2019 • 23 • 8

Print ISSN 2500-0462
Online ISSN 2500-3259

Научный рецензируемый журнал

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.
Периодичность 8 выпусков в год

Учредители

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

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

Сибирское отделение Российской академии наук

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

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

Редакционный совет

Л.И. Афтанас – академик РАН, д-р мед. наук (Россия)
В.С. Баранов – чл.-кор. РАН, д-р мед. наук (Россия)
Л.А. Беспалова – академик РАН, д-р с.-х. наук (Россия)
А. Бёрнер – д-р наук (Германия)
М.И. Воевода – академик РАН, д-р мед. наук (Россия)
И. Гроссе – д-р наук, проф. (Германия)
Г.Л. Дианов – д-р биол. наук, проф. (Великобритания)
Ю.Е. Дуброва – д-р биол. наук, проф. (Великобритания)
Н.Н. Дыгало – чл.-кор. РАН, 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)
И.Е. Керкис – д-р наук (Бразилия)
А.В. Кильчевский – чл.-кор. НАНБ, д-р биол. наук (Беларусь)
С.В. Костров – чл.-кор. РАН, д-р хим. наук (Россия)
А.В. Кочетов – чл.-кор. РАН, д-р биол. наук (Россия)
Ж. Ле Гуи – д-р наук (Франция)
Б. Люгтенберг – д-р наук, проф. (Нидерланды)
В.И. Молодин – академик РАН, д-р ист. наук (Россия)
В.П. Пузырев – академик РАН, д-р мед. наук (Россия)
А.Ю. Ржецкий – канд. биол. наук, проф. (США)
И.Б. Рогозин – канд. биол. наук (США)
А.О. Рувинский – д-р биол. наук, проф. (Австралия)
Е.А. Салина – д-р биол. наук, проф. (Россия)
К.Г. Скрябин – академик РАН, д-р биол. наук (Россия)
К.В. Славин – д-р наук, проф. (США)
В.А. Степанов – чл.-кор. РАН, д-р биол. наук (Россия)
И.А. Тихонович – академик РАН, д-р биол. наук (Россия)
Е.К. Хлесткина – д-р биол. наук, профессор (Россия)
Л.В. Хотылева – академик НАНБ, д-р биол. наук (Беларусь)
Э.К. Хуснутдинова – д-р биол. наук, проф. (Россия)
М.Ф. Чернов – д-р мед. наук (Япония)
С.В. Шестаков – академик РАН, д-р биол. наук (Россия)
Н.К. Янковский – академик РАН, д-р биол. наук (Россия)

Редакционная коллегия

Т.Г. Амстиславская – д-р биол. наук (Россия)
Е.Е. Андронов – канд. биол. наук (Россия)
Ю.С. Аульченко – д-р биол. наук (Россия)
Д.А. Афонников – канд. биол. наук, доцент (Россия)
Е.В. Березиков – канд. биол. наук, проф. (Нидерланды)
С.А. Боринская – д-р биол. наук (Россия)
П.М. Бородин – д-р биол. наук, проф. (Россия)
Т.А. Гавриленко – д-р биол. наук (Россия)
В.Н. Даниленко – д-р биол. наук, проф. (Россия)
С.А. Демаков – д-р биол. наук (Россия)
Е.А. Долгих – д-р биол. наук (Россия)
Ю.М. Константинов – д-р биол. наук, проф. (Россия)
О. Кребс – д-р биол. наук, проф. (Германия)
И.Н. Лаврик – канд. хим. наук (Германия)
Д. Ларкин – д-р биол. наук (Великобритания)
Л.А. Лутова – д-р биол. наук, проф. (Россия)
В.Ю. Макеев – чл.-кор. РАН, д-р физ.-мат. наук (Россия)
М.П. Мошкин – д-р биол. наук, проф. (Россия)
Е. Песцова – д-р биол. наук (Германия)
Н.А. Проворов – д-р биол. наук, проф. (Россия)
Д.В. Пышный – чл.-кор. РАН, д-р хим. наук (Россия)
А.В. Ратушный – канд. биол. наук (США)
М.Г. Самсонова – д-р биол. наук (Россия)
Е. Туруспеков – канд. биол. наук (Казахстан)
М. Чен – д-р биол. наук (Китайская Народная Республика)
Ю. Шавруков – д-р биол. наук (Австралия)

2019 • 23 • 8

Print ISSN 2500-0462
Online ISSN 2500-3259

Scientific Peer Reviewed Journal

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

Founded in 1997
Published 8 times annually

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

Editor-in-Chief
V.K. Shumny, 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

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

Editorial council

L.I. Aftanas, Full Member of the RAS, Dr. Sci. (Medicine), Russia
V.S. Baranov, Corr. Member of the RAS, Dr. Sci. (Medicine), Russia
L.A. Bespalova, Full Member of the RAS, Dr. Sci. (Agricul.), Russia
A. Börner, Dr. Sci., Germany
M.F. Chernov, Dr. Sci. (Medicine), Japan
G.L. Dianov, Professor, Dr. Sci. (Biology), Great Britain
Yu.E. Dubrova, Professor, Dr. Sci. (Biology), Great Britain
N.N. Dygalo, Corr. Member of the RAS, Dr. Sci. (Biology), Russia
J. Le Gouis, Dr. Sci., France
I. Grosse, Professor, Dr. Sci., Germany
S.G. Inge-Vechtomov, Full Member of the RAS, Dr. Sci. (Biology), Russia
I.E. Kerkis, Dr. Sci., Brazil
E.K. Khlestkina, Professor, Dr. Sci. (Biology), Russia
L.V. Khotyleva, Full Member of the NAS of Belarus, Dr. Sci. (Biology), Belarus
E.K. Khusnutdinova, Professor, Dr. Sci. (Biology), Russia
A.V. Kilchevsky, Corr. Member of the NAS of Belarus, Dr. Sci. (Biology), Belarus
A.V. Kochetov, Corr. Member of the RAS, Dr. Sci. (Biology), Russia
S.V. Kostrov, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia
B. Lugtenberg, Professor, Dr. Sci., Netherlands
V.I. Molodin, Full Member of the RAS, Dr. Sci. (History), Russia
V.P. Puzyrev, Full Member of the RAS, Dr. Sci. (Medicine), Russia
I.B. Rogozin, Cand. Sci. (Biology), United States
A.O. Ruvinsky, Professor, Dr. Sci. (Biology), Australia
A.Yu. Rzhetsky, Professor, Cand. Sci. (Biology), United States
E.A. Salina, Professor, Dr. Sci. (Biology), Russia
S.V. Shestakov, Full Member of the RAS, Dr. Sci. (Biology), Russia
K.G. Skryabin, Full Member of the RAS, Dr. Sci. (Biology), Russia
K.V. Slavin, Professor, Dr. Sci., United States
V.A. Stepanov, Corr. Member of the RAS, Dr. Sci. (Biology), Russia
I.A. Tikhonovich, Full Member of the RAS, Dr. Sci. (Biology), Russia
M.I. Voevoda, Full Member of the RAS, Dr. Sci. (Medicine), Russia
N.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

Editorial board

D.A. Afonnikov, Associate Professor, Cand. Sci. (Biology), Russia
T.G. Amstislavskaya, Dr. Sci. (Biology), Russia
E.E. Andronov, Cand. Sci. (Biology), Russia
Yu.S. Aulchenko, Dr. Sci. (Biology), Russia
E.V. Berezikov, Professor, Cand. Sci. (Biology), Netherlands
S.A. Borinskaya, Dr. Sci. (Biology), Russia
P.M. Borodin, Professor, Dr. Sci. (Biology), Russia
M. Chen, Dr. Sci. (Biology), People's Republic of China
V.N. Danilenko, Professor, Dr. Sci. (Biology), Russia
S.A. Demakov, Dr. Sci. (Biology), Russia
E.A. Dolgikh, Dr. Sci. (Biology), Russia
T.A. Gavrilenko, Dr. Sci. (Biology), Russia
Yu.M. Konstantinov, Professor, Dr. Sci. (Biology), Russia
O. Krebs, Professor, Dr. Sci. (Biology), Germany
D. Larkin, Dr. Sci. (Biology), Great Britain
I.N. Lavrik, Cand. Sci. (Chemistry), Germany
L.A. Lutova, Professor, Dr. Sci. (Biology), Russia
V.Yu. Makeev, Corr. Member of the RAS, Dr. Sci. (Physics and Mathem.), Russia
M.P. Moshkin, Professor, Dr. Sci. (Biology), Russia
E. Pestsova, Dr. Sci. (Biology), Germany
N.A. Provorov, Professor, Dr. Sci. (Biology), Russia
D.V. Pyshnyi, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia
A.V. Ratushny, Cand. Sci. (Biology), United States
M.G. Samsonova, Dr. Sci. (Biology), Russia
Y. Shavrukov, Dr. Sci. (Biology), Australia
E. Turuspekov, Cand. Sci. (Biology), Kazakhstan

