes and incubated for 30 min at 75 °C. Then, 1 ml dichloromethane was added to each tube and thoroughly mixed for 10 min, the tubes were centrifuged for 10 min at 6708 x g. The supernatant was transferred to a fresh tube and added with 0.2 volumes of 5x CTAB (5 % CTAB, 350 mM EDTA), mixed and incubated for 10 min at 65 °C. Then each tube was added with 1 ml dichloromethane, mixed for 10 min and centrifuged as described above. The supernatant was transferred to fresh tubes and DNA was precipitated adding equal volume of isopropanol, mixing and keeping at –20 °C for 1 h or more. Nucleic acids were precipitated by centrifuging as described above, washed twice with 70 % ethanol, dried and resuspended in 50 µl deionised water. For Polymerase Chain Reaction (PCR) 10-fold dilution (1 part nucleic acid solution: 9 parts of water) was used.

PCR reaction was performed in a volume of 20 µl with 2 µl of 10x ammonium-sulphate buffer, 2 µl of 25 mM MgCl₂, 0.2 µl of the Taq polymerase (5 U/µl), 0.15 µl BSA (10 mg/ml), 1 µl of forward and reverse primers (10 pM) each, and 2 µl of diluted DNA. Concentration of dNTPs in the reaction mixture was 0.2 mM each. PCR reaction was held under following conditions: initial denaturation 95 °C – 3 min; then 38 cycles including: denaturation at 94 °C – 30 s, primer annealing at 58 °C – 30 s, elongation at 72 °C – 60 s; terminal elongation at 72 °C – 5 min. To amplify the ITS1–ITS2 region, ITS-5m (5'-GGAAGGAGAAGTCGTAACAAGG) and ITS-4 (5'-TCCTCCGCTTATTGATATGC) primers were used (Sang et al., 1995).

For the sequencing reaction, the same primers were used as for amplification. In some cases, when the use of the primers ITS-5m and ITS-4 failed to produce chromatograms of the suitable quality, specially designed sequencing primers were used: seq-ITS-F (5'-GATGACTCTCGGCAACGG ATA) and seq-ITS-R (5'-CTCGATGGTTCACGGGATTCT). Sanger sequencing reaction was performed with the use of ABI PRISM® BigDye™ Terminator v3.1 Ready Reaction Cycle Sequencing Kit. Determination of the primary structure of the resulting products was done at the SB RAS Genomics Core Facility (Novosibirsk). Nucleotide sequences obtained in this study were deposited in GenBank under accession numbers MH427614 to MH427641.

Besides, sequences HE613425, HE613426, HE613427, HE613428, HE613433, HE613434, KF270926, KF270927, KF270928 and KF270929 were taken from public databases.

Sequences were aligned by ClustalW program incorporated into Mega 5 package (Thompson et al., 1994; Tamura et al., 2011). Estimations of pairwise divergence between sequences were conducted in MEGA 5 (Tamura et al., 2011). The trees were constructed by the Maximum Likelihood method based on the Tamura–Nei model by means of MEGA 5 package (Tamura, Nei, 1993). Numbers at the nodes represent bootstrap values as percentages out of 1000 replicates and are shown only for values greater than 50 %.

Results

Study of applicability of the ITS1–ITS2 region for species identification on the basis of sequences from public databases

To make sure that data on the primary structure of the ITS1–ITS2 region are applicable for species identification of the representatives of the genus *Stuckenia*, the following sequences were analysed: HE613433, HE613434, KF270928, KF270929 referred to as belonging to *S. vaginata*, and KF270926, HE613427, KF270927 HE613425, HE613426 HE613428, attributed to *S. pectinata* (McMullan et al., 2011; Kaplan et al., 2013). These sequences were obtained by two independent research teams who sequenced the ITS1–ITS2 region both in *S. pectinata*, and *S. vaginata*. The entries beginning with “KF” come from the Institute of Botany, Academy of Sciences of Czech Republic, and those beginning with “HE” were obtained by a research group from Great Britain. An analysis of the origin of the specimen studied showed that they were collected in rather distant geographic points. Namely, *S. vaginata* specimens were sampled in the Bothnia Bay near the coast of Sweden and Finland, in the Irkutsk region (Russia), and in the USA. *S. pectinata* samples were collected in USA, Netherlands, Great Britain, Italy, Russia and India.

Pairwise comparison of the above mentioned sequences showed that intraspecies differences within *S. pectinata* revealed in one of the studies coincided with those revealed in the other study (Table 1). Some ITS1–ITS2 sequences of *S. pectinata* obtained in different studies were identical. Pairwise comparison of ITS1–ITS2 sequence of *S. vaginata* demonstrated similar results. Thus, the data on the ITS1–ITS2 primary structure of a species obtained in one study are supported by those of the other study.

¹ Supplementary Material is available in the online version of the paper: <https://vavilovj-icg.ru/download/pict-2023-27/appx4.pdf>

Table 1. Matrix of pairwise uncorrected p-distances of the concatenated sequences of the ITS1–ITS2 region of *S. pectinata* and *S. vaginata* of different provenance, taken from GenBank

No.	GenBank ID	Species	1	2	3	4	5	6	7	8	9	10
1	HE613425	<i>S. pectinata</i>										
2	HE613427	<i>S. pectinata</i>	0.010									
3	HE613428	<i>S. pectinata</i>	0.000	0.010								
4	HE613426	<i>S. pectinata</i>	0.000	0.010	0.000							
5	KF270926	<i>S. pectinata</i>	0.010	0.000	0.010	0.010						
6	KF270927	<i>S. pectinata</i>	0.000	0.010	0.000	0.000	0.010					
7	HE613433	<i>S. vaginata</i>	0.039	0.029	0.039	0.039	0.029	0.040				
8	HE613434	<i>S. vaginata</i>	0.039	0.029	0.039	0.039	0.029	0.040	0.000			
9	KF270928	<i>S. vaginata</i>	0.039	0.029	0.039	0.039	0.029	0.039	0.003	0.003		
10	KF270929	<i>S. vaginata</i>	0.038	0.028	0.038	0.038	0.028	0.037	0.000	0.000	0.001	

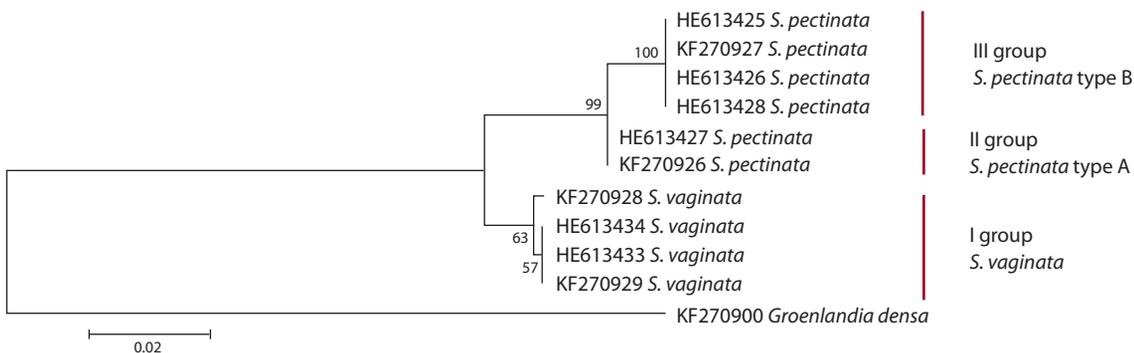


Fig. 1. Maximum likelihood phylogenetic tree constructed on the basis of the primary structure of the ITS1–ITS2 region of the representatives of *S. vaginata* and *S. pectinata* from public databases.

The sequences taken from public databases were used to construct a phylogenetic tree (Fig. 1). This tree contains three rather well supported groups. The entire group I is formed by the sequences referring to *S. vaginata* while groups II and III are formed by the sequences referring to *S. pectinata*.

All specimens belonging to the groups II and III were identified as *S. pectinata*, but had different structure of the ITS1–ITS2 region, this difference being supported by high bootstrap values. To distinguish between the genotypes within *S. pectinata* samples, those from group II were denoted as *S. pectinata* type A (*S. pectinata* genotype A), and those from group III – as *S. pectinata* type B (*S. pectinata* genotype B). Table 2 presents the alignments of the sequences of the three indicated groups. It can be seen that *S. pectinata* type A is represented by two haplotypes differing by a deletion at the position 136 and 138 (KF270926). Other samples of *S. vaginata* and *S. pectinata* type B have no such deletion. This haplotype with the deletion was not found in the samples studied in the present work, so this unique variant is not considered further. As seen from the alignments, *S. vaginata* differs from *S. pectinata* type A and *S. pectinata* type B by three indels (two one-nucleotide and one nine-nucleotide) and several nucleotide substitutions (sixteen one-nucleotide and two two-nucleotide). The differences between *S. pectinata* type A and *S. pectinata* type B are smaller and consist of five one-nucleotide and one two-nucleotide substitutions. Thus,

the primary structure of the ITS1–ITS2 region not only allows to identify known species *S. vaginata* and *S. pectinata* but also reveals the hitherto unknown type dichotomy of the latter for A and B types.

Sequencing and analysis of the ITS1–ITS2 region in the representatives of the genus *Stuckenia*

Primary structure of the ITS1–ITS2 region was determined in 28 samples (see Suppl. Material). Also, three sequences from public databases were involved into analysis to provide a reference: HE613427 representing “*S. pectinata* type A”, HE613428 representing “*S. pectinata* type B” and HE613434 representing “*S. vaginata*”. These sequences served as references to make species attribution of the sequenced samples to *S. vaginata*, *S. pectinata* type A or *S. pectinata* type B according to the primary structure of the ITS1–ITS2 region.

The so formed data array was used to reconstruct a phylogenetic tree (Fig. 2). The sequences formed three groups with good bootstrap support. The first group was formed by 2 specimens (Nos. 313 and 315) with the ITS1–ITS2 region typical of *S. vaginata*. On the basis of morphology, they were also classified as *S. vaginata*. The second group was formed by 2 specimens (Nos. 183 and 303) with the ITS1–ITS2 region of *S. pectinata* type A. On the basis of morphology, they were also classified as *S. pectinata*. The third group could be separated into two subgroups – III–I and III–II. The subgroup III–I

Table 2. Alignment of sequences of the ITS1–ITS2 region of *S. pectinata* and *S. vaginata* from public databases

Species	GenBank ID		Variable positions																		
			36	81	85	105	121	136	151	208	398	498	25	58	84	90	106	123	137	184	236
<i>S. pectinata</i> type B	KF270927	(1)	..C..M..A..G..CC..A..GC..G.T.C..G.T..T..A..K..G..T..T..C..																		
<i>S. pectinata</i> type B	HE613425	(17)	..C..C..A..G..CC..A..GC..G.T.C..G.T..T..A..G..G..T..T..C..																		
<i>S. pectinata</i> type B	HE613426	(17)	..C..C..A..G..CC..A..GC..G.T.C..G.T..T..A..G..G..T..T..C..																		
<i>S. pectinata</i> type B	HE613428	(17)	..C..C..A..G..CC..A..GC..G.T.C..G.T..T..A..G..G..T..T..C..																		
<i>S. pectinata</i> type A	KF270926	(13)	..C..C..A..C..CG..A..GC..G.T.C..-..T..A..G..G..C..T..C..																		
<i>S. pectinata</i> type A	HE613427	(17)	..C..C..A..C..CG..A..GC..G.T.C..G.T..T..A..G..G..C..T..C..																		
<i>S. vaginata</i>	KF270928	(7)	..A..C..T..C..TG..G..TT..C.A.G..GGT..C..T..G..A..C..G..T..																		
<i>S. vaginata</i>	KF270929	(4)	..A..C..T..C..TG..A..TT..C.A.G..GGT..C..T..G..A..C..G..T..																		
<i>S. vaginata</i>	HE613433	(17)	..A..C..T..C..TG..A..TT..C.A.G..GGT..C..T..G..A..C..G..T..																		
<i>S. vaginata</i>	HE613434	(17)	..A..C..T..C..TG..A..TT..C.A.G..GGT..C..T..G..A..C..G..T..																		

Species	GenBank ID		Variable positions									
			505	568	593	622	636	650	677	559	584	621
<i>S. pectinata</i> type B	KF270927		..C..A..A..A..G..AA.C..AA..G-----A..T....									
<i>S. pectinata</i> type B	HE613425		..C..A..A..A..G..AA.C..AA..G-----A..T....									
<i>S. pectinata</i> type B	HE613426		..C..A..A..A..G..AA.C..AA..G-----A..T....									
<i>S. pectinata</i> type B	HE613428		..C..A..A..A..G..AA.C..AA..G-----A..T....									
<i>S. pectinata</i> type A	KF270926		..C..G..A..A..G..TC.G..AA..G-----A..T....									
<i>S. pectinata</i> type A	HE613427		..C..G..A..A..G..TC.G..AA..G-----A..T....									
<i>S. vaginata</i>	KF270928		..-..G..C..T..G..TC.G..TC..ATTGTGGATC..T..G....									
<i>S. vaginata</i>	KF270929		..-..G..C..T..K..TC.G..TC..ATTGTGGATC..T..G....									
<i>S. vaginata</i>	HE613433		..-..G..C..T..T..TC.G..TC..ATTGTGGATC..T..G....									
<i>S. vaginata</i>	HE613434	(512)	..-..G..C..T..T..TC.G..TC..ATTGTGGATC..T..G....									

Note. Positions differing between *S. vaginata* and *S. pectinata* are marked with black, between *S. pectinata* type A and *S. pectinata* type B are marked with grey.

was formed by 18 specimens with the ITS1–ITS2 region of *S. pectinata* type B. On the basis of morphology, 9 specimens (Nos. 2, 15, 16, 46, 47, 52, 94, 95 and 117) were classified as *S. macrocarpa*, 5 – as *S. pectinata* (Nos. 4, 300, 301, 302 and 330), 4 – as *S. vaginata* (Nos. 3, 63, 93 and 100). The subgroup III–II was formed by 6 specimens. On the basis of morphology, 5 specimens (Nos. 1, 105, 317, 321 and 323) were classified as *S. chakassiensis*, 1 – as *S. vaginata* (No. 314).

Alignment of the sequences belonging to the subgroups III–I and III–II is given in Table 3. According to the nucleotides in the positions 102–103, the sequences of the subgroup III–I formed three clearly distinguishable batches. The first batch contained sequences from the samples Nos. 3, 63, 93, 100, 301, 302, 330 and the reference sequence HE613428. The second batch contained sequences from the samples Nos. 2, 46, 47, 52, 94, 95, 117, and 300. The third batch contained sequences from the samples Nos. 4, 15 and 16.

The samples Nos. 3, 63, 93 and 100 belonging to the first of the mentioned batch were identified as *S. vaginata*, the samples No. 302 and 301 with the same primary structure of the ITS1–ITS2 region were identified as *S. pectinata* type B. Identical primary structure was shared by the reference se-

quence HE613428. The sequence of the sample No. 330 differed from that of the above mentioned samples by a number of polymorphic positions (seen in the sequencing chromatograms as superimposed peaks) not found in the other samples, and in the positions 102–103 it had the same nucleotide composition as the samples Nos. 3, 63, 93, 100, 301, 302 and HE613428. Thus, this batch was formed by the samples identified on the basis of morphology as *S. vaginata* and *S. pectinata*, although on the basis of the primary structure of the ITS1–ITS2 region these samples should be classified as belonging to *S. pectinata* type B. Such discrepancy of the morphologic and molecular data could result from misidentification. Earlier in the comparative investigation of *S. vaginata* and *S. pectinata*, it was shown that *S. pectinata* had two recognition sites for the restrictase *CfoI* (GCGC) in the ITS1–ITS2 region, while *S. vaginata* had only one such site (King et al., 2001). In our data array, the samples Nos. 3, 63, 93, 100 and 314 have two recognition sites, which is typical of *S. pectinata*. Therefore, it is highly probable that species attribution of the samples Nos. 3, 63, 93 and 100 was erroneous and they should be considered as *S. pectinata*. Thus, it may be stated with a high degree of confidence that the first batch of the primary struc-

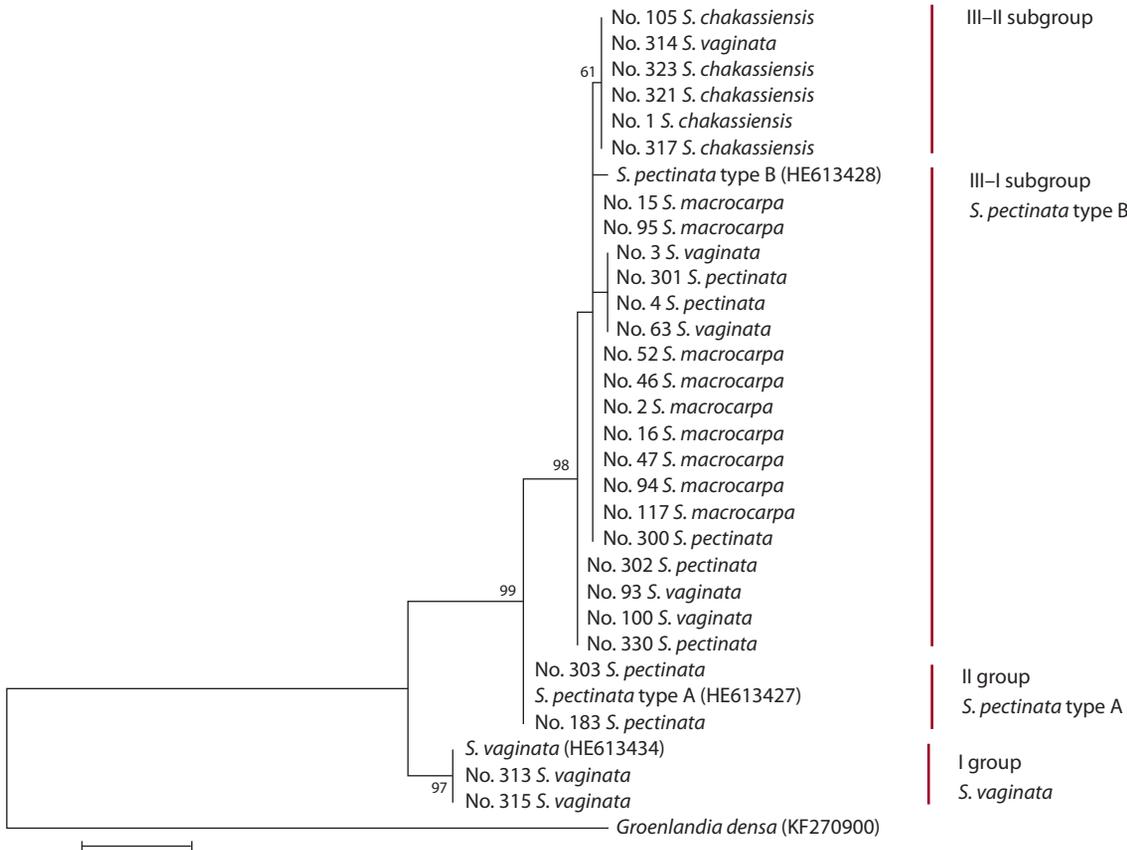


Fig. 2. Maximum likelihood tree reconstructed on the basis of the primary structure of the ITS1–ITS2 region in the samples with species attribution according to morphology (in combinations with the generic names as in the data source).

ture of the ITS1–ITS2 region is composed by the samples of *S. pectinata*, or more precisely, *S. pectinata* type B.