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

СОДЕРЖАНИЕ • 2019 • 23 • 8

957 ОТ РЕДАКТОРА

Генетика и селекция растений

958 ОБЗОР
Генетические механизмы акклиматизации чайного растения (*Camellia sinensis* (L.) Kuntze) к холодовому стрессу. *Л.С. Самарина, Л.С. Малюкова, М.В. Гвасалия, А.М. Ефремов, В.И. Маляровская, С.В. Лошкарёва, М.Т. Туов*

964 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Полиморфизм генов *Sdr*, регулирующих покой семян у *Triticum persicum* Vav. и *Triticum aethiopicum* Jakubz. *М.С. Баженов, Е.Д. Гусева, В.С. Рубец*

972 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Пластидный и митохондриальный геномы *Vavilovia formosa* (Stev.) Fed. и филогения родственных родов бобовых. *Н.В. Шацкая, В.С. Богданова, О.Э. Костерин, Г.В. Васильев, А.К. Кимеклис, Е.Е. Андронов, Н.А. Проворов* (на англ. языке)

981 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Экологическая стабильность боба овощного (*Vicia faba* L.) в условиях органического хозяйства. *Н.А. Георгиева, В.И. Косев* (на англ. языке)

Генетика и селекция животных

993 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Эффективность использования SNP-маркеров в гене *MSTN* в селекции кур пушкинской породы. *Н.В. Дементьева, А.Б. Вахрамеев, Т.А. Ларкина, О.В. Митрофанова*

999 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Ассоциация однонуклеотидного полиморфизма rs110861313 в межгенном районе хромосомы 23 с развитием лейкоза у крупного рогатого скота черно-пестрой породы. *Р.Б. Айтназаров, Е.В. Игнатьева, Т.А. Агаркова, Н.Г. Двоеглазов, Н.А. Осипова, В.В. Храмцов, Н.С. Юдин*

1006 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Влияние экзогенного хорионического гонадотропина человека на овуляцию у мышей. *С.Я. Амстиславский, С.В. Раннева, Д.С. Рагаева, Э.А. Чуйко, А.М. Попкова, Е.Ю. Брусенцев*

Медицинская генетика

1011 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Полиморфизм генов липидного обмена в некоторых популяциях Южной и Восточной Сибири. *Л.Э. Табиханова, Л.П. Осипова, Е.Н. Воронина, А.О. Брагин, М.Л. Филипенко*

1020 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Изменения липидного обмена мыши при одновременном воздействии антисмысловыми олигонуклеотидными производными к мРНК генов *apoB*, *PCSK9* и *apoCIII*. *С.И. Ошевский, Ю.И. Рагино, Е.В. Каштанова, Я.В. Полонская, Е.М. Стахнева, В.П. Николин, Н.А. Попова, Н.А. Колчанов, М.И. Воевода*

1026 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Ассоциация длины теломер лейкоцитов и уровня специфических антител после вакцинации против клещевого энцефалита. *Н.С. Юдин, В.А. Белявская, В.Н. Максимов, Д.Е. Иванощук, П.С. Орлов, М.И. Воевода*

Биоинформатика и системная биология

1032 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Метод главных компонент и его обобщения для последовательности любого типа (PCA-Seq). *В.М. Ефимов, К.В. Ефимов, В.Ю. Ковалева* (на англ. языке)

1037 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
In silico поиск генов, контролирующих ишемическую болезнь сердца. *И.В. Зоркольцева, Н.М. Белоногова, Г.Р. Свищёва, А.В. Кириченко, Т.И. Аксенович*

1047 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Кандидатные SNP-маркеры ревматоидного полиартрита, которые могут достоверно изменять сродство TATA-связывающего белка к промоторам генов человека. *И.В. Чадаева, Д.А. Рассказов, Е.Б. Шарыпова, И.А. Драчкова, Е.А. Ощепкова, Л.К. Савинкова, П.М. Пономаренко, М.П. Пономаренко, Н.А. Колчанов, В.А. Козлов*

Экологическая генетика

1059 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Разнообразие *mariner*-подобных элементов Orthoptera. *К.В. Устьянцев, М.Ю. Бирюков, И.С. Сухих, Н.В. Шацкая, В. Фет, А.Г. Блинов, И.Д. Конопацкая* (на англ. языке)

1067 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Внутри- и межвидовая изменчивость *Mentha arvensis* L. и *M. canadensis* L. *М.В. Семёнова, О.Л. Енина, О.В. Шелепова*

1076 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Метод газовой хроматографии-масс-спектрометрии для таксономии мискантуса. *Н.М. Слынько, Н.В. Бурмакина, О.М. Поцелуев, С.Ю. Капустянчик, Г.Ю. Галицын, Т.Н. Горячковская, Л.В. Куйбида, С.В. Шеховцов, С.Е. Пельтек, В.К. Шумный*

1082 Алфавитный указатель авторов статей, опубликованных в журнале в 2019 г.

© ИЦиГ СО РАН, 2019
© Вавиловский журнал генетики и селекции, 2019
© Сибирское отделение Российской академии наук, 2019

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2019 • 23 • 8

957 FROM THE EDITOR

Plant genetics and breeding

958 REVIEW
Genes underlying cold acclimation in the tea plant (*Camellia sinensis* (L.) Kuntze). *L.S. Samarina, L.S. Malyukova, M.V. Gvasaliya, A.M. Efremov, V.I. Malyarovskaya, S.V. Loshkareva, M.T. Tuov*

964 ORIGINAL ARTICLE
Polymorphism of *Sdr* genes regulating seed dormancy in *Triticum persicum* Vav. and *Triticum aethiopicum* Jakubz. *M.S. Bazhenov, E.D. Guseva, V.S. Rubets*

972 ORIGINAL ARTICLE
The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera. *N.V. Shatskaya, V.S. Bogdanova, O.E. Kosterin, G.V. Vasiliev, A.K. Kimeklis, E.E. Andronov, N.A. Provorov*

981 ORIGINAL ARTICLE
Ecological stability of broad bean (*Vicia faba* L.) in organic farming conditions. *N.A. Georgieva, V.I. Kosev*

Animal genetics and breeding

993 ORIGINAL ARTICLE
Efficiency of using SNP markers in the *MSTN* gene in the selection of the Pushkin breed chickens. *N.V. Dementeva, A.B. Vakhrameev, T.A. Larkina, O.V. Mitrofanova*

999 ORIGINAL ARTICLE
Single nucleotide polymorphism rs110861313 in the intergenic region of chromosome 23 is associated with the development of leukosis in the Russian Black Pied cattle. *R.B. Aitnazarov, E.V. Ignatieva, T.A. Agarkova, N.G. Dvoeglazov, N.A. Osipova, V.V. Khramtsov, N.S. Yudin*

1006 ORIGINAL ARTICLE
Effect of exogenous human chorionic gonadotropin on ovulation in mice. *S.Ya. Amstislavsky, S.V. Ranneva, D.S. Ragaeva, E.A. Chuyko, A.M. Popkova, E.Yu. Brusentsev*

Medical genetics

1011 ORIGINAL ARTICLE
Polymorphism of lipid exchange genes in some populations of South and East Siberia. *L.E. Tabikhanova, L.P. Osipova, E.N. Voronina, A.O. Bragin, M.L. Filipenko*

1020 ORIGINAL ARTICLE
Changes induced in mouse lipid metabolism by simultaneous impact of antisense oligonucleotide derivatives to *apoB*, *PCSK9*, and *apoCIII* mRNAs. *S.I. Oshevski, Y.I. Ragino, E.V. Kashtanova, Y.V. Polonskaya, E.M. Stakhneva, V.P. Nikolin, N.A. Popova, N.A. Kolchanov, M.I. Voevoda*

1026 ORIGINAL ARTICLE
Association between leukocyte telomere length and specific antibody levels after vaccination against tick-borne encephalitis. *N.S. Yudin, V.A. Belyavskaya, V.N. Maksimov, D.E. Ivanoshchuk, P.S. Orlov, M.I. Voevoda*

Bioinformatics and systems biology

1032 ORIGINAL ARTICLE
Principal component analysis and its generalizations for any type of sequence (PCA-Seq). *V.M. Efimov, K.V. Efimov, V.Y. Kovaleva*

1037 ORIGINAL ARTICLE
In silico mapping of coronary artery disease genes. *I.V. Zorkoltseva, N.M. Belonogova, G.R. Svishcheva, A.V. Kirichenko, T.I. Axenovich*