Out of 8 samples of the second batch, 7 (Nos. 2, 46, 47, 52, 94, 95 and 117), were identified on the basis of morphology as *S. macrocarpa* and 1 (No. 300) – as *S. pectinata*. Since all sequences of this batch are identical, it is probable that the sample No. 300 was misidentified, and the second batch is composed by the samples of *S. macrocarpa*.

The sequences composing the third batch (samples Nos. 4, 15 and 16) were identical and characterized by heterogeneity for the positions 102 and 103. Position 102 contained both G, as in *S. pectinata* type B, and T, as in *S. macrocarpa*. Position 103 contained both C, as in *S. pectinata* type B, and A, as in *S. macrocarpa*. Thus, these samples may be considered as interspecies hybrids between *S. pectinata* type B and *S. macrocarpa*. According to morphologic traits, the samples Nos. 15 and 16 were identified as *S. macrocarpa*, and the sample No. 4 – as *S. pectinata*.

Six samples forming the subgroup III–II had identical ITS1–ITS2 region, 5 of them, according to morphology, had been identified as *S. chakassiensis*, and one – as *S. vaginata* (sample No. 314). However, it is highly improbable that it really represents *S. vaginata*, since the structure of its ITS1–ITS2 region, in particular, the presence of two recognition sites for the *CfoI* restrictase is not typical of *S. vaginata*. If this sample is excluded from consideration, the subgroup III–II is constituted by the samples identified on the basis of morphology as *P. chakassiensis*. Therefore, it can be supposed that

this entire subgroup is formed by representatives of the latter species. As seen from the alignments of the sequences of the ITS1–ITS2 region of the samples belonging to the subgroups III–I and III–II (see Table 3), the only difference between the mentioned subgroups consists in one nucleotide substitution, T/C in the position 524. Thus, according to the nucleotide in this position, samples of *S. chakassiensis* can be unequivocally identified by the primary structure of the ITS1–ITS2 region.

Discussion

As noted above, the growing of representatives of the genus *Potamogeton* and *Stuckenia* under different ecologic conditions showed that a large part of morphologic traits basic for the taxonomy of the genus varies along with the growth conditions (Kaplan, 2002). Since the majority of investigators describe new taxa based solely on morphologic traits without any study as to the stability of their manifestation in different environments, ecological modifications were often described as new species (Kaplan, 2002). Thus, morphologic traits turned to be not too reliable for species identification in pondweeds, and there arises a need for developing markers applicable for species identification in the genus *Stuckenia*. Primary structure of the ITS1–ITS2 region is suggested here as a suitable marker for this purpose.

An analysis of the sequences of the ITS1–ITS2 region of *S. vaginata* and *S. pectinata* from public databases showed that these sequences differed, thus making the primary structure of this region a promising marker with respect to identification of

Table 3. Alignment of sequences of the ITS1–ITS2 region of the samples from III–I and III–II subgroups

Subgroup	Isolate	Sample		Variable positions				Species attribution according to ITS1–ITS2
				102	386	547	609	
				103	524	610		
				612				
III–II	No. 1	<i>S. chakassiensis</i>	(24)	..TA...T...T...A...AA.C...				<i>S. chakassiensis</i>
	No. 105	<i>S. chakassiensis</i>	(1)	..TA...T...T...A...AA.C...				
	No. 314	<i>S. vaginata</i>	(24)	..TA...T...T...A...AA.C...				
	No. 317	<i>S. chakassiensis</i>	(24)	..TA...T...T...A...AA.C...				
	No. 321	<i>S. chakassiensis</i>	(24)	..TA...T...T...A...AA.C...				
	No. 323	<i>S. chakassiensis</i>	(24)	..TA...T...T...A...AA.C...				
III–I		<i>S. pectinata</i> type B (HE613428)	(28)	..GC...T...C...A...AA.C...				<i>S. pectinata</i> type B
	No. 3	<i>S. vaginata</i>	(24)	..GC...T...C...A...AA.C...				
	No. 63	<i>S. vaginata</i>	(24)	..GC...T...C...A...AA.C...				
	No. 93	<i>S. vaginata</i>	(24)	..GC...T...C...A...AA.C...				
	No. 100	<i>S. vaginata</i>	(24)	..GC...T...C...A...AA.C...				
	No. 301	<i>S. pectinata</i>	(24)	..GC...T...C...A...AA.C...				
	No. 302	<i>S. pectinata</i>	(24)	..GC...T...C...A...AA.C...				
	No. 330	<i>S. pectinata</i>	(24)	..GC...Y...C...A...WM.S...				
	No. 2	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				<i>S. macrocarpa</i>
	No. 46	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				
	No. 47	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				
	No. 52	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				
	No. 94	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				
	No. 95	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				
	No. 117	<i>S. macrocarpa</i>	(24)	..TA...T...C...A...AA.C...				
	No. 300	<i>S. pectinata</i>	(24)	..TA...T...C...A...AA.C...				
	No. 4	<i>S. pectinata</i>	(24)	..KM...Y...C...R...AA.C...				F1 hybrid between <i>S. macrocarpa</i> and <i>S. pectinata</i>
	No. 15	<i>S. macrocarpa</i>	(24)	..KM...T...C...A...AA.C...				
	No. 16	<i>S. macrocarpa</i>	(24)	..KM...T...C...A...AA.C...				

Note. Positions differing between *S. pectinata* type B and *P. macrocarpus* are marked with black, differing *P. chakassiensis* from *S. pectinata* type B and *P. macrocarpus* are marked with grey.

the mentioned species. Moreover, two genotypes of this region were revealed in *S. pectinata*, designated as *S. pectinata* type A and *S. pectinata* type B.

These results were used to analyze the region ITS1–ITS2 in 28 samples of the genus *Stuckenia*, identified on the basis of morphology as *S. vaginata*, *S. pectinata*, *S. chakassiensis* and *S. macrocarpa*. Out of seven samples classified as *S. vaginata*, only two could be unequivocally attributed to this species according to the primary structure of their ITS1–ITS2 region, while the other five samples according to the primary structure of the studied region fit *S. pectinata* more. Such discrepancy of species attribution made on the basis of morphologic traits and molecular data may be due to original misidentification of the studied samples. Two samples classified as *S. vaginata* have only one recognition site for the *CfoI* restrictase in their ITS1–ITS2 region that is typical of *S. vaginata*, and rest of the samples classified as *S. vaginata* have two such recognition sites, which is typical of *S. pectinata*. This example demonstrates that species attribution made solely on the basis of morphology does not guarantee correct species identifica-

tion, for the purpose of which other approaches should be supplemented, in particular, the use of molecular data on the ITS1–ITS2 region appears to be appropriate.

The botanical assessment of representatives of the genus *Stuckenia* showed that *S. pectinata* is a polymorphic species and even an opinion that it might be a composite species was put forward, although no evidence was provided (Maemets, 1979; Kashina, 1988). Our result revealing the existence of two groups of the primary structure of the ITS1–ITS2 region in *S. pectinata* (*S. pectinata* type A and *S. pectinata* type B) favours the view that *S. pectinata* includes several species that are indistinguishable or hardly distinguishable at the level of morphology but clearly differ at the level of the ITS1–ITS2 region. Also, it should be noted that no intermediate forms between A and B type of *S. pectinata* were revealed. The existence of two cryptic species hidden under the name *S. pectinata* supposed on the basis of the structure of the ITS1–ITS2 region is supported by literature data coming from the RAPD analysis of *S. pectinata* samples of different origin (Mader et al., 1998). The RAPD spectra of the samples from the Pechora

River delta differed from those of the samples collected in Italy, Germany, Poland, France and the Saint-Petersburg surroundings. The samples from Spain and Egypt also differed from each other as well as from the above mentioned ones. P.A. Volkova et al. (2017), who studied the same ITS1–ITS2 region, revealed its uniformity in Europe but high differentiation in southern Siberia. All this allows to suppose that there exist at least two and perhaps more “forms” of *S. pectinata*. Their phylogenetic relationships are clear, but their taxonomic status is obscure. Whether they represent cryptic species, subspecies or merely intraspecific polymorphism requires further investigation.

Some of the samples analysed in the present work have been classified as *S. chakassiensis*. In the phylogenetic tree (see Fig. 2), these samples formed a separate subgroup. These samples were categorized as a subgroup because the bootstrap value did not permit to consider them as a separate group. The primary structure of the ITS1–ITS2 region was identical in all representatives of the subgroup, it was also identical in a sample classified on the basis of morphology as *S. vaginata* and discarded as such according to the structure of the ITS1–ITS2 region. This allows to suppose that the species *S. chakassiensis* really exists although investigators can experience difficulties in its identification. At the same time, the data on the primary structure of the ITS1–ITS2 region allows to identify this species. P.A. Volkova et al. (2017) studied the same ITS1–ITS2 region and also the plastid *rpl32-trnL* spacer and found no correspondence between the sequence data and diagnostic morphological character of *S. chakassiensis* (which is not as convincing, being the presence of sclerenchyma strands in leaves). Actually, our data evidence for the same with respect to the above mentioned specimen identified as *S. vaginatus*. Taken together, the results of P.A. Volkova et al. (2017) and of the present study can be interpreted such that the species *S. chakassiensis* does exist but its only diagnostic morphological character proposed is unreliable and may lead to misidentifications.

The existence of *S. chakassiensis* and its difference from *S. pectinata* is indirectly supported by the data coming from the study of metal contents in pondweeds and common reed (*Phragmites australis* Trin. ex Steud) from brackish lake Shira and freshwater reservoir Bugach (Ivanova et al., 2015). Differences in the contents of metals in the plants from different water bodies were shown for pondweeds but not for the common reed. At the same time, pondweeds collected in a desalinated part of Lake Shira did not differ from those collected in more salty water of the same lake. These paradoxical results can be easily interpreted if to suppose that the pondweed from the Shira Lake, with accordance to our data, belonged to *S. chakassiensis*, while the pondweed from the Bugach reservoir represented *S. pectinata*, that is, in fact, two species have been mixed. One of them grows in salt water, while the other in fresh or brackish water. In contrast to the pondweeds, common reed is adapted to the growth in fresh and brackish water as well as salt water, both studied lakes harbor the same species, the populations of which do not differ in metal contents. It should be noted that mineralization in Lake Shira is 15.9 g/l (Guseva et al., 2012). This is above the limit of the level of salinity that plants of *S. pectinata* withstand, over which their death begins (Coffey, 2001).

Especially interesting are the samples from the subgroup III–I classified on the basis of morphology as *S. vaginata*, *S. pectinata* and *S. macrocarpa*. According to the primary structure of the ITS1–ITS2 region, they can be categorized into three batches. The first of them is composed by samples that harbor the GC dinucleotide in the positions 102–103 (see Table 3); this is a characteristic of the reference sequence HE613428, *S. pectinata* type B. The sequences from the second batch harbor TA in these positions; the majority of the samples from this batch were classified as *S. macrocarpa* on the basis of morphology. The last batch is composed by the samples which harbor both G and T in the position 102 and both C and A in the position 103. That is, the third batch can be obtained by a mixture of any sequence from the first batch with any sequence from the second one. This allows to suppose on the basis of the molecular data on the primary structure of the ITS1–ITS2 region that the III–I subgroup is composed by the samples of *S. pectinata* type B, *S. macrocarpa* and interspecific hybrids between *S. pectinata* type B and *S. macrocarpa*. This can be experimentally tested by comparative analysis of morphology and anatomy of presumed original species and their hybrids. If molecular data find support in the morphology, this may be interpreted as evidence for the existence of *S. macrocarpa* as a separate species. Earlier, P.A. Volkova et al. (2017) also obtained somewhat confusing molecular results with respect to three analysed specimens morphologically identified as *S. macrocarpa*: they shared a specific ITS1–ITS2 haplotype but had a haplotype of the *rpl32-trnL* spacer found also in three other species.

In conclusion, the data obtained in the present work demonstrate the applicability of the primary structure of the ITS1–ITS2 region for species attribution and revealing species misidentification in the genus *Stuckenia*, which in some cases may be more reliable than morphological data.

References

- Coffey W.W. The feasibility of submerged macrophyte re-establishment in Kaituna Lagoon, Lake Ellesmere (Te Waihora). A thesis submitted in partial fulfilment of the requirements for the degree of Master of Applied Science in Resource Management/Ecology at Lincoln University. Lincoln University, N.Z. 2001. <https://researcharchive.lincoln.ac.nz/handle/10182/2271>
- Desalle R. Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff. *Conserv. Biol.* 2006;20(5): 1545–1547. DOI 10.1111/j.1523-1739.2006.00543.x.
- Fant J.B., Kamau E., Preston C.D. Chloroplast evidence for the multiple origins of the hybrid *Potamogeton x sudermanicus* Hagstr. *Aquat. Bot.* 2003;75(4):351–356. DOI 10.1016/S0304-3770(03)00004-4.
- Fant J.B., Kamau E., Preston C.D. Chloroplast evidence for the multiple origins of the hybrid *Potamogeton x fluitans*. *Aquat. Bot.* 2005; 83(2):154–160. DOI 10.1016/j.aquabot.2005.06.004.
- Fazekas A.J., Kuzmina M.L., Newmaster S.G., Hollingsworth P.M. DNA barcoding methods for land plants. *Methods Mol. Biol.* 2012; 858:223–252. DOI 10.1007/978-1-61779-591-6_11.
- Guseva N.V., Kopylova Yu.G., Khvashchevskaya A.A., Smetanina I.V. Chemical content of salt lakes of the Severo-Minusinsk depression, Khakasiya. *Bull. Tomsk Polytech. Univ.* 2012;321(1):163–168. (in Russian)
- Haynes R.R., Les D.H., Kral M. Two new combinations in *Stuckenia*, the correct name for *Coleogeton* (Potamogetonaceae). *Novon.* 1998a;8(3):241. DOI 10.2307/3392010.
- Haynes R.R., Les D.H., Holm-Nielsen L.B.I. Potamogetonaceae. In: Kubitzki K. (Ed.). The families and genera of vascular plants, Flowering Plants: Monocotyledons: Alismatanae and Commelinanae

- (except Gramineae). Vol. IV. Berlin: Springer-Verlag, 1998b;408-414. DOI 10.1007/978-3-662-03531-3_39.
- Hollingsworth P.M., Preston C.D., Gornall R.J. Euploid and aneuploid evolution in *Potamogeton* (Potamogetonaceae): a factual basis for interpretation. *Aquat. Bot.* 1998;60(4):337-358. DOI 10.1016/S0304-3770(97)00101-0.
- Hollingsworth P.M., Preston C.D., Gornall R.J. Isozyme evidence for the parentage and multiple origins of *Potamogeton* × *suecicus* (*P. pectinatus* × *P. filiformis*, Potamogetonaceae). *Plant Syst. Evol.* 1996;202(3):219-232.
- Holub J. *Stuckenia* Börner 1912 – the correct name for *Coleogeton* (Potamogetonaceae). *Preslia.* 1997;69,361-366.
- Iida S., Kosuge K., Kadono Y. Molecular phylogeny of Japanese *Potamogeton* species in light of noncoding chloroplast sequences. *Aquat. Bot.* 2004;80(2):115-127. DOI 10.1016/j.aquabot.2004.08.005.
- Ivanova Ye.A., Anishchenko O.V., Zuyev I.V., Avramov A.P. Content of metals in *Phragmites australis* Trin. ex Steud and *Potamogeton pectinatus* L. from water bodies of different salinity. *J. Sib. Fed. Univ. Biol.* 2015;8(3):347-361. DOI 10.17516/1997-1389-2015-8-3-347-361. (in Russian)
- Kaplan Z. A taxonomic revision of *Stuckenia* (Potamogetonaceae) in Asia, with notes on the diversity and variation of the genus on a worldwide scale. *Folia Geobot.* 2008;43(2):159-234. DOI 10.1007/s12224-008-9010-0.
- Kaplan Z., Fehrer J., Hellquist C.B. New hybrid combinations revealed by molecular analysis: the unknown side of North American pondweed diversity (*Potamogeton*). *Syst. Bot.* 2009;34(4):625-642. DOI 10.1600/036364409790139745.
- Kaplan Z. Hybridization of *Potamogeton* species in the Czech Republic: diversity, distribution, temporal trends and habitat preferences. *Preslia.* 2010;82(3):261-287.
- Kaplan Z., Jarolimova V., Fehrer J. Revision of chromosome numbers of Potamogetonaceae: a new basis for taxonomic and evolutionary implications. *Preslia.* 2013;85:421-482.
- Kaplan Z. Phenotypic plasticity in *Potamogeton* (Potamogetonaceae). *Folia Geobot.* 2002;37(2):141-170. DOI 10.1007/BF02804229.
- Kaplan Z., Stepanek J. Genetic variation within and between populations of *Potamogeton pusillus* agg. *J. Plant Syst. Evol.* 2003;239(1):95-112. DOI 10.1007/s00606-002-0252-7.
- Kashina L.I. Potamogetonaceae. In: Krasnoborov I.M. (Ed). Flora Sibiri. *Lycopodiaceae – Hydrocharitaceae*. Novosibirsk: Nauka, 1988;93-105. (in Russian)
- King R.A., Gornall R.J., Preston C.D., Croft J.M. Molecular confirmation of *Potamogeton* × *bottnicus* (*P. pectinatus* × *P. vaginatus*, Potamogetonaceae) in Britain. *Bot. J. Lin. Soc.* 2001;135(1):67-70. DOI 10.1111/j.1095-8339.2001.tb02370.x.
- Kress W.J., Wurdack K.J., Zimmer E.A., Weigt L.A., Janzen D.H. Use of DNA barcodes to identify flowering plants. *PNAS.* 2005;102(23):8369-8374. DOI 10.1073/pnas.0503123102.
- Les D.H. Taxonomic implications of aneuploidy and polyploidy in *Potamogeton* (Potamogetonaceae). *Rhodora.* 1983;85(843):301-323.
- Les D.H., Haynes R.R. *Coleogeton* (Potamogetonaceae), a new genus of pondweeds. *Novon.* 1996;6:389-391. DOI 10.2307/3392046.
- Lindqvist C., De Laet J., Haynes R.R., Aagesen L., Keener B.R., Albert V.A. Molecular phylogenetics of an aquatic plant lineage, Potamogetonaceae. *Cladistics.* 2006;22(6):568-588. DOI 10.1111/j.1096-0031.2006.00124.x.
- Mader E., van Vierssen W., Schwenk K. Clonal diversity in the submerged macrophyte *Potamogeton pectinatus* L. inferred from nuclear and cytoplasmic variation. *Aquat. Bot.* 1998;62(3):147-160. DOI 10.1016/S0304-3770(98)00096-5.
- McMullan J.J., Gornall R.J., Preston C. ITS rDNA polymorphism among species and hybrids of *Potamogeton* subgenus *Coleogeton* (Potamogetonaceae) in north-western Europe. *New J. Bot.* 2011;1(2):111-115. DOI 10.1179/204234811X13194453002788.
- Myaemets A.A. Potamogeton. In: Fedorova A.A. (Ed.). Flora Flora of the European Part of the USSR. Vol. 4. Leningrad: Nauka, 1979; 176-192. (in Russian)
- Rogers S.O., Bendich A.J. Extraction of DNA from milligram amounts of fresh herbarium and mummified plant tissues. *Plant Mol. Biol.* 1985;5(2):69-76. DOI 10.1007/BF00020088.
- Sang T., Crawford D.J., Stuessy T.F. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *PNAS.* 1995;92(15):6813-6817. DOI 10.1073/pnas.92.15.6813.
- Shneyer V.S., Rodionov A.V. Plant DNA barcodes. *Biol. Bull. Rev.* 2019;9(4):295-3900. DOI 10.1134/S207908641904008X.
- Sorsa P. Pollen morphology of *Potamogeton* and *Groenlandia* (Potamogetonaceae) and its taxonomic significance. *Ann. Bot. Fenn.* 1988;25(2):179-199.
- Tamura K., Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 1993;10(3):512-526. DOI 10.1093/oxfordjournals.molbev.a040023.
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* 2011;28(10):2731-2739. DOI 10.1093/molbev/msr121.
- Thompson J.D., Higgins D.G., Gibson T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994;22(22):4673-4680. DOI 10.1093/nar/22.22.4673.
- Tsvelev N.N. On the species of the subgenus *Coleogeton* of the genus *Potamogeton* (Potamogetonaceae) in North-Western Russia. *Botanicheskii Zhurnal.* 1996;81(7):88-91. (in Russian)
- Uehara K., Tanaka N., Momohara A., Zhou Z.-K. Genetic diversity of an endangered aquatic plant. *Potamogeton lucens* subspecies *sini-cus*. *Aquat. Bot.* 2006;85(4):350-354. DOI 10.1016/j.aquabot.2006.06.008.
- Volkova P.A., Kipriyanova L.M., Maltseva S.Y., Bobrov A.A. In search of speciation: Diversification of *Stuckenia pectinata* sl (Potamogetonaceae) in southern Siberia (Asian Russia). *Aquat. Bot.* 2017;143:25-32. DOI 10.1016/j.aquabot.2017.07.003.
- Wang Q.D., Zhang T., Wang J.B. Phylogenetic relationships and hybrid origin of *Potamogeton* species (Potamogetonaceae) distributed in China: insights from the nuclear ribosomal internal transcribed spacer sequence (ITS). *Plant Syst. Evol.* 2007;267(1-4):65-78. DOI 10.1007/s00606-006-0499-5.
- Whittall J.B., Hellquist C.B., Schneider E.L., Hodges S.A. Cryptic species in an endangered pondweed community (*Potamogeton*, Potamogetonaceae) revealed by AFLP markers. *Am. J. Bot.* 2004;91(12):2022-2029. DOI 10.3732/ajb.91.12.2022.
- Wiegleb G. Notes on pondweeds – outlines for a monographical treatment of the genus *Potamogeton* L. *Fedded Repertorium.* 1988;99(7-8):249-266.
- Wiegleb G., Kaplan Z. An account of the species of *Potamogeton* L. (Potamogetonaceae). *Folia Geobot.* 1998;33(3):241-316. DOI 10.1007/BF03216205.

ORCID ID

A.V. Mglinets orcid.org/0000-0001-6040-6801

Acknowledgements. The authors are grateful to L.M. Kipriyanova for providing plant material and to V.S. Bodganova for the help with English. This work was supported by the grant of Russian Foundation for the Fundamental Research (RFFI) project No. 13-04-02055-a and the scientific project FWN-2022-0017.

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

Received February 10, 2022. Revised November 14, 2022. Accepted November 17, 2022.

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

A genogeographic study of the Kyrgyz mountain merino via microsatellite markers

A.B. Bekturov¹, Zh.T. Isakova²✉, V.N. Kipen³, T.Dzh. Chortonbaev¹, S.B. Mukeeva², S.K. Osmonaliev⁴, K.A. Aitbaev²

¹ Kyrgyz National Agrarian University named after K.I. Skryabin, Bishkek, Kyrgyz Republic

² Research Institute of Molecular Biology and Medicine, Bishkek, Kyrgyz Republic

³ Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus

⁴ Kyrgyz Research Institute of Animal Husbandry and Pastures, Sokuluk District, Kyrgyz Republic

✉ jainagul@mail.ru

Abstract. The aim was to ascertain the genetic and geographical structure of the Kyrgyz mountain merino (KMM). We analyzed DNA samples of 109 Kyrgyz mountain merino specimens, bred in three state breeding factories (STB), including "Orgochor" in the Issykul Province, "Katta-Taldyk" in the Osh Province and STB named after Luschikhin in the Talas Province. We identified 126 alleles in 12 microsatellite markers (*McM042*, *INRA006*, *McM527*, *ETH152*, *CSRD247*, *OarFCB20*, *INRA172*, *INRA063*, *MAF065*, *MAF214*, *INRA005*, *INRA023*). There were 6 to 16 alleles in each locus (mean 10.500 ± 0.957 alleles per locus). We identified 67 rare alleles (prevalence less than 5.0 %), which made up 53.2 % of all alleles found. The greatest number of rare alleles was found in STR-markers of *CSRD247*, *INRA023*, *INRA005*, *INRA006*, *MAF214* and *OarFCB20*. For each group, there were individual differences in the distribution of allele frequencies across all the STR loci studied. The most significant of them were as follows: with regard to the *McM042* locus, allele 87 was major in the TALAS and OSH groups (35.6 and 45.7 %, respectively), whereas allele 95 was major in the ISSYK-KUL group (36.2 %); allele 154 was major in all groups with regard to the *INRA172* locus, but it was 1.25 times less prevalent in the ISSYK-KUL and 1.66 times less prevalent in the OSH groups compared to TALAS (55.2 and 41.4 %, respectively), whereas alleles 156 and 158 were found only in the ISSYK-KUL group. Considering the *ETH152* locus, 186 allele prevalence in the TALAS group was 51.1 %, but allele 190 was also markedly prevalent in the ISSYK-KUL and OSH groups, 34.5 and 34.3 %, respectively. The genetic division of the studied groups of KMM (with K from 3 to 10) was homogeneous – the contribution of each subcluster was equivalent. The AMOVA analysis revealed that the groups are located equidistantly. To conclude, the genetic diversity of the Kyrgyz mountain merino in three state breeding factories of the Kyrgyz Republic was high and comparable with each other.

Key words: Kyrgyz mountain merino; genotyping; STR markers.

For citation: Bekturov A.B., Isakova Zh.T., Kipen V.N., Chortonbaev T.Dzh., Mukeeva S.B., Osmonaliev S.K., Aitbaev K.A. A genogeographic study of the Kyrgyz mountain merino via microsatellite markers. *Vavilovskii Zhurnal Genetiki i Selektsii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):162-168. DOI 10.18699/VJGB-23-22

Геногеографическое исследование киргизского горного меринуса с использованием микросателлитных маркеров

А.Б. Бектуров¹, Ж.Т. Исакова²✉, В.Н. Кипень³, Т.Д. Чортонбаев¹, С.Б. Мукеева², С.К. Осмоналиев⁴, К.А. Айтбаев²

¹ Кыргызский национальный аграрный университет им. К.И. Скрябина, Бишкек, Кыргызская Республика

² Научно-исследовательский институт молекулярной биологии и медицины, Бишкек, Кыргызская Республика

³ Институт генетики и цитологии Национальной академии наук Беларуси, Минск, Республика Беларусь

⁴ Кыргызский научно-исследовательский институт животноводства и пастбищ, Сокулук, Кыргызская Республика

✉ jainagul@mail.ru

Аннотация. Проведено геногеографическое изучение породы овец киргизский горный меринос (КГМ). Проанализированы образцы ДНК 109 овец данной породы, разводимых в трех государственных племенных заводах (ГПЗ) в Республике Кыргызстан: ГПЗ «Оргочор» (Иссык-Кульская область), ГПЗ «Катта-Талдык» (Ошская область) и ГПЗ им. М.Н. Луцихина (Таласская область). В 12 исследованных микросателлитных маркерах (*McM042*, *INRA006*, *McM527*, *ETH152*, *CSRD247*, *OarFCB20*, *INRA172*, *INRA063*, *MAF065*, *MAF214*, *INRA005*, *INRA023*) идентифицировано 126 аллелей. Число аллелей в каждом локусе варьировало от 6 до 16 при среднем значении 10.500 ± 0.957 аллелей на локус. Определено 67 редких аллелей (с частотой встречаемости менее 5.0 %), что составляет 53.2 % от общего количества выявленных аллелей. Наибольшее количество редких аллелей установлено для STR-маркеров *CSRD247*, *INRA023*, *INRA005*, *INRA006*, *MAF214* и *OarFCB20*. Для каждой группы имеются индивидуальные различия в профиле распределения частот аллелей по всем исследуемым STR-локусам, наиболее значимые из которых следующие: в группах TALAS и OSH для локуса *McM042* в мажорном состоянии находится аллель 87 (35.6 и 45.7 % соответственно), в то время как для группы ISSYK-KUL наибольшее количество аллелей

шую распространенность получил аллель 95 (36.2 %); для локуса *INRA172* во всех группах мажорным аллелем был 154, однако в сравнении с группой TALAS его распространенность была меньше в 1.25 (ISSYK-KUL) и 1.66 (OSH) раза – 55.2 и 41.4 % соответственно, а аллели 156 и 158 встречались только в группе ISSYK-KUL; для локуса *ETH152* частота встречаемости аллеля 186 в группе TALAS составила 51.1 %, для групп ISSYK-KUL и OSH значительную распространенность приобретает аллель 190 – 34.5 и 34.3 % соответственно. При оценке генетической подразделенности исследуемых выборок КГМ (при K от 3 до 10) показана однородность структуры – вклад каждого субкластера равноценный. При анализе AMOVA обнаружено, что выборки расположены равноудаленно. Таким образом, генетическое разнообразие овец породы КГМ среди трех государственных племенных заводов Кыргызской Республики достаточно высокое и сопоставимое между собой.

Ключевые слова: киргизский горный меринос; генотипирование; STR-маркеры.

Introduction

The sheep breeds of Kyrgyz mountain merino (KMM) are common in all regions of the Kyrgyz Republic, which differ in natural and climatic conditions. In order to improve the breeding and productive qualities of KMM sheep, intra-breed (zonal) types of sheep were created (Bekturov et al., 2017).

The Kyrgyz mountain merino was created in 1990–2006 on the basis of the Kyrgyz fine-wool breed using sheep of the Australian merino breed and approved in 2006. The genetic structure of the breed includes 5 factory types and 24 factory lines. KMM sheep wool has high technological properties and has attributed to the highest quality categories of merino fine wool. The sheep are also known for high meat properties.

Each breed or animal type shows some heterogeneity in morphological, productive and technological qualities. Microsatellite loci (short tandem repeat, STR) can be used to solve breeding tasks related to the determination of breed affiliation or breed type (Deniskova et al., 2018; Isakova et al., 2019, 2021; Kharzinova, Zinovieva, 2020; Nosova et al., 2020; Lemesh et al., 2021).

To assess the condition and preserve the features of the KMM gene pool, genogeographic studies are needed. The preservation and further improvement of the breed should be controlled by the genetic dynamics studies both in the breed as a whole and in the main breeding farms engaged in KMM breeding. We have previously shown that local breeds of farm animals (in particular, the Kyrgyz horse) are characterized with a high genetic diversity, but local differentiation is also present, and the differences are significant for a number of high-altitude experimental zones (Isakova et al., 2021). In this regard, studies of similar structure are needed.

The information obtained during the molecular genetic analysis will complement the morphometric characteristics of breeding rams, repair rams and ewes, which will allow breeders to develop new and modify existing selection algorithms and schemes to maintain the inbred KMM genetic diversity, as well as preserve the genetic identity of this breed. In the future, they plan a number of measures to improve the breeding qualities of KMM breed sheep.

Thus, the purpose of this study was to conduct a genogeographic study of the Kyrgyz mountain merino sheep breed.

Materials and methods

The biological material for molecular genetic research was the blood samples of Kyrgyz mountain merino (KMM) sheep obtained from an adult population of 109 animals bred

in three state breeding plants (SBF), including 29 animals from SBF “Orgochor” (village Orgochor, Jety-Oguz district, Issyk-Kul region) (ISSYK-KUL sample), 35 animals from SBF “Katta-Taldyk” (village Bash-Bulak, Karasu district, Osh region) (OSH sample) and 45 animals from SBF named after M.N. Lushchikhin (village of Dzhun-Tube, Kara-Burinsky district, Talas region) (TALAS sample). The sampling sites are shown in Fig. 1.

DNA was isolated by phenol-chloroform extraction (Sambrook, Russel, 2001). The samples were genotyped using 12 microsatellite markers recommended by the International Society for Animal Genetics (ISAG): *McM042*, *INRA006*, *McM527*, *ETH152*, *CSR247*, *OarFCB20*, *INRA172*, *INRA063*, *MAF065*, *MAF214*, *INRA005*, *INRA023*, and also by the *AMEL* sex-specific locus.

Genotyping was carried out using a set of COrDIS Sheep (LLC “GORDIZ”, Russia) reagents for multiplex analysis according to the manufacturer’s recommendations. To correctly determine the genotype in the studied animals (amplicon size in bp), a sample with a control genotype included in the COrDIS Sheep kit was used. PCR were analyzed by capillary high-resolution electrophoresis using an automatic genetic analyzer Applied Biosystems 3500 (ThermoFisher, USA).

GenAIEx v. 6.503 (Peakall, Smouse, 2012), STRUCTURE v. 2.3.4 (Pritchard et al., 2000), Past v. 4.03 (Hammer et al., 2001) software was used for statistical analysis.

GenAIEx v. 6.503 was used to calculate the average number of alleles per locus (N_a), the effective number of alleles (N_e), the levels of expected (H_e) and observed (H_o) heterozygosity and the F_{IS} coefficient (Excoffier, 1991). STRUCTURE v. 2.3.4 allowed to calculate the Q criterion, which attributed each individual animal to the corresponding cluster (Pritchard et al., 2000). PPHelper v. 1.0.10 web application (Francis, 2016) was used for graphical interpretation of the results obtained in STRUCTURE v. 2.3.4.

We used GenAIEx 6.503 software (Peakall et al., 2012) to analyze population genetic parameters, the degree of genetic differentiation based on matrices of pairwise F_{ST} values, followed by visualization in Past v. 4.03 (Hammer et al., 2001).

The genetic structure of the studied samples of the KMM sheep breed was evaluated using principal component analysis (PCA) via clustering in STRUCTURE v. 2.3.4 (Pritchard et al., 2000) using a mixed model (the number of assumed K clusters from 3 to 10; the length of the burn-in period 50K; the Markov chain model Monte Carlo 5K). Ten iterations were completed for each K value. We also determined the



Fig. 1. The sampling sites: 1 – SBF “Orgochor” (village Orgochor, Jety-Oguz district, Issyk-Kul region); 2 – SBF “Katta-Taldyk” (village Bash-Bulak, Karasu district, Osh region); 3 – SBF named after M.N. Lushchikhin (village of Dzhun-Tube, Kara-Burinsky district, Talas region).

optimal number of clusters (ΔK) in POPHELPER v. 1.0.10 web application, using the method proposed in (Evanno et al., 2005).

All applicable international, national and/or institutional principles for the care and use of animals have been observed.

Results and discussion

The modern KMM sheep breed demonstrated a high level of inbreeding genetic variability, when 126 alleles were identified in the 12 microsatellite markers studied. The number of alleles in each locus varied from 6 to 16 (mean 10.500 ± 0.957). Sixty-seven rare alleles (with a prevalence less than 5.0 %) were identified, 53.2 % of the total number of identified alleles. The greatest number of rare alleles was found for the STR markers *CSRD247*, *INRA023*, *INRA005*, *INRA006*, *MAF214* and *OarFCB20*.

In order to analyze KMM inbreeding genetic subdivision bred in three geographically isolated zones, we computed N_a , N_e , H_o , H_e , I values and the F_{IS} coefficient, shown in Table 1.

The mean number of alleles per N_a locus varied from 8.000 to 8.500 (mean 8.306 ± 2.595), whereas the maximum value was noted in the TALAS group from the M.N. Lushchikhin SBF. The number of effective N_e alleles was the highest in the OSH sample from the Katta-Taldyk SBF. Shannon index,

reflecting the complexity of the community structure, averaged 1.657 ± 0.333 with the highest value in the OSH sample from the Katta-Taldyk SBF. The observed heterozygosity H_o as an indicator of the variability (polymorphism) of the population reflecting the proportion of heterozygous genotypes in the experiment ranged from 0.693 to 0.764. The expected heterozygosity of H_e as an indicator of the proportion of heterozygous genotypes, expected in the Hardy–Weinberg equilibrium, ranged from 0.730 to 0.770. Maximum values of H_o and H_e were in OSH from the Katta-Taldyk SBF. The mean value of F_{IS} index was the most neutral (0.006) in this group and indicated a balanced prevalence of heterogeneous genotypes, i.e. the level of related mating of individuals in the subpopulation was the least significant compared to the remaining two groups. In general, when comparing N_a , N_e , H_o , H_e , I and the F_{IS} coefficient, we found no statistically significant differences between three studied samples as of the Student’s t -test.

To assess the genetic subdivision of the KMM samples using STRUCTURE v. 2.3.4, we computed the Q criterion, which characterized the stratification of each individual animal in the corresponding group. A Q value of 75 % or higher confirms the individual’s attribution to its cluster. Fig. 2 graphically demonstrates (using the POPHELPER v. 1.0.10 web application (<http://pophelper.com/>)) the results of the analysis carried out in STRUCTURE v. 2.3.4 (automatic sorting was carried out based on the attribution of a particular sample to a major cluster).

The genetic material of KMM sheep from three geographically isolated zones was used in the study (see Fig. 1). For all samples within clusters $K = (3-10)$, there is a general uniformity of structure, whereas the contribution of each subcluster is equivalent. A pairwise comparison of the mean values of Q for three samples at $K = 2$ using analysis of variance showed no statistically significant differences. Thus, $F = 0.112$, $p = 0.739$ was for the pair TALAS/ISSYK-KUL; $F = 0.023$, $p = 0.881$, for the pair ISSYK-KUL/OSH; and $F = 0.267$, $p = 0.607$ was for the pair TALAS/OSH. This may result from the fact that the KMM subpopulations studied have common ancestors (for example, sheep producers); however, other factors may also have an effect.