1047 ORIGINAL ARTICLE
Candidate SNP-markers of rheumatoid arthritis that can significantly alter the affinity of the TATA-binding protein for human gene promoters. *I.V. Chadaeva, D.A. Rasskazov, E.B. Sharypova, I.A. Drachkova, E.A. Oshchepkova, L.K. Savinkova, P.M. Ponomarenko, M.P. Ponomarenko, N.A. Kolchanov, V.A. Kozlov*

Ecological genetics

1059 ORIGINAL ARTICLE
Diversity of *mariner*-like elements in Orthoptera. *K. Ustyantsev, M. Biryukov, I. Sukhikh, N.V. Shatskaya, V. Fet, A. Blinov, I. Konopatskaia*

1067 ORIGINAL ARTICLE
Intra- and interspecific variability of *Mentha arvensis* L. and *M. canadensis* L. *M.V. Semenova, O.L. Enina, O.V. Shelepova*

1076 ORIGINAL ARTICLE
Gas chromatography-mass spectrometry in the taxonomy of Miscanthus. *N.M. Slynko, N.V. Burmakina, O.M. Potseluyev, S.Yu. Kapustyanchik, G.Yu. Galitsin, T.N. Goryachkovskaya, L.V. Kuybida, S.V. Shekhovtsov, S.E. Peltek, V.K. Shumny*

1082 Alphabetical author index for the list of papers published in the journal in 2019

© Institute of Cytology and Genetics SB RAS, 2019
© Vavilov Journal of Genetics and Breeding, 2019
© Siberian Branch RAS, 2019

Уважаемые коллеги, дорогие читатели! Предлагаем вашему вниманию последний в уходящем году выпуск «Вавиловского журнала генетики и селекции».

Открывает номер рубрика «Генетика и селекция растений», включающая одну обзорную и три экспериментальные статьи. В обзоре рассмотрены литературные данные о генетических механизмах и роли ключевых генов в формировании ответа растений чая на холодовой стресс. Сравнительный анализ полиморфизма генов, влияющих на предуборочное прорастание семян у тетраплоидных видов пшеницы, выполнен в оригинальном исследовании. Обращаем внимание читателей на оригинальную работу, в которой приведены полученные впервые данные полных последовательностей пластидного и митохондриального геномов представителя семейства бобовых *Vavilovia formosa*. В статье болгарских ученых изучена выраженность агрономических признаков овощного боба *Vicia faba* в условиях органического земледелия, без применения пестицидов и удобрений.

Оригинальное исследование раздела «Генетика и селекция животных» знакомит с результатами использования маркер-ориентированной селекции в птицеводстве. Авторами показано, что отбор кур пушкинской породы с помощью маркеров однонуклеотидного полиморфизма в гене миостатина на протяжении пяти поколений способствует увеличению живой массы и изменению экстерьера взрослых особей. В другом исследовании проведена проверка информативности однонуклеотидных полиморфизмов, выявленных ранее у крупного рогатого скота голштинской породы и их гибридов с различной чувствительностью к вирусу лейкоза.

Рубрика «Медицинская генетика» включает три оригинальные статьи, в первой из которых рассмотрены этнические особенности в распределении полиморфных вариантов генов липидного обмена в коренных популяциях бурят, телеутов и русских, проживающих в Восточной Сибири. В следующей работе сделана экспериментальная попытка модификации активности генов липидного обмена человека с помощью антисмысловых однонуклеотидных производных, полученных на основе мРНК этих генов. Авторы третьей статьи обсуждают проблемы выявления перспективных маркеров для прогноза иммунологической реакции пациентов в ответ на вакцинацию против клещевого энцефалита.

Раздел «Биоинформатика и системная биология» открывает статья, в которой описан новый статистический метод PCF-Seq, позволяющий для любой последовательности, в том числе молекулярной, получить главные компоненты в числовом виде и визуализировать их в виде графиков. Два других оригинальных исследования посвящены проблемам поиска генов-кандидатов, контролирующих ишемическую болезнь сердца и ревматоидный артрит. Работы выполнены с помощью полногеномного поиска ассоциаций и современных биоинформатических подходов.

Заключительный раздел знакомит читателя с оригинальными работами по экологической генетике. Рассмотрено разнообразие мобильных элементов у 16 видов из отряда насекомых Orthoptera; на основе данных секвенирования высказано предположение о горизональном переносе генетического материала этих видов в удаленные таксоны насекомых из других отрядов. В следующей работе применен комплексный подход по использованию морфологических, биохимических признаков и молекулярных маркеров для разделения видов мяты *Mentha* ssp. с перекрывающимися признаками. Авторы последней статьи представили таксономический анализ некоторых видов и гибридных форм *Miscanthus* Anderss. В качестве критериев разделения образцов использованы результаты метаболомного анализа и нуклеотидные последовательности фрагментов пластидного генома.

Редколлегия «Вавиловского журнала генетики и селекции» поздравляет читателей и авторов журнала с наступающим 2020 годом. Желаем успехов в новом году, оригинальных результатов и публикаций с высоким уровнем цитирования!

Академик В.К. Шумный

ГЕНЕТИКА И СЕЛЕКЦИЯ РАСТЕНИЙ
Обзор / Review
УДК 633.72:631.527

Вавиловский журнал генетики и селекции. 2019;23(8):958-963
DOI 10.18699/VJ19.572

Генетические механизмы акклиматизации чайного растения (*Camellia sinensis* (L.) Kuntze) к холодовому стрессу

Л.С. Самарина ✉, Л.С. Малюкова, М.В. Гвасалия, А.М. Ефремов, В.И. Маляровская, С.В. Лошкарёва, М.Т. Туов

Всероссийский научно-исследовательский институт цветоводства и субтропических культур, Сочи, Россия
✉ e-mail: q11111w2006@yandex.ru

Проведен обзор публикаций о генетических механизмах, лежащих в основе холодоустойчивости чая и других видов высших растений. Холодовой стресс, включающий охлаждение (0…+15 °C) и заморозки (< 0 °C), нарушает метаболизм в клетках и тканях и ингибирует рост растений. Показано, что в последние десятилетия достигнут большой прогресс в понимании генетических механизмов ответа растений на холодовой стресс, были открыты ключевые гены – *ICE* (inducer of *CBF* expression), *CBF* (C-repeat-binding factor), *COR* (cold-regulated genes) – и их сигнальные пути. Установлено, что накопление транскриптов *CBF* происходит уже через 15 мин после начала воздействия низких температур, +4 °C, они играют важнейшую роль в холодовой акклиматизации чайного растения. Однако существует и *CBF*-независимый путь, включающий различные гены и транскрипционные факторы, такие как *HSFC1*, *ZAT12*, *CZF1*, *PLD* (фосфолипаза D), *WRKY*, *HD-Zip*, *CsLEA*, *LOX*, *NAC*, *HSP*, которые широко распространены у растений и вовлечены в базовые механизмы устойчивости чая к холоду и заморозкам. Обнаружено повышенное накопление транскриптов генов *CsDHN1*, *CsDHN2* и *CsDHN3* у устойчивых генотипов чая в сравнении с неустойчивыми сортами в период заморозков. Определена важная роль микроРНК в механизмах ответа на охлаждение и заморозки у чая. Генетический ответ растений на охлаждение и заморозки не одинаков, и экспрессия генов ответа носит генотип-специфический характер. Приведенные результаты исследований подчеркивают необходимость дальнейшего изучения механизмов, посредством которых различные гены регулируют устойчивость чая к холодовому стрессу, для выявления генетических маркеров устойчивости.
Ключевые слова: *Camellia sinensis*; морозоустойчивость; регуляторные гены; *CBF*; транскрипционные факторы; генетические маркеры; селекция.