Based on the analysis of F_{ST} genetic distances calculated using the AMOVA algorithm for 12 STR markers, a PCR graph

Table 1. Genetic and population characteristics of three independent KMM samples based on 12 STR markers

Sample		N_a	N_e	I	H_o	H_e	F_{IS}
TALAS	Mean	8.500	4.151	1.619	0.693	0.730	0.052
	Standard deviation	0.774	0.392	0.104	0.033	0.029	0.025
ISSYK-KUL	Mean	8.000	4.381	1.634	0.750	0.741	-0.015
	Standard deviation	0.769	0.468	0.105	0.030	0.027	0.027
OSH	Mean	8.417	4.617	1.718	0.764	0.770	0.006
	Standard deviation	0.763	0.387	0.084	0.017	0.015	0.020

Note. N_a – No. of different alleles per locus; N_e – No. of effective alleles; I – Shannon’s information index; H_o – observed heterozygosity; H_e – expected heterozygosity; F_{IS} – fixation index.

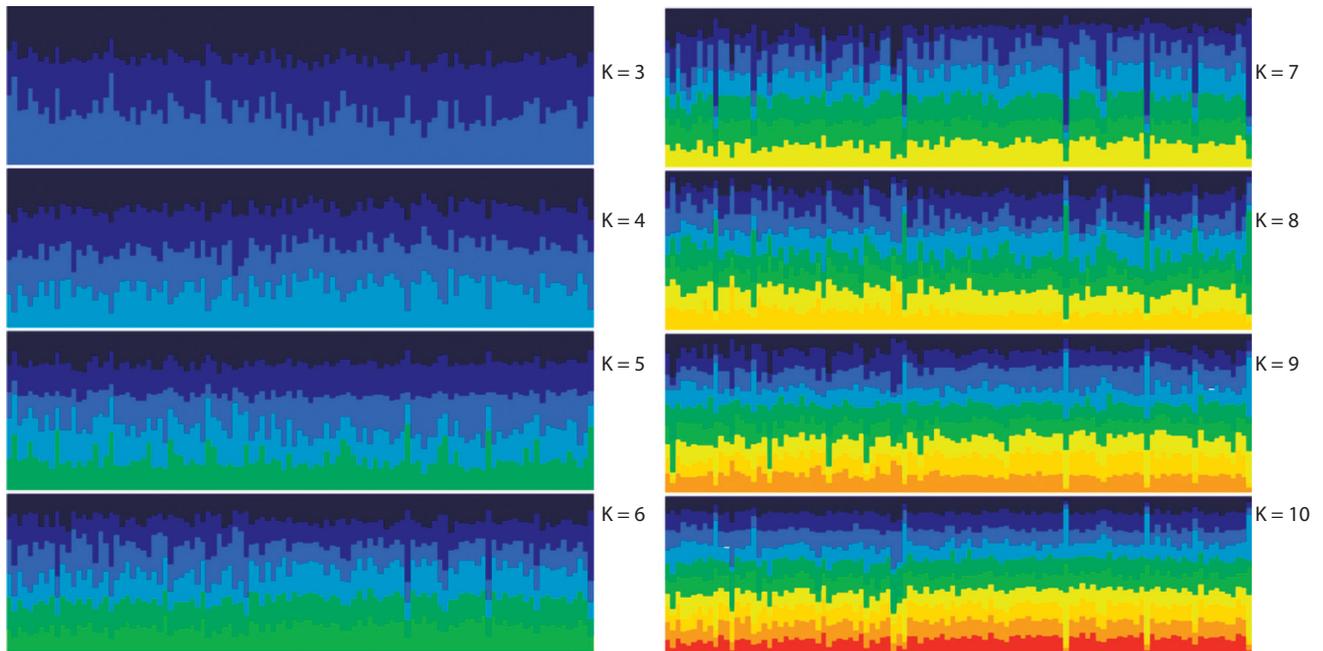


Fig. 2. The analysis of the genetic structure of the studied KMM samples for the most probable number of clusters (K) from 3 to 10. X – axis is the ID of the animal; Y – axis is the proportion in the corresponding cluster; Q values are calculated using the method of (Pritchard et al., 2000).

was constructed reflecting the mutual similarity/difference of the studied samples (Fig. 3).

The information presented in Fig. 2 and 3 allows to conclude that the studied samples of KMM did not differ significantly from each other. However, each sample had features that arose from the differences in the allele's prevalence in the studied STR loci, as well as the presence of rare and private (found only in one of the studied groups) alleles (Tables 2 and 3, respectively).

Among individuals from the M.N. Lushchikhin SBF (the TALAS sample), rare alleles accounted for 18.9, 12.2 and 10.0 %, respectively, for the *CSRD247*, *INRA005* and *INRA023* STR markers; among individuals from the Orgochor SBF (the ISSYK-KUL sample), high prevalence of rare alleles was found for STR markers *CSRD247* and *MAF214*, 12.1 and 10.3 %, respectively; and among individuals from the Katta-Taldyk SBF (OSH sample) – for *CSRD247* (15.7%), *MAF214* (15.7 %) and *OarFCB20* (14.3 %).

In general, we found individual differences in the distribution profile of allele frequencies across all the studied STR loci for each group. The most significant of those were allele 87 in the major state in the *McM042* locus (35.6 and 45.7 %, respectively) in the TALAS and OSH groups, whereas allele 95 was most prevalent (36.2 %) in the group ISSYK-KUL; major allele 154 for the *INRA172* locus in all groups, however, in comparison with the TALAS group, its prevalence was 1.25 (ISSYK-KUL) and 1.66 (OSH) times lower, 55.2 and 41.4 %, respectively, and alleles 156 and 158 were found only in the ISSYK-KUL group; the prevalence of 186 allele in the *ETH152* locus in the TALAS group was 51.1 %, whereas 190 allele was highly prevalent in ISSYK-KUL and OSH, 34.5 and 34.3 %, respectively.

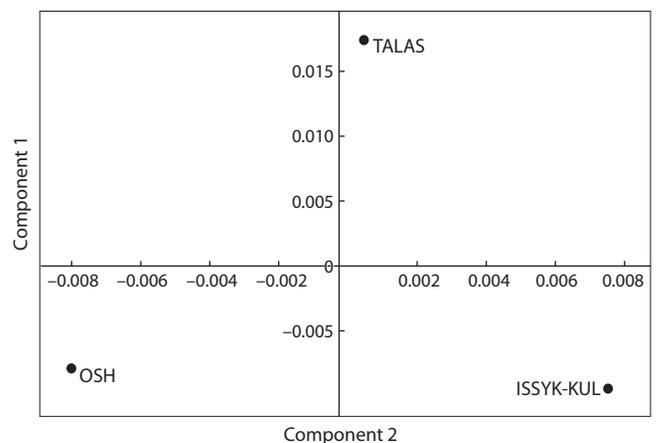


Fig. 3. Results of the analysis of the main components (as of 12 STR-markers in total).

We also identified specific peculiarities of private alleles. Among the sheep from the M.N. Lushchikhin SBF (TALAS) those were determined with regard to seven loci (a total of 10 alleles), including *INRA006*, *McM527*, *ETH152*, *CSRD247*, *INRA063*, *MAF214* and *INRA005*; and for the *INRA006* locus, the 124 allele was detected in 7.8 %. Seven private alleles in five STR markers were identified for sheep from the Orgochor SBF (ISSYK-KUL), the most common being *INRA172* (allele 156, frequency – 6.9 %) and *INRA023* (208, 6.9 %), as opposed to *INRA005* (147, 5.7 %) for individuals from the Katta-Taldyk SBF (OSH).

The highest calculated F_{ST} values are shown for the loci *McM042*, *INRA172* and *ETH152*, although in general the

Table 2. Total prevalence of rare alleles (prevalence less than 5 %) in the studied KMM samples in %

STR-marker	Sample		
	TALAS	ISSYK-KUL	OSH
<i>CSRD247</i>	18.9 (209/215/217/225/229/233/235/237/241)	12.1 (211/217/225/233/237)	15.7 (209/215/223/225/229)
<i>ETH152</i>	2.2 (200)	3.4 (192)	2.9 (198)
<i>INRA005</i>	12.2 (113/131/139/141)	8.6 (113/137/145)	11.4 (129/133/141/143)
<i>INRA006</i>	3.3 (114/118/134)	3.4 (120/134)	2.9 (118/126)
<i>INRA023</i>	10.0 (212/214/218)	3.4 (206/214)	5.7 (192/210/212)
<i>INRA063</i>	7.8 (173/179/197/199)	8.6 (189/199/201)	10.0 (187/195/199/201)
<i>INRA172</i>	3.3 (166/168)	8.6 (144/15/162/166)	1.4 (168)
<i>MAF065</i>	4.4 (131/135/137)	5.2 (123/135)	5.7 (123/131/137)
<i>MAF214</i>	7.8 (183/221/225/255/261)	10.3 (183/223/225/255)	15.7 (183/225/255/261/269)
<i>McM042</i>	8.9 (99/103)	3.4 (103)	7.1 (81/97)
<i>McM527</i>	2.2 (158/176)	3.4 (176)	–
<i>OarFCB20</i>	8.9 (95/103/111/113)	6.9 (107/111/113)	14.3 (77/83/95/103/107/113)

Table 3. The prevalence of private alleles in the KMM studied samples

Sample	STR-marker	Allele	Frequency, %
TALAS	<i>INRA006</i>	114	1.1
TALAS	<i>INRA006</i>	124	7.8
TALAS	<i>McM527</i>	158	4.4
TALAS	<i>ETH152</i>	200	2.2
TALAS	<i>CSRD247</i>	235	2.2
TALAS	<i>CSRD247</i>	241	1.1
TALAS	<i>INRA063</i>	167	5.6
TALAS	<i>INRA063</i>	197	1.1
TALAS	<i>MAF214</i>	221	1.1
TALAS	<i>INRA005</i>	139	1.1
ISSYK-KUL	<i>INRA006</i>	120	1.7
ISSYK-KUL	<i>CSRD247</i>	211	1.7
ISSYK-KUL	<i>CSRD247</i>	243	5.2
ISSYK-KUL	<i>INRA172</i>	156	6.9
ISSYK-KUL	<i>INRA172</i>	158	1.7
ISSYK-KUL	<i>MAF214</i>	223	1.7
ISSYK-KUL	<i>INRA023</i>	208	6.9
OSH	<i>INRA006</i>	126	1.4
OSH	<i>ETH152</i>	198	2.9
OSH	<i>OarFCB20</i>	77	2.9
OSH	<i>OarFCB20</i>	83	1.4
OSH	<i>INRA063</i>	187	1.4
OSH	<i>INRA063</i>	195	1.4
OSH	<i>MAF214</i>	269	4.3
OSH	<i>INRA005</i>	147	5.7
OSH	<i>INRA023</i>	210	1.4

F_{ST} values for all loci were not high and did not exceed 0.05 ($p < 0.001$).

We also conducted a comparative analysis of N_a and H_e parameters for KMM and fine-wool sheep breeds bred in Kazakhstan (Dossybayev et al., 2019), Russia (Deniskova et al., 2016), Pakistan (Ahmed et al., 2014) and Poland (Szumiec et al., 2018) (Table 4).

We found that the mean N_a in KMM (in the context of the STR markers studied in this paper) was the maximum in comparison with other studies. The calculated H_o index also turned out to be one of the largest and was comparable with the values obtained for the breeds Wielkopolskaya (Poland), Olkuska (Poland), Kail (Pakistan) and Kazakh fine-haired (Kazakhstan) (Ahmed et al., 2014; Szumiec et al., 2018; Dossybayev et al., 2019). The high rates of KMM genetic diversity are directly related to the multi-stage breeding processes that this breed underwent during the late XX–early XXI century.

Conclusion

Taken together, the genetic diversity of KMM breed sheep of the three state breeding plants of the Kyrgyz Republic is quite high and comparable to each other. We found it impossible to single out a group for which a qualitatively different (high or low) genetic diversity would be different compared the other two groups.

Nevertheless, it cannot be denied that for Kyrgyz mountain merino sheep from the M.N. Lushchikhin SBF, there was still a slight shift towards inbreeding processes – $F_{IS} = 0.052 \pm 0.025$ (the maximum individual values of this indicator were found for STR markers of *INRA023* – 0.120, *McM527* – 0.136, *McM042* – 0.142 and *MAF214* – 0.215). In this regard we assume that the positive shift of these markers (lack of heterozygotes) occurred due to the purposeful selection of individuals according to the economically valuable characteristics of wool, i.e. resulted from the association of these STR markers with the loci of quantitative traits QTL. However, such a relationship can only be assessed in further studies.

Table 4. Genetic characteristics of fine-wool breeds sheep samples based on STR loci genotyping

Breed (n)	STR	N_a	H_o	Reference
Pakistan				
Kail (47)	11	5.27 ± 1.49	0.766 ± 0.248	Ahmed et al., 2014
Russia				
Grozny (30)	11	9.00 ± 1.14	0.540 ± 0.089	Deniskova et al., 2016
Stavropol (32)		9.20 ± 0.92	0.575 ± 0.061	
Manych merino (30)		8.20 ± 0.90	0.647 ± 0.055	
Soviet merino (23)		8.00 ± 0.75	0.651 ± 0.060	
Salskaya (30)		8.50 ± 0.92	0.512 ± 0.089	
Volgogradskaya (30)		8.90 ± 1.22	0.525 ± 0.082	
Dagestan mountain breed (30)		9.00 ± 1.07	0.560 ± 0.079	
Transbaikalian fine-fleece (30)		8.90 ± 0.77	0.891 ± 0.018	
Kulunda (30)		7.20 ± 0.98	0.489 ± 0.095	
Poland				
Stavropol merino (93)	11	7.18 ± 1.94	0.663 ± 0.167	Szumiec et al., 2018
Olkuska(88)		5.64 ± 1.29	0.689 ± 0.138	
Wielkopolska (100)		7.82 ± 2.23	0.710 ± 0.065	
Kazakhstan				
Kazakh argali merino (15)	12	7.08 ± 0.64	0.678 ± 0.051	Dossybayev et al., 2019
Kazakh fine-haired (15)		7.92 ± 0.56	0.744 ± 0.048	
Kyrgyzstan				
Kyrgyz mountain merino (109)	12	10.50 ± 0.96	0.731 ± 0.023	Current study

An indirect confirmation of the inbreeding processes in this breeding plant may be the presence of six pairs of individuals among those selected for molecular genetic analysis, which were likely close relatives to each other (within pairs), because there were matching alleles in each of the 12 STR markers. In this regard, we propose to have a closer look at the intensity of inbreeding in the future. Four similar pairs were also identified among the individuals from the Orgochor and Katta-Taldyk SBFs, and it was possible that breeding events for the exchange of breeding sheep or repair sheep between these enterprises took place relatively recently.

References

- Ahmed Z., Babar M.E., Hussain T., Awan F.I. Genetic diversity analysis of Kail sheep by using microsatellite markers. *J. Anim. Plant Sci.* 2014;24(5):1329-1333.
- Bekturov A.B., Chortonbayev T.D., Chebodayev D.V. Tien Shan type of Kyrgyz mountain merino sheep and their performance. *Vestnik Altayskogo Gosudarstvennogo Agrarnogo Universiteta = Bulletin of the Altai State Agrarian University.* 2017;5(151):100-103. (in Russian)
- Deniskova T.E., Dotsev A.V., Okhlopov I.M., Bagirov V.A., Brem G., Zinovieva N.A., Kramarenko A.S. Characterization of the genetic structure of snow sheep (*Ovis nivicola lydekkeri*) of the Verkhoyansk mountain chain. *Russ. J. Genet.* 2018;54(3):328-334. DOI 10.1134/S1022795418030031.
- Deniskova T.E., Selionova M.I., Gladyr E.A., Dotsev A.V., Bobryshova G.T., Kostyunina O.V., Brem G., Zinovieva N.A. Variability of microsatellites in sheep breeds raced in Russia. *Selskokhozyaystvennaya Biologiya = Agricultural Biology.* 2016;51(6):801-810. DOI 10.15389/agrobiol.2016.6.801rus. (in Russian)
- Dossybayev K., Orazymbetova Z., Mussayeva A., Saitou N., Zhapbasov R., Makhatov B., Bekmanov B. Genetic diversity of different breeds of Kazakh sheep using microsatellite analysis. *Arch. Anim. Breed.* 2019;62(1):305-312. DOI 10.5194/aab-62-305-2019.
- Evanno G., Regnaut S., Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 2005;14(8):2611-2620. DOI 10.1111/j.1365-294X.2005.02553.x.
- Excoffier L. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics.* 1992;131(2):479-491. DOI 10.1093/genetics/131.2.479.
- Francis R.M. POPHELPER: An R package and web app to analyse and visualise population structure. *Mol. Ecol. Resour.* 2016;17(1):27-32. DOI 10.1111/1755-0998.12509.
- Hammer Q., Harper A.T., Ryan P.D. PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 2001;4(1):1-9.
- Isakova Zh.T., Isaev M.A., Kipen V.N., Kalinkova L.V., Aitbaev K.A., Arzybaev M.A., Mukeeva S.B., Osmoykul k. Meerim, Aldasheva N.M. Genetic diversity of the Kyrgyz horse breed using microsatellite markers – extended genogeographic study. *Russ. J. Genet.* 2021;57(4):438-445. DOI 10.1134/S1022795421040037.

- Isakova Zh.T., Toktosunov B.I., Kipen V.N., Kalinkova L.V., Talaibekova E.T., Aldasheva N.M., Abdurasulov A.H. Phylogenetic analysis of Kyrgyz Horse using 17 microsatellite markers. *Russ. J. Genet.* 2019;55:100-104. <https://doi.org/10.1134/S1022795419010071>.
- Kharzinova V.R., Zinovieva N.A. The pattern of genetic diversity of different breeds of pigs based on microsatellite analysis. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2020;24(7):747-754. DOI 10.18699/VJ20.669. (in Russian)
- Lemesh V.A., Ageets V.Yu., Nosonova A.Yu., Kipen V.N., Tsar N.I., Sergeeva T.A., Savicheva E.A. Genetic structure of the carp population (*Cyprinus carpio carpio*) grown in aquaculture in the Republic of Belarus. *Doklady Natsionalnoj Akademii Nauk Belarusi = Reports of the National Academy of Sciences of Belarus.* 2021; 65(1):68-75. DOI 10.29235/1561-8323-2021-65-1-68-75. (in Russian)
- Nosova A.Yu., Kipen V.N., Tsar A.I., Lemesh V.A. Differentiation of hybrid progeny of Silver Carp (*Hypophthalmichthys molitrix* Val.) and Bighead Carp (*H. nobilis* Rich.) based on microsatellite polymorphism. *Russ. J. Genet.* 2020;56:317-323. <https://doi.org/10.1134/S1022795420030126>.
- Peakall R., Smouse P.E. GenAlEx 6.503: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics.* 2012;28:2537-2539. DOI 10.1093/bioinformatics/bts460.
- Pritchard J.K., Stephens M., Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155(2):945-959. DOI 10.1093/genetics/155.2.945.
- Sambrook J., Russell D.W. *Molecular Cloning: A Laboratory Manual.* 3rd ed. New York: Cold Spring Harbor Laboratory Press, 2001.
- Szumiec A., Radko A., Koseniuk A., Rubis D., Bugno-Poniewierska M. Application of 11 STR markers for the evaluation of genetic variation in sheep. *ICAR Tech. Ser.* 2018;23:141-145.

Acknowledgements. The study was carried out with the financial support of the Ministry of Education and Science of the Kyrgyz Republic within the framework of the scientific project "The use of molecular genetic research methods and the creation of information resources for the intensification of sheep breeding in the Kyrgyz Republic".

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

Received March 2, 2022. Revised July 8, 2022. Accepted August 22, 2022.

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

Genetic profile of domestic cat (*Felis catus* L.) population of Aoshima Island (Japan)

S.K. Kholin

Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia

✉ h.axyridis@mail.ru

Abstract. The paper analyzes the genetic profile of the domestic cat population of the Aoshima Island. The population has been established in the middle of the last century, after a small group of animals was imported for rodent control. Based on three photographs, the genotypes of the cats in three overlapping groups (75, 56, and 70 individuals) were determined. The mutant allele frequencies of the sex-linked *O* (*Orange*) locus and the three autosomal loci *a*, *W*, and *I* (*Agouti*, *White*, and *Long hair*) responsible for coat color and length were estimated. The population lacks the mutant alleles *d* (*Dilution* locus), *W* and *w^g* (*White*), *ta^b* (*Tabby*), *Ti^A* (*Ticked*) present in other populations of Japan. This is an almost monomorphic population with prevailing red and tortoiseshell individuals. Most cats have interrupted stripes (genotype $Ti^+Ti^+Ta^{M-}$). The island's population differs from the other populations of the Japanese islands in the frequencies of two mutant alleles, *O* and *a*. The frequency of the *O* allele ($q(O) = 0.580$) is one of the highest in the region, while the frequency of the *a* allele ($q(a) = 0.276$) is two times lower than in the other populations. In both cases, the differences in frequencies between the neighbouring populations are significant ($p < 0.0001$). An independent study of the same population revealed a similar genetic structure. However, it detected the presence of the *d* allele, the frequency of the *a* allele was higher (0.534 versus 0.276, $p < 0.020$). The genetic profile, frequencies of mutant alleles in the population, and history of its origin indicate a significant influence of the founder effect on the genetic structure of the island's domestic cat population.

Key words: *Felis catus*; genetic profile; islands; founder effect; Japan.

For citation: Kholin S.K. Genetic profile of domestic cat (*Felis catus* L.) population of Aoshima Island (Japan). *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(2):169-176. DOI 10.18699/VJGB-23-23

Генетический профиль популяции домашней кошки (*Felis catus* L.) острова Аошима (Япония)

С.К. Холин

Федеральный научный центр биоразнообразия наземной биоты Восточной Азии Дальневосточного отделения Российской академии наук, Владивосток, Россия

✉ h.axyridis@mail.ru

Аннотация. Проанализирован генетический профиль популяции домашней кошки острова Аошима (Япония). Популяция возникла в середине прошлого века в результате завоза небольшой группы животных для борьбы с грызунами. По трем фотографиям определен фенотипический состав групп животных (75, 56 и 70 особей). Рассчитаны частоты мутантных аллелей сцепленного с полом локуса *Orange* и трех аутосомных локусов (*Agouti*, *White* и *Long hair*), отвечающих за окраску и длину шерсти. В популяции отсутствуют мутантные аллели *d* (локус *Dilution*), *W* и *w^g* (*White*), *ta^b* (*Tabby*), *Ti^A* (*Ticked*), отмеченные в других популяциях Японии. Фенотипически это практически мономорфная популяция с преобладающим большинством рыжих и черепаховых особей. Большинство кошек имеют прерванные полосы (генотип $Ti^+Ti^+Ta^{M-}$). Популяция острова резко отличается от остальных популяций островов Японии по частотам двух мутантных аллелей, *O* и *a*. Частота аллеля *O* ($q(O) = 0.580$) – одна из самых высоких частот в регионе, а аллеля *a* ($q(a) = 0.276$) – меньше в два раза, чем в других популяциях. В обоих случаях отличие от частот в окружающих портовых популяциях статистически значимо ($p < 0.0001$). Генетический состав, частоты мутантных аллелей в популяции и характер ее происхождения указывают на значительное влияние эффекта основателя на генетическую структуру популяции домашней кошки о-ва Аошима. Независимое исследование популяции кошек о-ва Аошима обнаружило сходную генетическую структуру. Однако в популяции найдены особи, носители аллеля *d*. Возможно, это может быть вызвано различиями в методике определения фенотипа таких кошек. Частота аллеля *a* статистически значимо выше (0.534 против 0.276, $p < 0.020$). Все это, однако, не влияет на общий вывод о действии эффекта основателя при возникновении популяции кошек этого острова.

Ключевые слова: *Felis catus*; генетический профиль; острова; эффект основателя; Япония.

Introduction

The domestic cat is a unique object of genetic research due to the high polymorphism of its populations for a number of traits (color, texture and length of coat, and some other features of external morphology). Freely living in human settlements, cats do not depend on humans for their reproduction. In this sense, they are similar to true natural populations. Cats' phenotypes can be easily identified at a distance. This makes it possible to collect sufficient data on allele frequencies without much effort. Such data can be used in various kinds of population genetic analysis, e. g., it has been shown that every cat population has its own genetic profile determined by its origin, location and population history (Lloyd, 1987).

From the point of view of population genetics, the history of studying the domestic cat (*Felis catus* L., 1758) dates back more than 65 years when the first investigation on the frequencies of the mutant alleles responsible for the color and length of coat in London's cat population was published (Searle, 1949). A surge of interest in the subject occurred in the late 1960s – early 1980s, when the research covered most of Europe and North America, Southeast Asia, and Australia (Kholin, 2018).

There are various methods for obtaining data on mutant allele frequencies in domestic cat populations, each having certain advantages and drawbacks (Twedt, 1983; Schüler, Borodin, 1992). Before the advent of digital cameras, the main method was direct observation of animals on the streets, in the yards or door-to-door survey. H. Todd and P. Jeanne (1972) were the first to use the photographic method. The use a single photo of a group of about 100 cats in Sao Paulo (Brazil) to carry out a detailed analysis of the cats' phenotypes and calculate the frequency of mutant alleles with varying degrees of accuracy. Digital photography has greatly facilitated the collection of the data and made it possible to obtain samples of sufficient size (several images of one individual), to accurately describe a cat's phenotype in the lab.

In Asia, it is the Japanese domestic cat that has been most intensively studied. Data on 105 (Nozawa et al., 1990), 174 (Nozawa et al., 2000) and 141 (Nozawa, Kawamoto, 2013) cat populations from small villages to megapolises in all prefectures of the four largest and 35 small islands have been obtained (Nozawa, 2019).

One of these islands, Aoshima, is about 400 hectares in size and located off the northern coast of Kyushu Island in the Inland Sea of Japan. The island is currently populated by no more than 10 people and 200 cats¹. The cats were brought there in the middle of the last century to reduce the number of rats damaging fishing nets. Eventually, the fishing industry fell into decay, but the cats remained, now being fed by the locals and the tourists arriving on the island.

The aim of the present study was to describe the genetic profile of the island's domestic cat population by analyzing the photographs of a group of cats and to compare the findings with the previous study of the Aoshima cats (Nozawa, 2019)² and other population studies.

Materials and methods

In March 2015, the photos of the island's cats were published in the media and on the Internet (Photos by Thomas Peter, Reuters, February 25, 2015)³. The quality of the published images was high enough to accurately describe the cat phenotypes. In this study, they were used as samples to assess the genetic profile of the cat population.

The photographs show individual cats and groups on the pier waiting for a boat to arrive. Three photos were selected to contain the largest numbers of animals (see the Figure and Supplementary Material)⁴. Each photograph was considered a separate sample from the same set of cats to estimate the stability of the assessment of the phenetic (genetic) composition of the population. The sample size for each locus depended on the visibility of each animal in the photograph. In total, the data of 75, 56 and 70 individuals were recorded (sample A, B, and C, respectively). The photographs were used to determine the number of individuals carrying a particular phenotype.

The data were used to calculate the frequencies of the *O* allele of the sex-linked *Orange* locus, as well as the frequencies of two alleles of autosomal loci: *Agouti* (*a*) and *Long hair* (*l*) (Table 1). The first two loci control the color of the coat, and the last, its length. The pattern of inheritance of these traits was described by R. Robinson (1993a, b). No individuals carrying the *d* allele of the *Dilution* locus were detected in any of the photographs.

The stripe pattern determined by autosomal loci *Tabby* (*Ta*) and *Ticked* (*Ti*) was also analyzed. The inheritance of this trait was described by E. Eizirik et al. (2010). The analysis demonstrated that among the cats there were no individuals homozygous for the *ta^b* allele producing marble coloration (*Blotched tabby*). Most of the cats had interrupted stripes (genotype *Ti⁺Ti⁺Ta^{M-}*). The cats carrying the dominant *Ti^A* allele are characterized by a complete absence of stripes on the body and the presence of stripes only on the head, paws and tail. Two individuals (see the Figure 1, B; Supplementary Material), (Nos. 29 and 42) in the foreground have a similar phenotype. However, the *Ti^A* allele was excluded from the analysis due to the difficulty of its unambiguous identification in the available photographs.

Previously, the *Piebald white spotting* (*S*) and *Dominant white* (*W*) were considered mutations at different loci.

² Unfortunately, the paper does not specify how the data were obtained. It is only known that the observations were carried out before 2015.

³ <http://www.theatlantic.com/photo/2015/03/a-visit-to-aoshima-a-cat-island-in-japan/386647/>. Last accessed March 12, 2023.

⁴ Supplementary Material is available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Kholin_Engl_27_2.pdf.

¹ <https://www.nippon.com/ru/behind/fnn20181019001/?pnum=2>. Last accessed March 12, 2023.



Aoshima cats selected as samples A and B (source: Reuters/Pixstream).



Aoshima cats selected as sample C (source: Reuters/Pixstream).

Table 1. Phenotypes and their corresponding genotypes in the domestic cat

Locus	Mutant allele	Phenotype	Genotype	Note
Sex-linked allele				
<i>Orange</i>	<i>O</i>	Red Tortoiseshell	<i>OO</i> and <i>OY</i> <i>Oo</i>	Cream in presence of <i>dd</i> genotype
Autosomal alleles				
<i>Aguti</i>	<i>a</i>	Black	<i>aa</i>	Not expressed in presence <i>OO</i> and <i>OY</i> genotype. Blue in presence of <i>dd</i> genotype
<i>Dilution</i>	<i>d</i>	Diluted	<i>dd</i>	
<i>Tabby</i>	<i>Ta^M</i>	Striped Spotted	<i>Ti⁺Ti⁺Ta^MTa-</i> <i>Ti⁺Ti⁺Ta^MTa+</i>	Not expressed in presence of <i>aa</i> genotype
	<i>Ta^b</i>	Marble	<i>Ti⁺Ti⁺ta^bta^b</i>	
<i>Ticked</i>	<i>Ti^A</i>	Abyssinian or ticked	<i>Ti^ATi^ATa^{M-}</i> or <i>Ti^ATi⁺Ta^{M-}</i>	Not expressed in presence of <i>aa</i> genotype
<i>White</i>	<i>W</i>	Dominant white	<i>WW</i> или <i>W-</i>	Epistatic to all other genotypes
	<i>W^S</i> <i>w^g</i>	Piebald Gloving or 'White socks'	<i>W^SW^S</i> или <i>W^{S-}</i> <i>w^gw^g</i>	
<i>Long hair</i>	<i>l</i>	Long hair	<i>ll</i>	

However, modern data indicate that they are semi-dominant (W^S) and dominant (W) mutations at the same W (KIT) locus (David et al., 2014). The former is responsible for the piebald coloration, the carriers of the latter are completely white. There is a third, previously unknown, recessive mutation w^g (*Gloving*) producing ‘white socks’ in homozygous cats. None of the examined photos show the cats carrying any of the two last alleles (w^g or W).

The frequencies of recessive alleles (q) were calculated as the square root of the frequencies of the corresponding phenotypes, and of dominant (p) – as $p = 1 - q$.

The standard errors (SE) were calculated as $\sqrt{\frac{1-q^2}{4n}}$ and $\sqrt{\frac{p(2-p)}{4n}}$, respectively (Robinson, Manchenko, 1981; Goncharenko et al., 1985).

Since the sex of the animals was not determined, the O allele frequencies of the sex-linked *Orange* locus were estimated using the maximum likelihood method assuming an equal sex ratio (Adalsteinsson, Blumenberg, 1984). In the first approximation, the formula

$$q = \frac{2a+b}{2n},$$

was applied where a and b are the numbers of red (genotype $O/-$) and tortoiseshell ($O/+$) cats, and n is the sample size ($n = a + b + c$, where c is the number of nonorange ($+/-$) individuals) (Robinson, 1972). To get a more accurate estimate an iterative algorithm $q_{i+1} = q_i + \frac{dL}{dq_i} \text{Var}(q_i)$ was used,

$$\text{where } \frac{dL}{dq} = \frac{a}{1+q} + \frac{a+b}{q} - \frac{c}{2-q} - \frac{b+c}{1-q},$$

$$\frac{1}{\text{Var}(q)} = 0.5N \left\{ \frac{q}{1+q} + \frac{3-q}{q} + \frac{1-q}{2-q} + \frac{2+q}{1-q} \right\}.$$

Its SE was calculated as $\sqrt{\text{Var}(q)}$.

To estimate the random mating (panmixia), the expected numerical ratio of genotypes a , b and c was estimated using the formulas: $0.5qn(1+q)$, $qn(1-q)$ и $0.5n(2-q)(1-q)$, respectively.

Testing of the statistical hypotheses was carried out using the χ^2 - and G -tests, the last having a distribution similar to that of χ^2 but being more convenient for analyzing contingency tables. Pairwise comparison of samples for individual loci was carried out using the χ^2 -test and the arcsine-transformation of the allele frequencies (Zhivotovskiy, 1991).

To assess the genetic differentiation of F_{st} and G_{st} (Kuznetsov, 2020), a computational add-on for Excel GenAlEx 6.503 (Peakall, Smouse, 2012) was applied.

Results and discussion

Table 2 shows the results of testing for panmixia at the *Orange* locus. In all cases we observed a good correspondence between the observed and expected frequencies of the genotypes ($p > 0.20$). The test for heterogeneity in the ratio of the genotype frequencies indicated the absence of significant differences between the samples for this characteristic ($G = 1.232$, $df = 4$, $p > 0.85$). The O allele frequency in the samples under consideration ranged 0.570–0.589 (mean, 0.580 ± 0.052) and was homogeneous ($\chi^2 = 0.049$, $df = 2$, $p > 0.95$).

Table 3 shows the estimates of the frequencies of the other alleles calculated under the assumption of panmixia. Testing for heterogeneity in the frequencies of mutant phenotypes did not reveal significant differences between the samples ($p > 0.15$ in all cases).

A comparison with data on the frequencies of mutant alleles in the main islands of Japan showed that the allele O frequency (0.580) in Aoshima was two or more times higher than in neighboring populations ($q(O) = 0.232$ (0.154–0.412)), and the frequencies throughout Japan ($q(O) = 0.220$ (0.095–0.490)) (Nozawa, Kawamoto, 2013). In samples B and C, the differences were significant ($p < 0.0001$). In the case of allele a , the situation was reversed. Its frequency (0.276) was two or more times lower than in the neighboring populations ($q(a) = 0.691$ (0.614–0.783)) and throughout Japan ($q(a) = 0.697$ (0.463–0.839)), ($p < 0.0001$). Allele l

Table 2. Observed and expected genotype ratios of the *Orange* locus; the results of the χ^2 -test for panmixia; and the estimate of O allele frequency ($q(O)$) in the samples from the Aoshima Island

Sample	Genotypes ratio	Genotype			χ^2 , $df = 1^*$	$q(O) \pm SE$
		$O/-$	$O/+$	$+/-$		
A	Observed	36	13	26	1.668	0.570 ± 0.049
	Expected	33.56	18.38	23.06		
B	Observed	26	14	16	0.008	0.589 ± 0.057
	Expected	26.19	13.56	16.25		
C	Observed	34	13	23	0.930	0.581 ± 0.051
	Expected	32.21	17.03	20.76		

Note. “-” means the state of the second allele is unknown; * means $p > 0.20$ in all cases.

Table 3. Observed phenotype ratio (Obs.) and mutant alleles frequency estimates (q) in the island's samples

Genotype	Sample						Phenotype frequency homogeneity criterion,		Average allele frequency, $q \pm SE$
	A		B		C		G , $df=2$	p	
	Obs.	$q \pm SE$	Obs.	$q \pm SE$	Obs.	$q \pm SE$			
a/a	3/39	0.277 ± 0.077	3/30	0.316 ± 0.087	2/36	0.235 ± 0.081	0.218	0.896	0.276 ± 0.081
$+/-$	36/39		27/30		34/36				
$W^s/-$	40/63	0.396 ± 0.050	25/39	0.401 ± 0.064	32/66	0.282 ± 0.043	1.830	0.401	0.360 ± 0.052
$+/+$	23/63		14/39		34/66				
l/l	4/75	0.231 ± 0.056	4/56	0.267 ± 0.064	6/70	0.293 ± 0.057	0.400	0.818	0.264 ± 0.062
$+/-$	71/75		52/56		64/70				

Table 4. Indices of genetic differentiation between the insular and "mainland" populations

Index	<i>Orange</i>	<i>Agouti</i>	<i>White</i>	<i>Long hair</i>	Total
F_{st}	0.156	0.181	0.039	0.008	0.106
G_{st}	0.152	0.175	0.035	0.004	0.101

frequency (0.264) fit into the range of variability in the surrounding populations ($q(l) = 0.214$ (0.117–0.307)) and all Japanese islands ($q(l) = 0.181$ (0–0.412)), ($p > 0.15$).