Для цитирования: Самарина Л.С., Малюкова Л.С., Гвасалия М.В., Ефремов А.М., Маляровская В.И., Лошкарёва С.В., Туов М.Т. Генетические механизмы акклиматизации чайного растения (*Camellia sinensis* (L.) Kuntze) к холодовому стрессу. Вавиловский журнал генетики и селекции. 2019;23(8):958-963. DOI 10.18699/VJ19.572

Genes underlying cold acclimation in the tea plant (*Camellia sinensis* (L.) Kuntze)

L.S. Samarina ✉, L.S. Malyukova, M.V. Gvasaliya, A.M. Efremov, V.I. Malyarovskaya, S.V. Loshkareva, M.T. Tuov

Russian Research Institute of Floriculture and Subtropical Crops, Sochi, Russia
✉ e-mail: q11111w2006@yandex.ru

The article reviews the latest studies showing the diversity of genetic mechanisms and gene families underlying the increased cold and frost tolerance of tea and other plant species. It has been shown that cell responses to chilling (0…+15 °C) and freezing (< 0 °C) are not the same and gene expression under cold stress is genotype-specific. In recent decades, progress has been made in understanding the genetic mechanisms underlying the cold response of plants – *ICE1* (inducer of *CBF* expression 1), *CBF* (C-repeat-binding factor), *COR* (cold-regulated genes) pathways and signaling have been discovered. The *ICE*, *CBF* and *DHN* gene groups play a key role in the cold acclimation of the tea plant. The accumulation of *CBF* transcripts occurs after 15 min of chilling induction, and longer cold stress leads to accumulation of *CBF* transcripts. It is shown that the transcripts of the *CsDHN1*, *CsDHN2* and *CsDHN3* genes accumulate at a higher level in resistant genotypes of tea in comparison with susceptible cultivars during freezing. *CBF*-independent pathways include genes involved in metabolism and transcription factors such as *HSFC1*, *ZAT12*, *CZF1*, *PLD* (phospholipase D), *WRKY*, *HD-Zip*, *CsLEA*, *LOX*, *NAC*, *HSP*, which are widely distributed in plants and are involved in the basic mechanisms of tea resistance to cold and frost. The most recent studies show an important role of miRNA in the mechanisms of response to chilling and freezing in tea. The data obtained on different plant species may correlate with the mechanisms of frost tolerance of tea and are the basis for future studies of the signaling pathways of response to cold in the tea plant. The results of

© Самарина Л.С., Малюкова Л.С., Гвасалия М.В., Ефремов А.М., Маляровская В.И., Лошкарёва С.В., Туов М.Т., 2019

the research emphasize the need to further explore the ways in which various genes regulate the tolerance of tea to cold stress to find the molecular markers of frost tolerance.
Key words: *Camellia sinensis*; frost tolerance; regulatory genes; *CBF*; transcription factors; genetic markers; breeding.

For citation: Samarina L.S., Malyukova L.S., Gvasaliya M.V., Efremov A.M., Malyarovskaya V.I., Loshkareva S.V., Tuov M.T. Genes underlying cold acclimation in the tea plant (*Camellia sinensis* (L.) Kuntze). Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):958-963. DOI 10.18699/VJ19.572 (in Russian)

Введение

Чайное растение (*Camellia sinensis* (L.) Kuntze) служит сырьем для получения одного из наиболее распространенных в мире безалкогольных напитков. Ареал возделывания этой многолетней древесной субтропической культуры распространился от тропических до умеренных климатических регионов Земли. В более северных широтах это растение подвергается воздействию холодового стресса в период зимних заморозков и весенних резких понижений температуры, которые совпадают с периодом вегетации и первым урожаем (Vyas, Kumar, 2005; Yue et al., 2015). Холодовой стресс, включающий охлаждение (0…+15 °C) и заморозки (< 0 °C), нарушает клеточный метаболизм, вызывая снижение скорости поглощения питательных веществ, скорости фотосинтеза; ингибирование роста растений; снижение урожая и качества чайного листа, что приводит к существенным экономическим потерям (Zheng et al., 2015). При холодовом стрессе запускается ряд физиолого-биохимических и молекулярных защитных механизмов, препятствующих повреждению тканей и клеток растений.

Полное понимание этих механизмов для любой сельскохозяйственной культуры является ключом к созданию новых генотипов с повышенной устойчивостью. Поэтому проводятся многочисленные исследования в этой области, а в последние годы благодаря таким методам, как RNAseq, накоплен очень ценный материал, касающийся генетических механизмов устойчивости чая. Однако многие вопросы еще не выяснены, в частности различия механизмов устойчивости между генотипами, специфика механизмов ответа на охлаждение и заморозки, а также посттрансляционные процессы у чая (Zheng et al., 2015).

Морозоустойчивые сорта чая, как правило, имеют в основе своей устойчивости различные морфологические и физиолого-биохимические механизмы, обеспечивающие устойчивость к неблагоприятным абиотическим факторам. Некоторые сорта обладают более высокой концентрацией осмолитов в цитоплазме в период заморозков, другие характеризуются особыми приспособлениями на морфологическом уровне, а часть сортов включает механизмы адаптации на физиологическом уровне. При этом есть сорта, которые могут сочетать в себе сразу несколько механизмов, обеспечивающих их повышенную устойчивость (Dr. Janaki Mohotti, personal communication). В ряде исследований было показано, что экспрессия генов чая при холодовом ответе и механизмы устойчивости качественно отличаются у неустойчивых и устойчивых сортов, т. е. носят генотип-специфический характер (Pennycooke et al., 2008; Eriksson, Webb, 2011). Другие исследователи приходят к заключению, что у неустойчивых сортов при холодовом стрессе индуцируются те же процессы, что и у морозоустойчивых, но значительно медленнее (Zhang et al., 2011). Однако в работах по функциональной геномике чая обычно используют только два генотипа, один устойчивый и один неустойчивый, и их транскриптом сравнивают в условиях холодового стресса, но такой подход не дает в полной мере выявить различия механизмов устойчивости у разных сортов.

В настоящей статье была поставлена цель – обобщить результаты исследований генетических механизмов устойчивости чая и других древесных культур к низкотемпературному стрессу для понимания главных факторов, лежащих в основе этого ценного признака.

Генетические механизмы устойчивости растений чая к холодовому стрессу

В последние десятилетия достигнут большой прогресс в понимании генетических механизмов, обуславливающих ответ растений на холодовой стресс. Выявлены основные гены-регуляторы ответа: *COR* (cold-responsive), *KIN* (cold-induced), *LTI* (low temperature induced) и *RD* (responsive to dehydration). Продукты этих генов можно классифицировать на две группы: первая включает белки *LEA* (late embryogenesis abundant proteins), *HSP* (heat shock proteins), антифриз-белки, белки транспорта липидов, дегидрины и осморегуляторы (сахара, свободные стеролы, раффиноза, глюкозиды, пролин, глицин бетаин); вторая – различные транскрипционные факторы, регулирующие передачу сигналов и экспрессию генов ответа на холод (Megha et al., 2018). Семейство генов *ICE* (inducer of *CBF* expression) – *CBF* (C-repeat-binding factor), *COR* (cold-regulated genes) – впервые изучили на модельном травянистом растении *Arabidopsis thaliana* (Thomashow, 1999), а затем его роль в ответе на различные виды стресса (холод, мороз, засуха и т. д.) была подтверждена у ряда древесных культур (Kitashiba et al., 2004; El Kayal et al., 2006), в том числе и у чая (Wang et al., 2012).

У чайного растения определено несколько ключевых генов, участвующих в процессах адаптации к низкотемпературному стрессу: *CsICE1* (FE861156), *CsCBF1* (EU563238), *CsCBF2* (KC702795) (Wang et al., 2012; Yuan et al., 2013). При сравнении ответных реакций контрастных по устойчивости к холоду сортов чая была установлена экспрессия различных генов: *CsCBF1* – у устойчивого к холоду сорта чая, *CsCBF2* – у неустойчивого (Wang et al., 2012; Ding et al., 2015). При этом высокий уровень экспрессии гена *CsCBF1* отмечен в результате подавления экспрессии *CsCBF2*, что свидетельствовало об антагонизме действия этих генов. Наряду с этим отсутствовали различия между сортами в экспрессии *CsICE1*, гена, участвующего в базовой регуляции морозостойкости растений.

Транскрипция генов группы *CBF*, в отличие от генов *ICE*, начиналась уже при +4 °C (Wang et al., 2012), что свидетельствовало о важной роли первых в процессе акклиматизации к холоду, а гены *ICE*, вероятно, больше отвечают за базовую морозостойкость растений. Возможно, что существуют какие-то *ICE*-подобные гены, чувствительные к изменению температуры и влияющие на транскрипцию *CBF1* и *CBF2*. Основная функция *CBF*, по-видимому, заключается в том, чтобы опосредовать адаптивные ответы при холодовом воздействии, т. е. вызвать изменения, имеющие решающее значение для повышения устойчивости к холодовому стрессу (Zhao et al., 2016; Zhao, Zhu, 2016). Выявлено, что накопление транскриптов *CBF* происходит уже через 15 мин после начала воздействия низких температур +4 °C, а отрицательные температуры приводят к дальнейшему их накоплению (Hua, 2016).

Гены группы *CBF* не только участвуют в повышении устойчивости к холоду, но также играют важную роль в ответе на другие абиотические факторы, в частности осмотический стресс (засуха, засоление и заморозки), приводящий к водному дефициту и повреждению клеточных мембран (Chinnusamy et al., 2003; Jia et al., 2016; Yin et al., 2016; Zhao et al., 2016).

Установлено, что мутанты арабидопсиса с гиперэкспрессией генов *CBF* имеют существенные различия по фенотипу. Отмечено ингибирование роста растений при повышении экспрессии этой группы генов (Achard et al., 2008). Y. Yin с коллегами (2016) выявили, что гиперэкспрессия гена *CsCBF3* у арабидопсиса выражалась повышенной фотосинтетической активностью, карликовым фенотипом и темно-зеленым цветом листьев, что подтверждено другими авторами (Welling, Palva, 2006; Sharabi-Schwager et al., 2010).