In their studies published earlier (Nozawa et al., 1990, 2000; Nozawa, Kawamoto, 2013) pursued outdated ideas about the genetics of piebald coloration in cats: the piebald cats were considered as carriers of the dominant S allele of a *Piebald white spotting* locus. However, since the island's cat population has no w^s allele, for ease its comparison against the surrounding populations, in our study a proportion of piebald cats was used. In the considered samples they comprised $63.5 \pm 6.1\%$, $64.1 \pm 7.7\%$ and $48.5 \pm 6.2\%$, respectively (mean $58.7 \pm 6.6\%$). According to the test results, the samples were homogeneous in terms of the frequency of this trait (see Table 3, $p = 0.401$). The proportion of piebald cats on the island was not much higher than that in the surrounding populations ($q = 0.456$ (0.346–0.523)) and fit within the variability range for all the Japanese islands ($q = 0.532$ (0.188–0.815), $p > 0.25$).

Examining other photos of the island's cats showed the presence of the so-called bobtail (short-tailed) cats. Unfortunately, the available photos did not allow to estimate their proportion in the population, while in Japan their ratio varied from 0 to 79.6% (mean 28.8%) (Nozawa, Kawamoto, 2013).

There are not so many publications (about 40) devoted to the population genetics of island domestic cat populations (Kholin, 2018). In most cases, the mutant allele frequencies of the populations correspond to those in the populations they originated from (Lloyd, 1987). This situation is com-

monly observed on the islands with stable large settlements of people such as the Azores whose cat population came from Portugal (Todd, Lloyd, 1984). However, there are cases when groups of cats introduced accidentally or deliberately to small islands become feral. In such populations, the original genetic profile of the founding group has been preserved as a result of the founder effect (Dreux, 1974; van Aarde, Robinson, 1980; Jones, Horton, 1984).

This is the case of the Aoshima Island where the cats, once brought to the island, lived their lives protecting fishing nets from rats and nothing has changed for them in this respect since the fishermen left the island. This is why this population has contrasting genetic differences when compared to the nearby populations it may have descended from. This is evidenced by the high genetic differentiation (Table 4) at two loci between the island population and that of their possible ancestors, populating the nearest port of Matsuyama City. This differentiation also indicates the founding group had a homogeneous phenotypic composition. With a greater probability, these were red cats, since among sailors and fishermen there is a belief that red cats bring good luck.

One of the indirect confirmations of the absence of significant migration to the island after the cat population was established is the following fact. In the populations of the main islands of Japan, relatively low frequencies of the d and ta^b alleles are observed due to the country's long-term historical isolation. In the postwar years, a steady increase in the proportion of cats carrying these alleles was noted for they became popular with the people of Japan (Nozawa,

Table 5. Mutant allele frequencies estimations of Aoshima's domestic cat population and their statistical comparison

Allele	This paper	K. Nozawa, 2019	χ^2	<i>p</i>
<i>O</i>	0.580	0.568	0.020	0.886
<i>a</i>	0.276	0.534	5.742	0.016
<i>d</i>	?	0.204	–	–
<i>W^S</i>	0.360	0.333	0.101	0.750
<i>l</i>	0.264	0.312	0.391	0.532

Kawamoto, 2013). However, in the Aoshima population there are still no cats with the “marble” phenotype homozygous for the *ta^b* allele.

The presented data and the results of another study (Nozawa, 2019) indicate the resemblance of the frequency estimates obtained by different observers since in both cases comparable sample sizes have been obtained (56–75 and 72 individuals, respectively). Table 5 shows the estimates of the mutant allele frequencies and the results of their statistical comparison, which indicate no statistically significant differences, except for the *a* allele, the frequency of which is significantly higher in the (Nozawa, 2019) sample. Another difference is the presence of cats of diluted color phenotype in the Nozawa (2019) sample and their absence in our samples. This may be due to differences in the method for determining the phenotype of such cats. Thus, the founder effect has played a main role in the formation of the genetic composition of the island cat population.

What is interesting is the population's future since most of the cats were neutered in 2018⁵. But how this will affect the genetic structure of this population and its condition as a whole would only be shown by future research.

Conclusion

The genetic profile of the domestic cat populating the Aoshima island differs sharply from that of the populations of the port cities surrounding the island, and the Japanese population as a whole. The island's cats lack the alleles common to other populations, and have one of the highest frequencies of the *O* allele ($q(O) = 0.580$) observed in Japan. Phenotypically, this is an almost monomorphic population mainly composed of red and tortoiseshell individuals, which is probably due to the single introduction by fishermen to the island of a small group of cats with a high frequency of *O* allele carriers. Thus, the founder effect had a large influence on the formation of the genetic composition of the island's cat population.

References

Adalsteinsson S., Blumenberg B. Simultaneous maximum likelihood estimation of the frequency of sexlinked orange and the male ratio in the cat. *Carnivore Genet. Newsl.* 1984;4:68-77.

⁵ <https://www.nippon.com/ru/behind/fnn20181019001/?pnum=2>. Last accessed March 12, 2023.

David V.A., Menotti-Raymond M., Wallace A.C., Roelke M., Kehler J., Leighty R., Eizirik E., Hannah S.S., Nelson G., Schäffer A.A., Connelly C.J., O'Brien S.J., Ryugo D.K. Endogenous retrovirus insertion in the *KIT* oncogene determines *White* and *White spotting* in domestic cats. *G3 (Bethesda)*. 2014;4(10):1881-1891. DOI 10.1534/g3.114.013425.

Dreux Ph. The cat population of péninsule Courbet, îles Kerguelen: an example of the founder effect. *Polar Rec.* 1974;17(106):53-54. DOI 10.1017/S0032247400031405.

Eizirik E., David V.A., Buckley-Beason V., Roelke M.E., Schäffer A.A., Hannah S.S., Narfström K., O'Brien S.J., Menotti-Raymond M. Defining and mapping mammalian coat pattern genes: Multiple genomic regions implicated in domestic cat stripes and spots. *Genetics*. 2010;184(1):267-275. DOI 10.1534/genetics.109.109629.

Goncharenko G.G., Lopatin O.E., Manchenko G.P. Mutant color genes in populations of domestic cats of Central Asia and European part of the Soviet Union. *Genetika (Moscow)*. 1985;21(7):1151-1158. (in Russian)

Jones E., Horton B.J. Gene frequencies and body weights of feral cats, *Felis catus* (L.), from five Australian localities and from Macquarie Island. *Austral. J. Zool.* 1984;32(2):231-237. DOI 10.1071/ZO9840231.

Kholin S.K. Bibliography on the population genetics of the domestic cat (*Felis catus* L.). Vladivostok: Federal Scientific Center of the East Asia Terrestrial Biodiversity Publ., 2018. (in Russian)

Kuznetsov V.M. Nei's methods for analyzing genetic differences between populations. *Problemy Biologii Productivnykh Zhivotnykh = Problems of Productive Animal Biology*. 2020;1:91-110. (in Russian)

Lloyd A.T. Cats from history and history from cats. *Endeavour*. 1987; 11(3):112-115. DOI 10.1016/0160-9327(87)90197-9.

Nozawa K. Genetic polymorphisms in coat color and other morphological traits of the Japanese feral cats. The 5th compilation of small-island data. *Rep. Soc. Res. Native Livest.* 2019;29:105-120.

Nozawa K., Kawamoto Y. Genetic polymorphisms in coat color and other morphological traits of the Japanese feral cats. The 4th compilation of mainland data. *Rep. Soc. Res. Native Livest.* 2013;26: 105-139.

Nozawa K., Maeda Y., Hasegawa Y., Kawamoto Y. Genetic polymorphisms in coat color and other morphological traits of the Japanese feral cats. Report of the 3rd compilation of data. *Rep. Soc. Res. Native Livest.* 2000;18:225-268. (in Japanese)

Nozawa K., Namikawa T., Kawamoto Y. Genetic polymorphisms in coat color and other morphological traits of the Japanese feral cats. *Rep. Soc. Res. Native Livest.* 1990;13:51-115. (in Japanese)

Peakall R., Smouse P.E. GenA1Ex 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.

Robinson R. Mutant gene frequencies in cats of Cyprus. *Theor. Appl. Genet.* 1972;42(7):293-296. DOI 10.1007/BF00277721.

Robinson R. Genetics of colors. In: *Cat Genetics*. Novosibirsk: Nauka Publ., 1993a;44-53. (in Russian)

- Robinson R. Coat structure genetics. In: Cat Genetics. Novosibirsk: Nauka Publ., 1993b;53-57. (in Russian)
- Robinson R., Manchenko G.P. Cat gene frequencies in cities of the USSR. *Genetica*. 1981;55(1):41-46. DOI 10.1007/BF00134003.
- Schüler L., Borodin P.M. Influence of sampling methods on estimated gene frequency in domestic cat populations of East-Germany. *Arch. Anim. Breed.* 1992;35(6):629-634.
- Searle A.G. Gene frequencies in London's cats. *J. Genet.* 1949;49(3): 214-220. DOI 10.1007/BF02986074.
- Todd N.B., Jeanne R.L. Some cats of São Paulo, Brazil. *J. Hered.* 1972;63(6):321-323. DOI 10.1093/oxfordjournals.jhered.a108307.
- Todd N.B., Lloyd A.T. Mutant allele frequencies in the domestic cats of Portugal and the Azores. *J. Hered.* 1984;75(6):495-497. DOI 10.1093/oxfordjournals.jhered.a109994.
- Twedt D.J. Influence of survey methods and sample sizes on estimated gene frequencies in domestic cat population. *J. Hered.* 1983;74(2): 121-123. DOI 10.1093/oxfordjournals.jhered.a109736.
- van Aarde R.J., Robinson T.J. Gene frequencies in feral cats on Marion Island. *J. Hered.* 1980;71(5):366-368. DOI 10.1093/oxfordjournals.jhered.a109391.
- Zhivotovsky L.A. Population Biometry. Moscow: Nauka Publ., 1991. (in Russian)

ORCID ID

S.K. Kholin orcid.org/0000-0002-3016-2217

Acknowledgements. The author is sincerely grateful to Dr. R. Fagen (Juneau, USA), who suggested the idea of using photos of cats published on the Internet; to Dr. A.P. Kryukov (Federal Scientific Center of the East Asia Terrestrial Biodiversity FEB RAS, Vladivostok) – for assistance in obtaining Japanese publications, as well as to Professor P.M. Borodin (Institute of Cytology and Genetics SB RAS, Novosibirsk) – for his help in preparing the publication.

Conflict of interest. The author declares no conflict of interest.

Received October 24, 2022. Revised December 25, 2022. Accepted December 27, 2022.

Original Russian text <https://vavilov-jcg.ru/>

Alterations in the social-conditioned place preference and density of dopaminergic neurons in the ventral tegmental area in *Clstn2*-KO mice

I.N. Rozhkova¹, S.V. Okotrub¹, E.Yu. Brusentsev¹, K.E. Uldanova¹, E.A. Chuyko¹, V.A. Naprimerov^{1,2}, T.V. Lipina³, T.G. Amstislavskaya⁴, S.Ya. Amstislavsky¹ ✉

¹ Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Novosibirsk State Agricultural University, Novosibirsk, Russia

³ University of Toronto, Toronto, Canada

⁴ Scientific Research Institute of Neurosciences and Medicine, Novosibirsk, Russia

✉ amstis@yandex.ru

Abstract. The incidence of autistic spectrum disorders (ASD) constantly increases in the world. Studying the mechanisms underlying ASD as well as searching for new therapeutic targets are crucial tasks. Many researchers agree that autism is a neurodevelopmental disorder. *Clstn2*-KO mouse strain with a knockout of calyntenin 2 gene (*Clstn2*) is model for investigating ASD. This study aims to evaluate the social-conditioned place preference as well as density of dopaminergic (DA) neurons in the ventral tegmental area (VTA), which belongs to the brain reward system, in the males of the *Clstn2*-KO strain using wild type C57BL/6J males as controls. Social-conditioned place preference test evaluates a reward-dependent component of social behavior. The results of this test revealed differences between the *Clstn2*-KO and the control males, as the former did not value socializing with the familiar partner, spending equal time in the isolation- and socializing-associated compartments. The *Clstn2*-KO group entered both compartments more frequently, but spent less time in the socializing-associated compartment compared to the controls. By contrast, the control males of the C57BL/6J strain spent more time in socializing-associated compartment and less time in the compartment that was associated with loneliness. At the same time, an increased number of DA and possibly GABA neurons labeled with antibodies against the type 2 dopamine receptor as well as against tyrosine hydroxylase were detected in the VTA of the *Clstn2*-KO mice. Thus, a change in social-conditioned place preference in *Clstn2*-KO mice as well as a higher number of neurons expressing type 2 dopamine receptors and tyrosine hydroxylase in the VTA, the key structure of the mesolimbic dopaminergic pathway, were observed.

Key words: *Clstn2*-KO mice; social behavior; brain; ventral tegmental area; dopaminergic neurons.

For citation: Rozhkova I.N., Okotrub S.V., Brusentsev E.Yu., Uldanova K.E., Chuyko E.A., Naprimerov V.A., Lipina T.V., Amstislavskaya T.G., Amstislavsky S.Ya. Alterations in the social-conditioned place preference and density of dopaminergic neurons in the ventral tegmental area in *Clstn2*-KO mice. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(2):177-184. DOI 10.18699/VJGB-23-14

Изменения в социальном предпочтении места и плотность дофаминергических нейронов в вентральном тегментуме у *Clstn2*-КО мышей

И.Н. Рожкова¹, С.В. Окоотруб¹, Е.Ю. Брусенцев¹, Е.Е. Ульданова¹, Э.А. Чуйко¹, В.А. Напримеров^{1,2}, Т.В. Липина³, Т.Г. Амстиславская⁴, С.Я. Амстиславский¹ ✉

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский государственный аграрный университет, Новосибирск, Россия

³ Университет Торонто, Торонто, Канада

⁴ Научно-исследовательский институт нейронаук и медицины Сибирского отделения Российской академии наук, Новосибирск, Россия

✉ amstis@yandex.ru

Аннотация. В мире наблюдается рост случаев расстройств аутистического спектра (РАС). Исследование механизмов и причин возникновения РАС, а также поиск мишеней для терапии этих расстройств являются актуальной задачей. Многие исследователи сходятся во мнении, что возникновение аутизма связано с нарушением развития нервной системы. Линия мышей *Clstn2*-КО, нокаут по гену кальсинтенин-2, полученная на основе C57BL/6J, моделирует симптомы РАС. Данное исследование было направлено на изучение у самцов *Clstn2*-КО социального предпочтения места и плотности дофаминергических нейронов в вентральном тегментуме, который представляет собой часть системы вознаграждения головного мозга, в сравнении с контрольной линией мышей C57BL/6J дикого типа. Тест «социально обусловленное предпочтение места» отражает социальное вознаграждение. Результаты этого теста по-

казали, что у самцов *Clstn2*-KO наблюдаются отклонения от контрольных мышей в социальном вознаграждении, так как они проводили одинаковое время в обоих отсеках установки, ассоциированных либо с изоляцией, либо с социализацией со знакомым партнером. При этом животные из группы *Clstn2*-KO заходили в обе части камеры значительно чаще, но проводили меньше времени в социально-ассоциированном отсеке по сравнению с контрольной группой. Самцы контрольной линии C57BL/6J, напротив, проводили больше времени в отсеке, ассоциированном с социализацией, где было взаимодействие с сородичем, и меньше в отсеке, в котором ранее особь находилась в одиночестве. В вентральном тегментуме, отвечающем за процессы, связанные с вознаграждением, у мышей *Clstn2*-KO было обнаружено повышенное число дофаминергических нейронов и, возможно, ГАМК-ергических нейронов, меченных антителами против дофаминового рецептора второго типа и тирозингидроксилазы. На основании полученных результатов можно заключить, что у мышей *Clstn2*-KO имеет место изменение значимости социального вознаграждения, а также обнаружено повышенное число нейронов, экспрессирующих дофаминовые рецепторы второго типа и тирозингидроксилазу, в одной из важных структур мезолимбического дофаминергического пути – вентральном тегментуме, который является частью системы вознаграждения.

Ключевые слова: мыши *Clstn2*-KO; социальное вознаграждение; мозг; вентральный тегментум; дофаминергические нейроны.

Introduction

Autism Spectrum Disorders (ASD) in children are characterized by impaired social interaction, low interest in peers, and difficulties in maintaining social contacts (Autism Spectrum Disorder, 2013). Many researchers agree that the ASD are developmental disorders of the nervous system (Bourgeron, 2009; Buxbaum, 2009; Marshall, Mason, 2019; Sawicka et al., 2019; Girault, Piven, 2020; Yang, Shcheglovitov, 2020). An imbalance between excitation and inhibition processes in various brain structures is characteristic for ASD (Canitano, 2007), which is caused by abnormal interactions between neurons and by impaired synaptic plasticity (Zoghbi, 2003). Mutations in the adhesion proteins genes, which play a key role in intercellular connections, including interneuronal and neuroglial contacts, have been identified in a number of ASD studies (Bourgeron, 2009; Buxbaum, 2009). In particular, impaired synthesis of neuroligins, neuroligins, contactins, and cadherins may be associated with the development of ASD in humans (Bourgeron, 2009; Buxbaum, 2009). Also, in the mouse strains modeling these disorders, the expression of genes responsible for the formation of these proteins may be impaired (Lipina et al., 2016; Zhang Q. et al., 2019).

Calcintenin-1, -2 and -3 (*Clstn1*, *Clstn2* and *Clstn3*), belonging to the cadherin family, are synaptic adhesion proteins that are able to bind Ca^{2+} ions and regulate their intracellular concentration. Of particular interest is *Clstn2*, which is specifically expressed in inhibitory interneurons (Hintsch et al., 2002) and is associated with verbal memory in adolescents (Jacobsen et al., 2009), as well as with semantic and cognitive characteristics in the elderly (Laukka et al., 2013). Moreover, genetic analysis of gene copy number variation in autistic patients revealed a deletion of the 2nd intron of the *Clstn2* gene (AlAyadhi et al., 2016). According to The Human Protein Atlas (<https://www.proteinatlas.org/>), *Clstn2* in mice is expressed in the hippocampus and some other brain structures, including the midbrain. To study the function of this protein, a *Clstn2* knockout (*Clstn2*-KO) mouse strain based on C57BL/6J was established (Lipina et al., 2016). As we have shown earlier, the absence of *Clstn2* in mice causes a selective deficit of inhibitory interneurons in the prefrontal cortex and hippocampus (Lipina et al., 2016). This is accompanied by the manifestation of ASD-like conditions, including stereotypy, insufficient social motivation, abnormal ultrasonic vocalization (Ranneva

et al., 2017; Klenova et al., 2021), as well as morphological changes in synapses (Ranneva et al., 2020).

Previously, structural and functional disorders in the mesolimbic dopaminergic pathway, which includes the ventral tegmental area (VTA) and nucleus accumbens, were found in children with ASD, and these changes in the reward system were demonstrated to be associated with underdevelopment of social skills (Supekar et al., 2018). Studies demonstrate that synaptic proteins associated with the development of ASD (Huguet et al., 2016) play an important role in the functioning of the mesolimbic pathway of the dopaminergic (DS) and GABAergic brain systems (Hart et al., 2012; Karayannis et al., 2014), one of the key midbrain components of which is the VTA (Lammel et al., 2008; Morales, Margolis, 2017). The VTA is a key structure of the reward brain system (Sesack, Grace, 2010) and regulates behavioral response to reward/punishment, including social reinforcement (Gunaydin, Deisseroth, 2014; Saunders et al., 2015). The VTA contains the bodies of dopaminergic (DA) neurons, as well as the glutamatergic and GABAergic neurons (Saunders et al., 2015). Terminals of DA neurons of the dopamine mesolimbic pathway are characterized by co-transmission, i. e. the ability to release various neurotransmitters, in particular, dopamine, glutamate, and GABA (Root et al., 2014; Zhang S. et al., 2015; Berrios et al., 2016).

One of the theories of autism is based on the notion that social motivation is reduced in autistic persons due to the alterations in the brain reward system (Kohls et al., 2012). Although the development of subcortical neuronal mechanisms of the brain is critical within the first months of life, the brain structures involved in the reward processes, that is, in the formation and correction of behavior through positive reactions to various stimuli, are functioning during the lifespan (Kohls et al., 2012, 2014). The imbalance between social and non-social motivation is the peculiar characteristic of the reward system in autistic persons (Kohls et al., 2014). This theory assumes that the reward system in ASD subjects is hyperactive in response to interests unrelated to socialization, while disruption of social behavior associates underactivity of the brain reward system in response to socially significant stimuli (Kohls et al., 2012, 2014).

The neurobiological reward system includes DA neurons of the VTA, which have projections mainly to the nucleus

accumbens and to the prefrontal cortex, and regulates social motivation (Saunders et al., 2015). It was demonstrated that DA neurons of the reward system increase their activity during the interaction of a mouse with a relative (Solie et al., 2022). A characteristic feature of DA neurons is that they release dopamine as a neurotransmitter and also contain the enzyme tyrosine hydroxylase (TH), which is necessary for its synthesis (Morales, Margolis, 2017). The study of Lammel et al. (2008) considers two types of DA neurons in the VTA (Lammel et al., 2008). Type 1 DA neurons express TH and a dopamine receptor type 2 (D2R), and their terminals end up within the shell of the nucleus accumbens and in the dorsolateral striatum. Type 2 DA neurons express TH, but not D2R, and their endings spread to the prefrontal cortex, the core and the medial zone of the shell of the nucleus accumbens, as well as the basolateral parts of the amygdala nuclei. The D2R, which can be expressed not only on DA but also on GABAergic neurons (Lammel et al., 2008; Margolis et al., 2012; Morales, Margolis, 2017), is associated with addictive behavior in which the brain reward system, the VTA in particular, is involved (Bello et al., 2011).

Mice express social behavior in a variety of contexts, including interactions with peers of the same and the opposite sex, it is also involved in early play behavior and in mother-offspring interactions (Chen, Hong, 2018). The social-conditioned place preference test evaluates social reward in young and adult mice when a certain context is associated with positive social interaction with a familiar partner (Panksepp, Lahvis, 2007; Lipina et al., 2013; Lan et al., 2019). Based on this, we hypothesized that ASD-like social behavior may be associated with impaired functioning of one or more elements of the mesolimbic dopaminergic pathway, which plays an important role in the regulation of social preference (Gunaydin, Deisseroth, 2014).

The aim of this work was to study social reward in *Clstn2* knockout mice (*Clstn2*-KO), as well as to study the density of neurons containing D2R and TH in the VTA.

Materials and methods

Experimental animals. Seven *Clstn2*-KO males, and five wild-type (C57BL/6J) males at the age of three months were used in this study. Animals were kept in the same-sex groups of 3–5 individuals in 36 × 25 × 14 cm (length × width × height) cages, in a conventional vivarium at the Institute of Neurosciences and Medicine (Novosibirsk) with sawdust bedding; 12D:12L cycle, at 20–22 °C, with free access to dry granulated food for laboratory rodents and to purified water. All studies were done in accordance with the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes (ETS No. 123).

Social-conditioned place preference test. The study was carried out as described previously (Panksepp, Lahvis, 2007; Lipina et al., 2013; Lan et al., 2019), with minor modifications. Briefly, on the eve of the experiment, the animals were kept individually for 24 hours. The experimental chamber consisted of three compartments. Two outer compartments (between which there was a third – an intermediate compartment) were separated by removable partitions. The floor, made of polypropylene, was of a different texture in the two outer compartments: rough and smooth. Before testing, the mice

were adapted to the experimental cage and the selection of the floor texture by the tested mouse was evaluated to exclude the possibility of the preference for one or another surface non-related with social interaction (session “adaptation”). The assessment was carried out visually using a stopwatch: the time spent (in seconds) in each of the compartments during 20 minutes. After adaptation, the test animals were housed in separate cages for 24 hours. Thereafter, the main experiment started.

The compartment with a rough surface was associated with social interaction, as the studied mouse was there in contact with its familiar relative of the same sex, age, and genotype, while the compartment with a smooth surface was associated with isolation, as the mouse was alone there. The procedure for establishing an “association” of the surface type with the compartment context took three days. On the first day of the experiment, the tested mouse was placed for 20 minutes in a compartment with a rough floor for socialization with a familiar partner. Three hours later, the animal was transferred to a compartment with a smooth floor, where it was left alone for 20 minutes. On the second day, the test mouse was first placed in the smooth surface compartment, where it was alone for 20 minutes, and after three hours, it was placed in the rough surface compartment with a partner for 20 minutes. On the third day of the experiment, the conditions were repeated as described for the first day. It is important to note that the familiar partner during the 20-minute socialization remained the same for each experimental animal during the three days of social reward formation. After each 20-minute session, surfaces were cleaned up with 70 % alcohol to remove odors and the surfaces were thoroughly dried. On the fourth day, the mice explored the set for 15 minutes (basic behavior session). On the fifth day of the experiment (“social reward test”), each mouse was placed in the central compartment of an empty setup, the partitions were removed to allow free movement, and the time spent in the compartments with a smooth and rough floor texture was recorded for 20 minutes. The evaluation was carried out visually using a stopwatch. The criterion for the presence of a mouse in a particular compartment was the presence of the entire body of the animal (all four paws) in the compartment, either with a rough or smooth floor covering.

Intracardiac perfusion. All animals used in the behavioral experiment were perfused the day after its completion through the circulatory system to fix the brain. Mice were anesthetized by intramuscular injection of 75 µL (per 10 g of weight) medetomidine hydrochloride (Meditin, 1 mg/ml; API-SAN, Russia) and 60 µL (per 10 g of weight) zoletil (Virbac, France). Thereafter, mice were injected through the circulatory system with 30–50 mL of phosphate-buffered saline (PBS), and then 10 % formalin solution based on PBS. After that, the brain was removed and placed in a 30 % sucrose solution in PBS at +4 °C for dehydration and further fixation for the next 3–4 weeks until the fixed material sank to the bottom of the flask. The fixed brain samples were frozen using Tissue-Tek O.C.T. (Sakura Finetek, USA) and stored at –70 °C.

Preparation of frozen brain slices. Three animals were randomly chosen for each group for the histological analysis. Frozen brain sections from each of the animals were made at a distance of –2.92 to –3.28 mm from the bregma, which

corresponds to the area of the VTA. Sections 10 μm thick were obtained on an HM550 OP Cryotome (Thermo Fisher Scientific, USA) at -25 °C and placed on Superfrost Plus, Menzel-Glaser glass slides (Thermo Fisher Scientific).

Immunohistochemical staining. Sample staining was performed according to the manufacturer's protocols with minor modifications. Briefly, after washing and exposure to Protein Block ab64226 (Abcam, UK), 50 μL of the corresponding antibody was added and left in a humid dark chamber overnight at +4 °C. The concentration of antibodies was: 1:400, 1:800 – anti-D2R-AF647 sc-5303 (Santa Cruz Biotechnology, USA) and anti-TH-AF488 MAB318-AF488 (Merck, Germany), respectively. Thereafter, the samples were washed in PBS-Tween, excess liquid was removed and placed in ProLong, Glass Antifade Mountant, Thermo P36982 (Thermo Fisher Scientific).

Analysis of the density of neurons. Images were obtained using a confocal laser scanning microscope LSM 780 (Carl Zeiss, Germany) equipped with a Plan-Apochromat 20x/0.8 M27 objective (Carl Zeiss) at the research facilities of the Center for Collective Use of Microscopic Analysis of Biological Objects of the Siberian Branch of the Russian Academy of Sciences (<https://ckp.icgen.ru/ckpmabo/>) to estimate the density of antibodies labeled neurons. The number of cells was counted manually: without the use of special programs for counting, in at least three sections per animal, in a field of view of 10000 μm² (one field of view per section). Since the VTA is a heterogeneous structure (Sanchez-Catalan et al., 2014), we took sections throughout the entire area, which correspond to a certain distance from the bregma, i.e., the rostral part of the VTA -2.92 mm, the central part -3.16 mm and caudal part -3.28 mm according to the atlas (Paxinos, Franklin, 2001). The ImageJ program was used to restrict the field of view (10000 μm²). The average number of cells from three sections for each animal and the average volume density (mm³) were calculated.

Statistical analysis. The analysis of the results was carried out using the STATISTICA v. 12.0 (StatSoft, Inc., USA) software package. All data were tested for normality using the Shapiro-Wilk W-test. Data on the behavioral parameters are presented as mean ± standard error of the mean (M ± SEM). Comparison between groups was performed using Student's *t*-test. Data on neuron density are presented as a median with the first and third quartiles – Me [Q1;Q3]. The density of labeled neurons between groups was compared using the Mann-Whitney U-test. The significance level was taken at *p* < 0.05.

Results

The preliminary testing of the control (C57BL/6J) and Clstn2-KO mice before the start of the main experiment did not reveal significant differences on the time spent in the compartments with smooth (499.8 ± 43.6 and 490.5 ± 37.0 sec, respectively) and rough (550.7 ± 17.8 and 472.8 ± 28.3 sec, respectively) floor; thus the preference for a certain compartment by mice of both studied groups was excluded. The results of the main test are presented in the Table. Mice of the control group spent more time (*p* < 0.05) in the socially associated compartment, where there was interaction with the conspecifics, compared with the compartment in which the individual was previously alone. Meanwhile, Clstn2-KO mice spent the same amount of time in both compartments. At the same time, animals from the Clstn2-KO group entered both parts of the chamber much more often (*p* < 0.001), but spent less time (*p* < 0.05) in the socially associated compartment compared to mice of the control group.

Data on the density of VTA neurons labeled with D2R and TH antibodies are presented in Figures 1 and 2. Statistical analysis revealed a higher (*p* < 0.001) density of neurons labeled with anti-D2R and anti-TH in Clstn2-KO knockout mice in the studied area as compared to controls.

Discussion

Previously, the social-conditioned place preference test was already used on mice of different strains (C57BL, DBA, BALB, Disc1-Q31L) (Panksepp, Lahvis, 2007; Lipina et al., 2013; Lan et al., 2019). It was shown that normally the animals spend more time in the compartment where they had previously contacted conspecifics; these findings are consistent with the notion that socially conditioned place preference reflects social rewards (Panksepp, Lahvis, 2007). In our study, we examined social place preference as well as the density of anti-D2R and anti-TH antibody-labeled neurons in the VTA of Clstn2-KO males and wild-type control (C57BL/6J) mice. In the social-conditioned place preference test, Clstn2-KO mice entered both compartments significantly more often, which is apparently due to their higher level of locomotor activity compared to the controls, which is consistent with the hyperactivity of these animals described in an earlier work (Lipina et al., 2016). It is possible that Clstn2-KO mice, due to their hyperactivity, were unable to form a reward caused by daily socialization with a familiar partner, and as a result, were unable to express their preference for the “social” compartment. The observed impairment of social place preference in Clstn2-KO mice is in good agreement with the previously

Social-conditioned place preference

Parameters	Strain (number of males)	
	C57BL/6J (<i>n</i> = 5)	Clstn2-KO (<i>n</i> = 7)
Isolation-associated compartment, sec	359.28 ± 71.71	457.54 ± 47.66
Social-associated compartment, sec	648.24 ± 74.56*	461.06 ± 24.55 ⁺
Isolation-associated compartment, <i>n</i>	14.80 ± 2.24	30.57 ± 2.79 ⁺⁺⁺
Social-associated compartment, <i>n</i>	15.80 ± 1.77	31.43 ± 2.52 ⁺⁺⁺

**p* < 0.05 as compared with time in isolation-associated compartment; ⁺*p* < 0.05 as compared with C57BL/6J; ⁺⁺⁺*p* < 0.001 as compared with C57BL/6J.

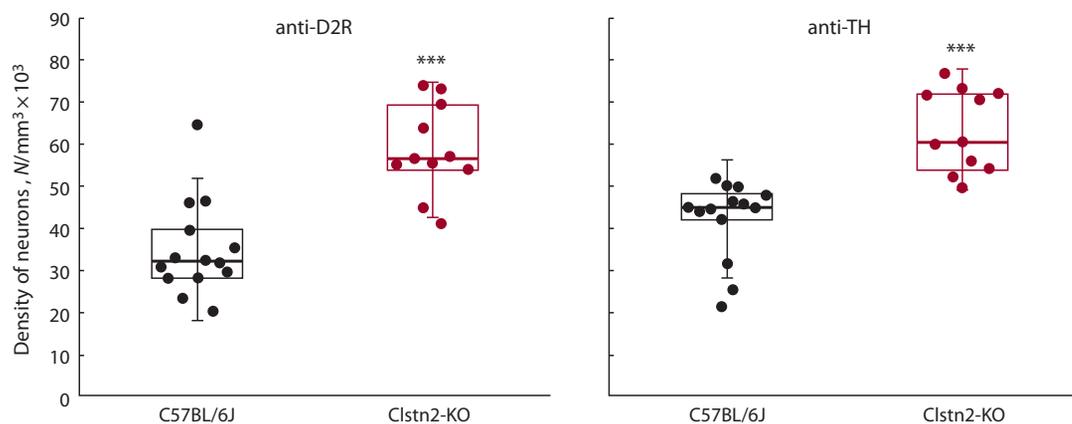


Fig. 1. The density of neurons labeled with anti-D2R, and anti-TH in the ventral tegmental area.

N – number of neurons in the field of interest. ● – density of neurons obtained per each slice. The upper and the lower bounds of the boxes correspond to the first and the third quartiles, respectively; bold horizontal line – median; vertical lines – standard deviation. ****p* < 0.001 as compared with C57BL/6J.

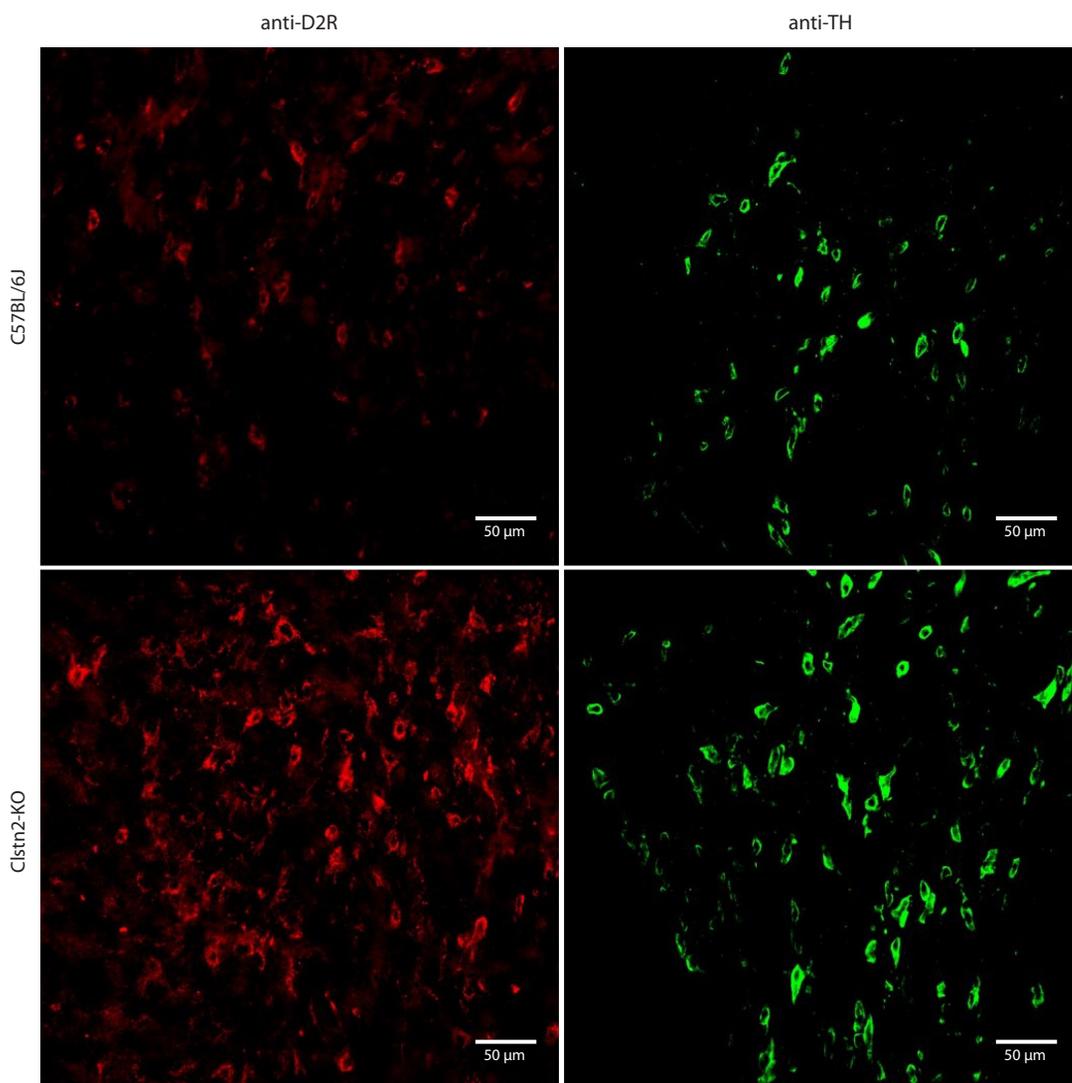


Fig. 2. The density of neurons labeled with antibodies against the second dopamine receptor (anti-D2R-AF647) and tyrosine hydroxylase (anti-TH-AF488) in males of C57BL/6J and Clstn2-KO strains in the ventral tegmental area.

reported impairment of social behavior for these mice (Raneva et al., 2017; Klenova et al., 2021).