До недавнего времени не представлялось возможным определить роль каждого отдельного гена группы *CBF*, так как они расположены последовательно один за другим, необходимо получить стабильные нулевые мутанты отдельно по этим генам для понимания роли каждого из них. Однако с появлением системы геномного редактирования CRISPR/Cas9 удалось определить вклад каждого из трех генов *CBF* в механизмы холодоустойчивости (Ban et al., 2017). С применением геномного редактирования было выявлено, что выключение всех трех генов приводило к существенному повышению чувствительности растений к заморозку у *A. thaliana*. Главную роль в механизмах холодоустойчивости у *A. thaliana* из трех генов играет *CBF2* (Zhao et al., 2016; Zhao, Zhu, 2016). Однако у чая было показано, что из трех генов *CBF* наибольший вклад в механизмы холодоустойчивости вносит *CBF1* (Ban et al., 2017).

Тем не менее мутанты, у которых выключены все три гена *CBF*, не теряют полностью способности к холодовой акклиматизации, что указывает на наличие *CBF*-независимого пути, вносящего вклад в механизмы холодоустойчивости (Jia et al., 2016; Zhao, Zhu, 2016).

Важное значение в ответе на холодовой стресс имеет группа генов *DHNs*, кодирующих белки-дегидрины, которые могут действовать как криопротекторы, молекулярные шапероны, а также антиоксиданты (Ban et al., 2017). Транскрипция генов *DHNs* коррелирует с повышенной холодостойкостью наряду с генами *CBF* (Paul, Kumar, 2013; Li Y.Y. et al., 2016). У чая были выделены гены *CsDHN1* (GQ228834.1), *CsDHN2* (FJ436978) и *CsDHN3* (KY270880) (Li Y.Y. et al., 2016) и показано, что их транскрипты накапливаются в большей степени у устойчивых генотипов чая, в сравнении с неустойчивыми сортами, в период заморозков (Li Y.Y. et al., 2016). Недавние исследования подтвердили, что ключевую роль у холодоустойчивых сортов чая играет высокая экспрессия *CsCBF1* в сочетании с генами группы *CsDHN* и накоплением сахарозы.

CBF-независимый путь может также включать другие гены метаболизма и транскрипционные факторы, такие как *HSFC1*, *ZAT12* и *CZF1* (Vogel et al., 2005; Park et al., 2015), *PLD* (фосфолипаза D) (Li et al., 2004), *WRKY* (Wang et al., 2016b), *HD-Zip* (Shen et al., 2018), *LOX* (Zhu et al., 2018), *LEA* (Wang et al., 2018), *NAC* (Wang et al., 2016a), которые играют важную роль в регуляции ответа на холодовой стресс. Семейства *CPK* (calcium dependent protein kinases) и *CIPK* (*CBL*-interacting protein kinases) – гены протеинкиназ – также рассматриваются как положительные регуляторы ответа при заморозках, поскольку экспрессия этих генов повышалась в период заморозков у *Camellia japonica* (Li Q. et al., 2016). Имеются сообщения о том, что у чая при холодовом ответе гены флавоноидного пути активизируются в большей степени, чем гены биосинтеза антоцианов (Wang et al., 2009).

Транскриптомный анализ обнаружил 33 транскрипционных фактора *HD-Zip* (гомеодоменсодержащих белков растений), участвующих в устойчивости к холоду чайного растения сорта Longjing43. Выявлено, что устойчивость к холоду может быть связана с генами *Cshdz* (Shen et al., 2018). Ряд транскрипционных факторов *bZIP* (с доменом типа «лейциновая застежка-молния», как правило, регулирует ответ на абсцизовую кислоту и гиббереллины) при холодовом стрессе у арабидопсиса, риса и сои и может вносить вклад в холодовой ответ у чайного растения. Гиперэкспрессия генов *GmbZIP1* и *AtABF3* из этой группы приводила к повышению у растений разных видов устойчивости к холоду, заморозкам, высоким температурам, окислительному стрессу и засухе (Kim et al., 2004; Gao et al., 2011). У чая ген *CsbZIP6* напрямую участвует в холодовом ответе, в то время как ген *AtbZIP63* может быть включен в регуляторные пути и его роль требует дальнейшего изучения. У чайного растения 18 генов *CsbZIP* также экспрессируются в ответ на холод, засуху, засоление и другие стрессы (Cao et al., 2015), однако экспрессия этих генов, как и многих других, находится под сезонной регуляцией циркадных ритмов, которые в большой степени влияют на ответ клеток на холодовой стресс (Wang et al., 2017).

Многие транскрипционные факторы семейства *WRKY* содержат так называемые цинковые пальцы и важны в регуляции ответа на низкие и высокие температуры, а также солевой стресс и засуху. У чайного растения был выделен новый ген этого семейства, *CsWRKY2*, экспрессия которого повышалась при холодовом стрессе (+4 °C) параллельно с повышением концентрации АБК (Wang et al., 2016b). Установлено также, что ингибитор биосинтеза АБК снижал экспрессию *CsWRKY2*, на основании чего

был сделан вывод о важности этого гена в сигнальных путях с абсцизовой кислотой.

Группа белков *LEA* (белки позднего эмбриогенеза животных и растений) – большое и разнообразное семейство полипептидов, играющих значительную роль в регуляции роста и развития растений. При холодовом стрессе у чая экспрессия генов этой группы повышалась, всего было выявлено 33 гена *CsLEA* (Wang et al., 2018).

Семейство белков NAC состоит из нескольких транскрипционных факторов и регулирует различные биологические процессы (передачу ауксинового сигнала и развитие меристем), в том числе адаптацию к различным стрессам. Было установлено, что экспрессия генов *ANAC019*, *ANAC055*, *ANAC072/RD26* у арабидопсиса индуцируется засухой, засолением, холодом и повышением содержания абсцизовой кислоты (Wang et al., 2016a). У двух изученных сортов чайного растения (Huangjinya – чувствительный и Yingshuang – устойчивый к холоду) при холодовом стрессе происходило усиление экспрессии большинства генов *CsNAC*. Экспрессия таких генов, как *CsNAC2*, *CsNAC9*, *CsNAC12*, *CsNAC30*, *CsNAC44* и *CsNAC45*, постепенно повышалась и достигала максимума через 12–24 ч при +4 °C (Wang et al., 2016a). Экспрессия *CsNAC12* и *CsNAC18* у обоих сортов возрастала с меньшей скоростью (в два раза), в отличие от других генов (Wang et al., 2016a). У устойчивого сорта Yingshuang экспрессия генов *CsNAC17* и *CsNAC26* постепенно уменьшалась при холодовом воздействии. При этом экспрессия гена *CsNAC32* снижалась у обоих сортов. Таким образом, выявленные профили экспрессии этих генов у чая позволяют предположить, что *CsNAC* контролируют комплекс регуляторных сетей генов и оказывают влияние на различные физиологические функции для адаптации к различным стрессам.

Исследования генов регуляции углеводного метаболизма в зимний период у чая показали, что в начале зимы происходит холодовая акклиматизация, сопровождающаяся быстрым распадом крахмала и постепенным повышением содержания моно- и дисахаридов (сахарозы, фруктозы, глюкозы) в цитоплазме. Это осуществляется на фоне повышения экспрессии генов углеводного обмена, таких как ген бета-амилазы (*CsBAM*), инвертазы (*CsINV5*), ген синтеза раффинозы (*CsRS2*), участвующих в метаболизме крахмала, сахарозы и раффинозы соответственно. Одновременно повышалась и экспрессия генов транспорта углеводов, за исключением генов транспортеров сахаров *CsSWEET2*, *3*, *16*, генов раннего ответа на дегидратацию *CsERD6* и транспортера инозитола *CsINT2*, экспрессия которых снижалась. Интересно, что гены гексокиназ, *CsHXK3* и *CsHXK2*, участвующие в регуляции метаболизма сахаров, на начальной стадии акклиматизации к холоду имели противоположный характер экспрессии. Экспрессия этих генов в зимний период приводит к повышению концентрации в цитоплазме простых сахаров, снижению содержания крахмала (Yue et al., 2015), соотношение которых меняется на обратное во время потепления. При акклиматизации растений чая к зимним заморозкам ряд генов, *CsSPS*, *CsRS2* и *CsINV5*, связанных с синтезом углеводов, также усиленно экспрессируется (Yue et al., 2015). У более устойчивых сортов транскрипты этих генов накапливаются в цитоплазме, а у неустойчивых экспрессия этих генов не меняется. Поэтому данные гены рассматриваются как генотип-специфичные маркеры холодостойкости и могут использоваться в селекции чая.