In our work, we focused on the study of neurons expressing D2R and TH in the VTA. In both *Clstn2*-KO and C57BL/6J mice, the number of neurons labeled with anti-TH antibodies was slightly higher than the number of neurons labeled against D2R. This is apparently due to the fact that not only the DA neurons in which TH is found but also GABAergic neurons of the VTA express D2R (Lammel et al., 2008; Morales, Margolis, 2017). Meanwhile, we found more neurons with both D2R and TH in the VTA of *Clstn2*-KO mice compared to controls.

It was found that *Clstn2*-KO mice have more neurons containing D2R, as well as TH in the VTA compared to C57BL mice. Our data, as well as the results obtained on other strains of mice modeling ASD (Squillace et al., 2014; Bariselli et al., 2016, 2018; Chao et al., 2020; Tassan Mazzocco et al., 2021), indicate changes in the mesolimbic dopaminergic pathway, which also plays an important role in human ASD (Supekar et al., 2018). In particular, in the work on mice of the BTBR strain, despite the fact that they did not reveal functional changes in D1R in the striatum, a sharp decrease in D2R functions was observed upon activation of DA neurons (Squillace et al., 2014). Also, in *Shank3* and *Nlgn3*-KO mice, a decrease in the activity of DA neurons in the VTA was revealed, which caused a behavioral deficit, including alterations of social preferences compared to C57BL controls (Bariselli et al., 2016, 2018). In another study, two strains of mice modeling different forms of ASD were studied: BTBR and *Fmr1*-KO (Chao et al., 2020). A general decrease in tyrosine hydroxylase expression was found in the substantia nigra, VTA and striatum and in BTBR mice compared to C57BL mice, but not in the *Fmr1*-KO strain (Chao et al., 2020). In a study of TKO mice, which is another model of autism, no changes were found in the VTA DA neurons (Tassan Mazzocco et al., 2021).

Thus, ASD is often, but not always, associated with disturbances in the DS in the VTA. A rather unexpected result is that in *Clstn2*-KO mice the DS in the VTA is changed, but in the direction of an increase in the number of neurons containing D2R and TH. The previously described hyperactivity of *Clstn2*-KO mice (Lipina et al., 2016), which was corroborated in the current work by the increased frequency of entering of the compartments in the social-conditioned place preference test, may be associated with an increased density of neurons expressing D2R. It has been shown in the hyperactive *Coloboma* mice, that knockout of the D2R dopamine receptor gene resulted in a decrease in locomotor activity compared to controls (Fan et al., 2010). Based on this, one may assume that the increase in neurons with D2R in the VTA of *Clstn2*-KO mice reported herein may be associated with an increased locomotor activity of these animals. It is also interesting to note that human studies have shown that nucleotide polymorphism in the *D2R* gene can be considered as a potential risk factor for the development of not only ASD, but also attention deficit hyperactivity disorder (Mariggio et al., 2021).

It was previously shown that male *Disc1*-Q31L mice with depression-like behavior, which were studied in the social-conditioned place preference test, unlike *Clstn2*-KO mice, preferred the compartment associated with isolation (Lipina

et al., 2013). It can be assumed that this test adequately assesses the alterations of social behavior different models of mental disorders. Indeed, a depressive-like state caused by a deficiency of monoamines, including DA, is characterized by a complete avoidance of social contacts, which was demonstrated for the *Disc1*-Q31L strain (Lipina et al., 2013). However, in our study on *Clstn2*-KO mice, which are a model of ASD, results of this test were different. Nevertheless, we cannot completely exclude the effect of impaired spatial long-term memory observed in the Morris test in *Clstn2*-KO mice (Lipina et al., 2016) on social preference, which needs to be considered in future studies.

The data obtained may indicate a decrease in motivation for interacting with conspecifics in mice with a knockout for the *Clstn2* gene, as the mice of this strain have not demonstrated preferences to social-associated compartment. Also, changes were found in the VTA, which plays an important role in social preference (Gunaydin, Deisseroth, 2014); in this brain structure, an increased number of neurons expressing D2R and TH was found in *Clstn2*-KO mice. Thus, it can be assumed that the *Clstn2* gene plays a certain role in dopamine-dependent processes of reward and motor activity, which may be associated with changes in the density of DA neurons in the VTA.

Conclusion

The results of this study suggest that *Clstn2* knockout mice, which can be considered as a model for studying autism spectrum disorders, demonstrate a change in the perception of social reward and an increased number of neurons expressing dopamine type 2 receptors and tyrosine hydroxylase in one of the important structures of the mesolimbic dopaminergic pathway – the ventral tegmental area, which is part of the reward system.

References

- AlAyadhi L.Y., Hashmi J.A., Iqbal M., Albalawi A.M., Samman M.I., Elamin N.E., Bashir S., Basit S. High-resolution SNP genotyping platform identified recurrent and novel CNVs in autism multiplex families. *Neuroscience*. 2016;339:561-570. DOI 10.1016/j.neuroscience.2016.10.030.
- Autism spectrum disorder. In: Diagnostic and Statistical Manual of Mental Disorders, 5th Edn. Washington, DC: American Psychiatric Association, 2013;50-59.
- Bariselli S., Hornberg H., Prevost-Solie C., Musardo S., Hatstatt-Burkle L., Scheiffele P., Bellone C. Role of VTA dopamine neurons and neuroligin 3 in sociability traits related to nonfamiliar conspecific interaction. *Nat. Commun.* 2018;9(1):3173. DOI 10.1038/s41467-018-05382-3.
- Bariselli S., Tzanoulinou S., Glangetas C., Prevost-Solie C., Pucci L., Viguie J., Bezzi P., O'Connor E.C., Georges F., Luscher C., Bellone C. SHANK3 controls maturation of social reward circuits in the VTA. *Nat. Neurosci.* 2016;19(7):926-934. DOI 10.1038/nn.4319.
- Bello E.P., Mateo Y., Gelman D.M., Noain D., Shin J.H., Low M.J., Alvarez V.A., Lovinger D.M., Rubinstein M. Cocaine supersensitivity and enhanced motivation for reward in mice lacking dopamine D2 autoreceptors. *Nat. Neurosci.* 2011;14(8):1033-1038. DOI 10.1038/nn.2862.
- Berrios J., Stamatakis A.M., Katak P.A., McElligott Z.A., Judson M.C., Aita M., Rougie M., Stuber G.D., Philpot B.D. Loss of UBE3A from TH-expressing neurons suppresses GABA co-release and enhances VTA-NAc optical self-stimulation. *Nat. Commun.* 2016;7:10702. DOI 10.1038/ncomms10702.

- Bourgeron T. A synaptic trek to autism. *Curr. Opin. Neurobiol.* 2009; 19(2):231-234. DOI 10.1016/j.conb.2009.06.003.
- Buxbaum J.D. Multiple rare variants in the etiology of autism spectrum disorders. *Dialogues Clin. Neurosci.* 2009;11(1):35-43. DOI 10.31887/DCNS.2009.11.1/jdbuxbaum.
- Canitano R. Epilepsy in autism spectrum disorders. *Eur. Child. Adolesc. Psychiatry.* 2007;16(1):61-66. DOI 10.1007/s00787-006-0563-2.
- Chao O.Y., Pathak S.S., Zhang H., Dunaway N., Li J.S., Mattern C., Nikolaus S., Huston J.P., Yang Y.M. Altered dopaminergic pathways and therapeutic effects of intranasal dopamine in two distinct mouse models of autism. *Mol. Brain.* 2020;13(1):111. DOI 10.1186/s13041-020-00649-7.
- Chen P., Hong W. Neural circuit mechanisms of social behavior. *Neuron.* 2018;98(1):16-30. DOI 10.1016/j.neuron.2018.02.026.
- Fan X., Xu M., Hess E.J. D2 dopamine receptor subtype-mediated hyperactivity and amphetamine responses in a model of ADHD. *Neurobiol. Dis.* 2010;37(1):228-236. DOI 10.1016/j.nbd.2009.10.009.
- Girault J.B., Piven J. The neurodevelopment of autism from infancy through toddlerhood. *Neuroimaging Clin. N. Am.* 2020;30(1):97-114. DOI 10.1016/j.nic.2019.09.009.
- Gunaydin L.A., Deisseroth K. Dopaminergic dynamics contributing to social behavior. *Cold Spring Harb. Symp. Quant. Biol.* 2014;79: 221-227. DOI 10.1101/sqb.2014.79.024711.
- Hart A.B., Engelhardt B.E., Wardle M.C., Sokoloff G., Stephens M., de Wit H., Palmer A.A. Genome-wide association study of *d*-amphetamine response in healthy volunteers identifies putative associations, including cadherin 13 (*CDH13*). *PLoS One.* 2012;7(8): e42646. DOI 10.1371/journal.pone.0042646.
- Hintsch G., Zurlinden A., Meskenaitė V., Steuble M., Fink-Widmer K., Kinter J., Sonderegger P. The calyntenins – a family of postsynaptic membrane proteins with distinct neuronal expression patterns. *Mol. Cell. Neurosci.* 2002;21(3):393-409. DOI 10.1006/mcne.2002.1181.
- Huguet G., Benabou M., Bourgeron T. The genetics of autism spectrum disorders. In: Sassone-Corsi P., Christen Y. (Eds.) *A Time for Metabolism and Hormones. Research and Perspectives in Endocrine Interactions.* Cham: Springer, 2016;101-130. DOI 10.1007/978-3-319-27069-2_11.
- Jacobsen L.K., Picciotto M.R., Heath C.J., Mencl W.E., Gelernter J. Allelic variation of calyntenin 2 (*CLSTN2*) modulates the impact of developmental tobacco smoke exposure on mnemonic processing in adolescents. *Biol. Psychiatry.* 2009;65(8):671-679. DOI 10.1016/j.biopsych.2008.10.024.
- Karayannis T., Au E., Patel J.C., Heron D., Salomon D., Glessner J., Restituito S., Gordon A., Rodriguez-Murillo L., Roy N.C., Gogos J.A., Rudy B., Rice M.E., Karayiorgou M., Hakonarson H., Keren B., Huguet G., Bourgeron T., Hoeffler C., Tsien R.W., Peles E., Fishell G. Cntnap4 differentially contributes to GABAergic and dopaminergic synaptic transmission. *Nature.* 2014;511(7508):236-240. DOI 10.1038/nature13248.
- Klenova A.V., Volodin I.A., Volodina E.V., Ranneva S.V., Amstislavskaja T.G., Lipina T.V. Vocal and physical phenotypes of calyntenin2 knockout mouse pups model early-life symptoms of the autism spectrum disorder. *Behav. Brain Res.* 2021;412:113430. DOI 10.1016/j.bbr.2021.113430.
- Kohls G., Chevallier C., Troiani V., Schultz R.T. Social 'wanting' dysfunction in autism: neurobiological underpinnings and treatment implications. *J. Neurodev. Disord.* 2012;4(1):10. DOI 10.1186/1866-1955-4-10.
- Kohls G., Yerys B.E., Schultz R.T. Striatal development in autism: repetitive behaviors and the reward circuitry. *Biol. Psychiatry.* 2014; 76(5):358-359. DOI 10.1016/j.biopsych.2014.07.010.
- Lammel S., Hetzel A., Hackel O., Jones I., Liss B., Roeper J. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron.* 2008;57(5):760-773. DOI 10.1016/j.neuron.2008.01.022.
- Lan A., Stein D., Portillo M., Toiber D., Kofman O. Impaired innate and conditioned social behavior in adult C57Bl/6 mice prenatally exposed to chlorpyrifos. *Behav. Brain Funct.* 2019;15(1):2. DOI 10.1186/s12993-019-0153-3.
- Laukka E.J., Lovden M., Herlitz A., Karlsson S., Ferencz B., Pantzar A., Keller L., Graff C., Fratiglioni L., Backman L. Genetic effects on old-age cognitive functioning: a population-based study. *Psychol. Aging.* 2013;28(1):262-274. DOI 10.1037/a0030829.
- Lipina T.V., Fletcher P.J., Lee F.H., Wong A.H.C., Roder J.C. Disrupted-in-schizophrenia-1 Gln31Leu polymorphism results in social anhedonia associated with monoaminergic imbalance and reduction of CREB and β -arrestin-1,2 in the nucleus accumbens in a mouse model of depression. *Neuropsychopharmacology.* 2013;38(3):423-436. DOI 10.1038/npp.2012.197.
- Lipina T.V., Prasad T., Yokomaku D., Luo L., Connor S.A., Kawabe H., Wang Y.T., Brose N., Roder J.C., Craig A.M. Cognitive deficits in calyntenin-2 deficient mice associated with reduced GABAergic transmission. *Neuropsychopharmacology.* 2016;41(3):802-810. DOI 10.1038/npp.2015.206.
- Margolis E.B., Toy B., Himmels P., Morales M., Fields H.L. Identification of rat ventral tegmental area GABAergic neurons. *PLoS One.* 2012;7(7):e42365. DOI 10.1371/journal.pone.0042365.
- Mariggio M.A., Palumbi R., Vinella A., Laterza R., Petruzzelli M.G., Peschechera A., Gabellone A., Gentile O., Vincenti A., Margari L. *DRD1* and *DRD2* receptor polymorphisms: genetic neuromodulation of the dopaminergic system as a risk factor for ASD, ADHD and ASD/ADHD overlap. *Front. Neurosci.* 2021;15:705890. DOI 10.3389/fnins.2021.705890.
- Marshall J.J., Mason J.O. Mouse vs man: organoid models of brain development & disease. *Brain Res.* 2019;1724:146427. DOI 10.1016/j.brainres.2019.146427.
- Morales M., Margolis E.B. Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. *Nat. Rev. Neurosci.* 2017;18(2): 73-85. DOI 10.1038/nrn.2016.165.
- Panksepp J.B., Lahvis G.P. Social reward among juvenile mice. *Genes Brain Behav.* 2007;6(7):661-671. DOI 10.1111/j.1601-183X.2006.00295.x.
- Paxinos G., Franklin K.B.J. *The Mouse Brain in Stereotaxic Coordinates.* 2nd Edn. Academic Press, 2001.
- Ranneva S.V., Maksimov V.F., Korostyshevskaja I.M., Lipina T.V. Lack of synaptic protein, calyntenin-2, impairs morphology of synaptic complexes in mice. *Synapse.* 2020;74(2):e22132. DOI 10.1002/syn.22132.
- Ranneva S.V., Pavlov K.S., Gromova A.V., Amstislavskaya T.G., Lipina T.V. Features of emotional and social behavioral phenotypes of calyntenin2 knockout mice. *Behav. Brain Res.* 2017;332:343-354. DOI 10.1016/j.bbr.2017.06.029.
- Root D.H., Mejias-Aponte C.A., Zhang S., Wang H.L., Hoffman A.F., Lupica C.R., Morales M. Single rodent mesohabenular axons release glutamate and GABA. *Nat. Neurosci.* 2014;17:1543-1551. DOI 10.1038/nn.3823.
- Sanchez-Catalan M.J., Kaufling J., George F., Veinante P., Barrot M. The antero-posterior heterogeneity of the ventral tegmental area. *Neuroscience.* 2014;282:198-216. DOI 10.1016/j.neuroscience.2014.09.025.
- Saunders B.T., Richard J.M., Janak P.H. Contemporary approaches to neural circuit manipulation and mapping: focus on reward and addiction. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2015;370(1677): 20140210. DOI 10.1098/rstb.2014.0210.
- Sawicka K., Hale C.R., Park C.Y., Fak J.J., Gresack J.E. FMRP has a cell-type-specific role in CA1 pyramidal neurons to regulate autism-related transcripts and circadian memory. *eLife.* 2019;8: e46919. DOI 10.7554/eLife.46919.
- Sesack S.R., Grace A.A. Cortico-Basal Ganglia reward network: microcircuitry. *Neuropsychopharmacology.* 2010;35(1):27-47. DOI 10.1038/npp.2009.93.

- Solie C., Girard B., Righetti B., Tapparel M., Bellone C. VTA dopamine neuron activity encodes social interaction and promotes reinforcement learning through social prediction error. *Nat. Neurosci.* 2022;25:86-97. DOI 10.1038/s41593-021-00972-9.
- Squillace M., Doderio L., Federici M., Migliarini S., Errico F., Napolitano F., Krashia P., Di Maio A., Galbusera A., Bifone A., Scattoloni M.L., Pasqualetti M., Mercuri N.B., Usiello A., Gozzi A. Dysfunctional dopaminergic neurotransmission in asocial BTBR mice. *Transl. Psychiatry.* 2014;4(8):e427. DOI 10.1038/tp.2014.69.
- Supekar K., Kochalka J., Schaefer M., Wakeman H., Qin S., Padmanabhan A., Menon V. Deficits in mesolimbic reward pathway underlie social interaction impairments in children with autism. *Brain.* 2018;141(9):2795-2805. DOI 10.1093/brain/awy191.
- Tassan Mazzocco M., Guarnieri F.C., Monzani E., Benfenati F., Valtorta F., Comai S. Dysfunction of the serotonergic system in the brain of synapsin triple knockout mice is associated with behavioral abnormalities resembling synapsin-related human pathologies. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* 2021;105:110135. DOI 10.1016/j.pnpbp.2020.110135.
- Yang G., Shcheglovitov A. Probing disrupted neurodevelopment in autism using human stem cell-derived neurons and organoids: an outlook into future diagnostics and drug development. *Dev. Dyn.* 2020;249(1):6-33. DOI 10.1002/dvdy.100.
- Zhang Q., Wu H., Zou M., Li L., Li Q., Sun C., Xia W., Cao Y., Wu L. Folic acid improves abnormal behavior *via* mitigation of oxidative stress, inflammation, and ferroptosis in the BTBR T+ tf/J mouse model of autism. *J. Nutr. Biochem.* 2019;71:98-109. DOI 10.1016/j.jnutbio.2019.05.002.
- Zhang S., Qi J., Li X., Wang H.L., Britt J.P., Hoffman A.F., Bonci A., Lupica C.R., Morales M. Dopaminergic and glutamatergic microdomains in a subset of rodent mesoaccumbens axons. *Nat. Neurosci.* 2015;18(3):386-392. DOI 10.1038/nn.3945.
- Zoghbi H.Y. Postnatal neurodevelopmental disorders: meeting at the synapse? *Science.* 2003;302:826-830. DOI 10.1126/science.1089071.

Acknowledgements. The work was partially funded by the budget project No. FWRN-2022-0023 of the Institute of Cytology and Genetics SB RAS (Novosibirsk, Russia) and the Russian Foundation for Basic Research (No. 20-015-00162). The studies were performed using the equipment of the Center for Genetic Resources of Laboratory Animals of the Federal Research Center of the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, supported by the Ministry of Science and Higher Education of the Russian Federation (Unique identifier of the project RFMEFI62119X0023) and the Center for Microscopic Analysis of Biological Objects of the Siberian Branch of the Russian Academy of Sciences. The authors are grateful to T.N. Igonina for helpful discussions on the results and for the assistance in preparing the text of the manuscript.

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

Received July 11, 2022. Revised October 21, 2022. Accepted October 21, 2022. Published online December 12, 2022.