Белки теплового шока *HSPs* (heat shock proteins) – вездесущие белки в клетках растений, роль которых сначала сводилась к тепловому стрессу, но теперь стало известно, что они играют важную роль в ответе на разные биотические и абиотические стрессы, помогают поддерживать стабильность клеточных мембран. Установлено, что у арабидопсиса экспрессия *HSP90* существенно повышается в условиях жары, холода, засоления, воздействия тяжелых металлов, окислительного стресса (Chen et al., 2018). У чая также было выявлено 47 генов *CsHSP* из семейств *HSP90*, *HSP70* и *sHSP* и отмечена их важная функция в регуляции роста и ответа на засуху и экстремальные температуры.

В последние годы многими исследователями подтверждено, что различные микроРНК активно участвуют в защитных реакциях растений (Zheng et al., 2015). У чайного растения описаны 233 микроРНК, дифференциально экспрессирующихся при холодовом стрессе. Показано, что чем сильнее холодовое воздействие, тем большее количество микроРНК вовлечено в ответ. Определена консервативная функция большинства идентифицированных микроРНК: при разной степени холодового воздействия выявлены разные микроРНК, перекрывания механизмов ответа на холод и заморозки не происходило. Обнаружено, что в большинстве случаев количество микроРНК, образованной при холодовом стрессе, обратно пропорционально количеству целевой мРНК.

Холодовая акклиматизация влияет на состав липидов в плазматической мембране клеток (Li et al., 2008). У *Camellia japonica* в ответ на заморозки наблюдалось повышение экспрессии генов, участвующих в биосинтезе ненасыщенных жирных кислот (Li Q. et al., 2016). Ненасыщенность жирных кислот контролируется транскрипцией ключевых генов десатуразы – *FAD2-3* (Kargiotidou et al., 2008). У холодоустойчивого сорта чая при индукции стресса экспрессия гена десатуразы *CsFAD8* была выше по сравнению с неустойчивым (Wang et al., 2012).

Семейство генов липоксигеназ *LOX* (*9-LOX*, *13-LOX* и *9/13 LOX*), как известно, отвечает за катаболизм липидов, синтез оксилипина, который выполняет защитную функцию опосредованно через ален-оксид-синтетазу (AOS) и гидропероксид лиазу (HPL) для синтеза жасмоновой кислоты и C6-альдегидов (Li Q. et al., 2016). При низких температурах жасмоновая кислота может предотвратить повреждение тканей, задерживая активные формы кислорода и перекиси водорода, регулируя аскорбат-глутатионовый цикл и активируя важные гены, такие как *MYC2* (включенный в сигнальные пути абсцизовой и жасмоновой кислот), *AP2/ERF* (большое семейство транскрипционных факторов, вовлеченных в передачу этиленового сигнала и регуляцию развития цветков и семян) (Wu et al., 2015) или *WRKY* (Li et al., 2017). Было показано, что экспрессия генов *LOX* индуцируется низкими температурами, у чайного растения гены *CsLOX6* и *CsLOX7* могут участвовать в синтезе жасмоновой кислоты. Абсцизовая кислота не влияла на экспрессию генов *CsLOX1*, *CsLOX6* и *CsLOX7*, что может свидетельствовать об отсутствии

АБК-зависимого элемента (*ABREs*) в промоутерах этих генов. В условиях холодового стресса при +4 °C у чая обнаружено накопление транскриптов генов *CsLOX1*, *CsLOX6* и *CsLOX7*, при этом наибольший уровень их экспрессии отмечался через 9, 6 и 12 ч соответственно (Zhu et al., 2018). У родственного вида *Camellia japonica* обнаружен похожий механизм ответа – экспрессия генов биосинтеза жасмоновой кислоты *LOX3*, *AOS1*, *AOS2*, *AOC2*, *AOC3*, *OPR3*, *OPCL1*, *ACX2* и *ACAA1* повышалась в условиях холодового стресса, как следствие, происходило накопление свободной жасмоновой кислоты в листьях.

Для выявления корреляции между ответом на холод и ответом на заморозки у чая выполнена функциональная оценка дифференциальной экспрессии мРНК. Анализ транскриптомов чая по базам KEGG подтвердил, что при низких положительных температурах активируются процессы углеводного обмена, фосфорилирования белков, биосинтеза гликана, гормональной передачи сигнала, регуляции фотосинтеза, MAPK-опосредованной передачи сигналов, метаболизма ксенобиотиков, метаболизма глутатиона. При ответе на заморозки активируются гены фосфорилирования, ауксин-опосредованный путь передачи сигнала, гены регуляции транскрипции, углеводного метаболизма, гены регуляции цветения, а также гены метаболизма и биосинтеза гликана. При холодовом стрессе, вызванном низкими положительными температурами, происходит активация транскрипции, а в ответ на заморозки – ее подавление (Zheng et al., 2015). Было показано, что некоторые механизмы ответа чая на холод и заморозки перекрываются, но другие ключевые биологические процессы, относящиеся к передаче сигналов (MAPK-каскад, АБК-активация), процессы углеводного обмена и метаболизм активных форм кислорода (распад перекиси водорода), а также фотосинтез при этих воздействиях протекали по-разному.

Заключение

Таким образом, в настоящее время экспериментально подтверждены многие механизмы устойчивости растений чая к холодовому стрессу и определена роль ключевых генов, лежащих в основе устойчивости. Тем не менее очевидна необходимость дальнейшего изучения путей, посредством которых различные гены регулируют устойчивость чая к холодовому стрессу и установлению вклада отдельных генов в этот признак. Каков относительный вклад семейств перечисленных генов в морозоустойчивость чая? Какова регуляторная архитектура их транскрипционных факторов? И как изменения экспрессии этих генов связаны с биохимическими, биофизическими и клеточными изменениями, которые в конечном итоге приводят к морозоустойчивости сортов? Эти вопросы остаются открытыми.

Список литературы / References

Achard P., Gong F., Cheminant S., Alioua M., Hedden P., Genschik P. The cold-inducible *CBF1* factor-dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. Plant Cell. 2008;20:2117-2129.

Ban Q., Wang X., Pan C., Wang Y., Kong L., Jiang H., Xu Y., Wang W., Pan Y., Li Y., Jiang Ch. Comparative analysis of the response and gene regulation in cold resistant and susceptible tea plants. PLoS One. 2017;12(12):e0188514. DOI 10.1371/journal.pone.0188514.

Cao H., Wang L., Yue C., Hao X., Wang X., Yang Y. Isolation and expression analysis of 18 *CsbZIP* genes implicated in abiotic stress responses in the tea plant (*Camellia sinensis*). Plant Physiol. Biochem. 2015;97:432-442.

Chen J., Gao T., Wan S., Zhang Y., Yang J., Yu Y., Wang W. Genome-wide identification, classification and expression analysis of the *HSP* gene superfamily in tea plant (*Camellia sinensis*). Int. J. Mol. Sci. 2018;19:2633. DOI 10.3390/ijms19092633.

Chinnusamy V., Ohta M., Kanrar S., Lee B.H., Hong X., Agarwal M., Zhu J.K. *ICE1*: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Genes Dev. 2003;17:1043-1054.

Ding Z., Li C., Shi H., Wang H., Wang Y. Pattern of *CsICE1* expression under cold or drought treatment and functional verification through analysis of transgenic *Arabidopsis*. Genet. Mol. Res. 2015;14:11259-11270.

El Kayal W., Navarro M., Marque G., Keller G., Marque C., Teulieres C. Expression profile of *CBF*-like transcriptional factor genes from *Eucalyptus* in response to cold. J. Exp. Bot. 2006;57:2455-2469.

Eriksson M.E., Webb A.A.R. Plant cell responses to cold are all about timing. Curr. Opin. Plant Biol. 2011;14:731-737.

Gao S.Q., Chen M., Xu Z.S. The soybean *GmbZIP1* transcription factor enhances multiple abiotic stress tolerances in transgenic plants. Plant Mol. Biol. 2011;75:537-553.

Hua J. Defining roles of tandemly arrayed *CBF* genes in freezing tolerance with new genome editing tools. New Phytol. 2016;212:301-302.

Jia Y., Ding Y., Shi Y., Zhang X., Gong Z., Yang S. The *cbf*s triple mutants reveal the essential functions of *CBFs* in cold acclimation and allow the definition of CBF regulons in *Arabidopsis*. New Phytol. 2016;212:345-353.

Kargiotidou A., Deli D., Galanopoulou D., Tsaftaris A., Farmaki T. Low temperature and light regulate *delta 12* fatty acid desaturases (FAD2) at a transcriptional level in cotton (*Gossypium hirsutum*). J. Exp. Bot. 2008;59:2043-2056. DOI 10.1093/jxb/ern065.

Kim J., Kang J.Y., Kim S.Y. Over-expression of a transcription factor regulating ABA-responsive gene expression confers multiple stress tolerance. Plant Biotechnol. J. 2004;2:459-466.

Kitashiba H., Ishizaka T., Isuzugawa K., Nishimura K., Suzuki T. Expression of a sweet cherry *DREB1/CBF* ortholog in *Arabidopsis* confers salt and freezing tolerance. J. Plant Physiol. 2004;161:1171-1176.

Li L., Lu X., Ma H., Lyu D. Jasmonic acid regulates the ascorbate–glutathione cycle in *Malus baccata* Borkh. roots under low root-zone temperature. Acta Physiol. Plant. 2017;39:174.

Li Q., Lei S., Du K., Li L., Pang X., Wang Zh., Wei M., Fu S., Hu L., Xu L. RNA-seq based transcriptomic analysis uncovers α-linolenic acid and jasmonic acid biosynthesis pathways respond to cold acclimation in *Camellia japonica*. Sci. Rep. 2016;7(6):36463. DOI 10.1038/srep36463.

Li W.Q., Li M.Y., Zhang W.H., Welti R., Wang X.M. The plasma membrane-bound phospholipase D delta enhances freezing tolerance in *Arabidopsis thaliana*. Nat. Biotechnol. 2004;22:427-433. DOI 10.1038/nbt949.

Li W.Q., Wang R.P., Li M.Y., Li L.X., Wang C.M., Welti R., Wang X. Differential degradation of extraplastidic and plastidic lipids during freezing and post-freezing recovery in *Arabidopsis thaliana*. J. Biol. Chem. 2008;283:461-468. DOI 10.1074/jbc.M706692200.

Li Y.Y., Zhou Y.Q., Xie X.F., Shu X.T., Deng W.W., Jiang C.J. Cloning and transcription analysis of dehydrin gene (C*sDHN*) in tea plant (*Camellia sinensis*). J. Agric. Biotechnol. 2016;24:332-341.

Megha S., Basu U., Kav N.N.V. Regulation of low temperature stress in plants by microRNAs. Plant Cell Environ. 2018;41:1-15.

Park S., Lee C.M., Doherty C.J., Gilmour S.J., Kim Y., Thomashow M.F. Regulation of the *Arabidopsis CBF* regulon by a complex low-temperature regulatory network. Plant J. 2015;82:193-207.

Paul A., Kumar S. Dehydrin2 is a stress-inducible, whereas Dehydrin1 is constitutively expressed but up-regulated gene under varied cues

in tea [*Camellia sinensis* (L.) O. Kuntze]. Mol. Biol. Rep. 2013;40: 3859-3863. DOI 10.1007/s11033-012-2466-2.

Pennycooke J.C., Cheng H., Stockinger E.J. Comparative genomic sequence and expression analyses of *Medicago truncatula* and alfalfa subspecies falcata *COLD-ACCLIMATION-SPECIFIC* genes. Plant Physiol. 2008;146:1242-1254. DOI 10.1104/pp.107.108779.

Sharabi-Schwager M., Samach A., Porat R. Overexpression of the *CBF2* transcriptional activator in *Arabidopsis* counteracts hormone activation of leaf senescence. Plant Signal Behav. 2010;5(3):296-309.

Shen W., Li H., Teng R., Wang Y., Wang W., Zhuang J. Genomic and transcriptomic analyses of *HD-Zip* family transcription factors and their responses to abiotic stress in tea plant (*Camellia sinensis*). Genomics. 2018. DOI 10.1016/j.ygeno.2018.07.009.

Thomashow M.F. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. 1999;50:571-599.

Vogel J.T., Zarka D.G., Van Buskirk H.A., Fowler S.G., Thomashow M.F. Roles of the *CBF2* and *ZAT12* transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. Plant J. 2005;41:195-211.

Vyas D., Kumar S. Tea (*Camellia sinensis* (L.) O. Kuntze) clone with lower period of winter dormancy exhibits lesser cellular damage in response to low temperature. Plant Physiol. Biochem. 2005;43: 383-388.

Wang L., Cao H., Qian W., Yao L., Hao X., Li N., Yang Y., Wang X. Identification of a novel *bZIP* transcription factor in *Camellia sinensis* as a negative regulator of freezing tolerance in transgenic *Arabidopsis*. Ann. Bot. 2017;119:1195-1209.

Wang L., Li X., Zhao Q., Jing Sh., Chen Sh., Yuan H. Identification of genes induced in response to low-temperature treatment in tea leaves. Plant Mol. Biol. Rep. 2009;27:257-265. DOI 10.1007/s11105-008-0079-7.

Wang W., Gao T., Chen J., Yang J., Huang H., Yu Y. The late embryogenesis abundant gene family in tea plant (*Camellia sinensis*): Genome-wide characterization and expression analysis in response to cold and dehydration stress. Plant Physiol. Biochem. 2018;135:277-286. DOI 10.1016/j.plaphy.2018.12.009.

Wang Y., Jiang C.J., Li Y.Y., Wei C.L., Deng W.W. CsICE1 and CsCBF1: two transcription factors involved in cold responses in *Camellia sinensis*. Plant Cell Rep. 2012;31:27-34. DOI 10.1007/s00299-011-1136-5.

Wang Y.-X., Liu Z.-W., Wu Z.-J., Li H., Zhuang J. Transcriptome-wide identification and expression analysis of the *NAC* gene family in tea plant [*Camellia sinensis* (L.) O. Kuntze] PLoS One. 2016a; 11(11):e0166727. DOI 10.1371/journal.pone.0166727.

Wang Y., Shu Z., Wang W., Jiang X., Li D., Pan J., Li X. *CsWRKY2*, a novel *WRKY* gene from *Camellia sinensis*, is involved in cold and drought stress responses. Biol. Plant. 2016b;60:443-451. DOI 10.1007/s10535-016-0618-2.

Welling A., Palva E.T. Molecular control of cold acclimation in trees. Physiol. Plant. 2006;127:167-181.

Wu Zh., Li X., Liu Zh., Li H., Wang Y., Zhuang J. Transcriptome-based discovery of *AP2/ERF* transcription factors related to temperature stress in tea plant (*Camellia sinensis*) Funct. Integr. Genomics. 2015; 15(6):741-752. DOI 10.1007/s10142-015-0457-9.

Yin Y., Ma Q., Zhu Z., Cui Q., Chen Ch., Chen X., Fang W., Li X. Functional analysis of *CsCBF3* transcription factor in tea plant (*Camellia sinensis*) under cold stress. Plant Growth Regul. 2016;80:335. DOI 10.1007/s10725-016-0172-0.

Yuan H.Y., Zhu X.P., Zeng W., Yang H.M., Sun N., Xie S.X., Cheng L. Isolation and transcription activation analysis of the *CsCBF1* gene from *Camellia sinensis*. Acta Botanica Boreali-Occidentalia Sinica. 2013;110:147-151.

Yue C., Cao H.L., Wang L., Zhou Y.H., Huang Y.T., Hao X.Y., Wang Y.C., Wang B., Yang Y.J., Wang X.C. Effects of CA on sugar metabolism and sugar-related gene expression in tea plant during the winter season. Plant Mol. Biol. 2015;88:591-608. DOI 10.1007/s11103-015-0345-7.

Zhang L.L., Zhao M.G., Tian Q.Y., Zhang W.H. Comparative studies on tolerance of *Medicago truncatula* and *Medicago falcata* to freezing. Planta. 2011;234:445-457. DOI 10.1007/s00425-011-1416-x.

Zhao Ch., Zhang Zh., Xie Sh., Si T., Li Y., Zhu J. Mutational evidence for the critical role of CBF transcription factors in cold acclimation in *Arabidopsis*. Plant Physiol. 2016;171:2744-2759.

Zhao Ch., Zhu J. The broad roles of *CBF* genes: From development to abiotic stress. Plant Signal. Behav. 2016;11:8. DOI 10.1080/15592324.2016.1215794.

Zheng C., Zhao L., Wang Y., Shen J., Zhang Y., Jia S., Li Y., Ding Z. Integrated RNA-Seq and sRNA-Seq analysis identifies chilling and freezing responsive key molecular players and pathways in tea plant (*Camellia sinensis*). PLoS One. 2015;10(4):e0125031. DOI 10.1371/journal.pone.0125031.

Zhu J., Wang X., Guo L., Xu Q., Zhao S., Li F., Yan X., Liu Sh., Wei Ch. Characterization and alternative splicing profiles of the lipoxygenase gene family in tea plant (*Camellia sinensis*). Plant Cell Physiol. 2018;59(9):1765-1781. DOI 10.1093/pcp/pcy091.

ORCID ID
L.S. Samarina orcid.org/0000-0002-0500-1198

Благодарности. Исследования проводятся при поддержке РНФ, проект № 18-76-10001.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 06.11.2018. После доработки 21.05.2019. Принята к публикации 14.06.2019.

ГЕНЕТИКА И СЕЛЕКЦИЯ РАСТЕНИЙ
Оригинальное исследование / Original article
УДК 631.523.11:[633.112.5+633.112.9]

Вавиловский журнал генетики и селекции. 2019;23(8):964-971
DOI 10.18699/VJ19.573

Полиморфизм генов *Sdr*, регулирующих покой семян у *Triticum persicum* Vav. и *Triticum aethiopicum* Jakubz.

М.С. Баженов[1, 2] ✉, Е.Д. Гусева[1], В.С. Рубец[1]

[1] Российский государственный аграрный университет – МСХА им. К.А. Тимирязева, Москва, Россия
[2] Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии, Москва, Россия
✉ e-mail: mikhabazhenov@gmail.com

Предуборочное прорастание зерна пшеницы, периодически наблюдаемое во многих регионах возделывания этой культуры, приводит к ухудшению его продовольственных и посевных качеств. Покой семян считается основным компонентом устойчивости к предуборочному прорастанию. Это физиологическое состояние регулируется множеством генов и сильно зависит от условий окружающей среды. Один из регуляторов покоя семян злаков – ген *Sdr4* (*Seed dormancy 4*), впервые изученный у риса. У мягкой пшеницы гомологи этого гена, *TaSdr-A1* и *TaSdr-B1*, также участвуют в регуляции покоя семян. Поиск ценных аллелей генов у местных сортов и эндемичных форм считается перспективным направлением исследований, нацеленных на повышение устойчивости сельскохозяйственных культур к неблагоприятным факторам окружающей среды. В настоящем исследовании гены *Sdr* были секвенированы у нескольких образцов двух тетраплоидных видов пшеницы, имеющих ограниченные ареалы возделывания, – пшеницы карталинской (*Triticum persicum* Vav.) и пшеницы эфиопской (*Triticum aethiopicum* Jakubz.). В результате у этих видов были найдены те же варианты аллелей генов *Sdr-A1* и *Sdr-B1*, которые ранее были обнаружены у пшеницы мягкой. При этом у пшеницы карталинской встречается только аллель *Sdr-A1a*, а у пшеницы эфиопской – аллель *Sdr-A1b*. При анализе гибридов F_2, полученных от скрещивания данных тетраплоидных видов, аллель *Sdr-A1b* был связан с меньшим индексом прорастания зерна, чем аллель *Sdr-A1a*, что не согласуется с результатами предшествующих ассоциативных исследований. В промоторе гена *Sdr-B1* у изучаемых образцов были обнаружены ранее неизвестные полиморфизмы. В 3′-конце гена *TraesCS2B02G215200*, расположенного на комплементарной цепи ДНК близко к 3′-концу гена *Sdr-B1*, обнаружена делеция 16 нуклеотидов. Обсуждается возможное влияние найденных полиморфизмов на экспрессию генов *Sdr*.
Ключевые слова: предуборочное прорастание; тетраплоидные пшеницы; маркеры; межвидовая гибридизация; секвенирование.

Для цитирования: Баженов М.С., Гусева Е.Д., Рубец В.С. Полиморфизм генов *Sdr*, регулирующих покой семян у *Triticum persicum* Vav. и *Triticum aethiopicum* Jakubz. Вавиловский журнал генетики и селекции. 2019;23(8):964-971. DOI 10.18699/VJ19.573

Polymorphism of *Sdr* genes regulating seed dormancy in *Triticum persicum* Vav. and *Triticum aethiopicum* Jakubz.

M.S. Bazhenov[1, 2] ✉, E.D. Guseva[1], V.S. Rubets[1]

[1] Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, Moscow, Russia
[2] All-Russia Research Institute of Agricultural Biotechnology, Moscow, Russia
✉ e-mail: mikhabazhenov@gmail.com

Preharvest sprouting of wheat grain, sporadically observed in many regions of cultivation of this crop, leads to deterioration of its food and sowing qualities. Seed dormancy is considered to be the main component of resistance to preharvest sprouting. This physiological state of seeds is regulated by many genes, and it depends heavily on environmental conditions. One of the regulators of seed dormancy in cereals is the *Sdr4* gene (*Seed dormancy 4*), which was first studied in rice. In common wheat, the homologues of this gene (*TaSdr-A1* and *TaSdr-B1*) are also involved in the regulation of seed dormancy. The search for valuable alleles in local varieties and endemic forms is a promising area of research aimed at increasing the resistance of crops to adverse environmental factors. In this study, *Sdr* genes were sequenced in several accessions of two tetraploid wheat species with limited cultivation areas: Persian wheat (*Triticum persicum* Vav.) and Ethiopian wheat (*Triticum aethiopicum* Jakubz.). As a result, the same *Sdr-A1* and *Sdr-B1* variants that had been found in common wheat were detected in these species. The Persian wheat accessions possessed only the *Sdr-A1a* allele, while Ethiopian ones, only *Sdr-A1b*. The analysis of F_2 hybrids obtained from crossing these tetraploid species showed that the *Sdr-A1b* allele was associated with a lower germination index of grain than *Sdr-A1a*. This result was inconsistent with earlier association studies. Previously unknown polymorphisms were found in the promoter of the *Sdr-B1* gene in the studied accessions. A deletion of 16 nucleotides was detected in the 3′-terminal region of the

© Баженов М.С., Гусева Е.Д., Рубец В.С., 2019

TraesCS2B02G215200 gene, located on the complementary DNA chain close to the 3′-end of the *Sdr-B1* gene. Possible effects of the detected polymorphisms on the expression of *Sdr* genes are discussed.
Key words: preharvest sprouting; tetraploid wheats; markers; interspecific hybridization; sequencing.

For citation: Bazhenov M.S., Guseva E.D., Rubets V.S. Polymorphism of *Sdr* genes regulating seed dormancy in *Triticum persicum* Vav. and *Triticum aethiopicum* Jakubz. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):964-971. DOI 10.18699/VJ19.573 (in Russian)

Введение

Предуборочное прорастание пшеницы, приводящее к снижению урожайности и качества зерна, остается проблемой во многих регионах возделывания этой культуры по всему миру (Olaerts, Courtin, 2018).

Покой семян – основной компонент устойчивости к предуборочному прорастанию. Это особое физиологическое состояние семян, наблюдаемое после их созревания и связанное с задержкой или полным отсутствием прорастания даже в условиях достаточной для этого влажности. Покой семян контролируется множеством генов и сильно зависит от условий окружающей среды. В настоящее время локусы количественных признаков (QTL), связанные с устойчивостью к предуборочному прорастанию, были картированы практически на всех хромосомах пшеницы. Наиболее значимые локусы располагаются на хромосомах 2B, 3A и 4A. В некоторых из этих QTL были выявлены гены-кандидаты, в том числе *TaMKK3*, *TaPHS1* (*TaMFT*), *TaVp1*, *Tamyb10*, *TaSdr* (Nakamura, 2018; Vetch et al., 2018).

Sdr4 (*Seed dormancy 4*) впервые был идентифицирован как ген-кандидат в одном из основных QTL устойчивости к предуборочному прорастанию у риса. Было выявлено, что его экспрессия снижается у мутантных форм по гену *OsVp1*. Показано, что *Sdr4* положительно регулирует гены, связанные с покоем семян (*OsDOG1-like*), и отрицательно – гены, активирующиеся при прорастании (*OsGA20ox-1*, *OsEXPB3*). Белок, кодируемый геном *Sdr4*, не имеет сходства с другими белками известной функции, но обладает сигналом ядерной локализации (Sugimoto et al., 2010). Биоинформатический анализ показал возможное наличие в его структуре суперспирали и доменов цинковых пальцев (Zhang et al., 2017). Предположительно, белок SDR4 играет роль специфического регулятора покоя семян, действующего как транскрипционный фактор под управлением более общего регулятора созревания семян – белка VP1 (Sugimoto et al., 2010).

У мягкой пшеницы были клонированы гомологи *OsSdr4* – гены *TaSdr-A1* и *TaSdr-B1*, локализованные на хромосомах 2A и 2B соответственно. В последовательностях этих генов обнаружены точечные мутации. Их связь с индексом прорастания продемонстрирована на обширной коллекции китайских сортов пшеницы и рекомбинантных инбредных линиях Yangxiaomai × Zhongyou 9507. Было также показано, что аллели генов *TaSdr*, придающие более высокую устойчивость к предуборочному прорастанию, часто встречаются у китайских и японских сортов и почти не обнаруживаются у российских сортов пшеницы (Zhang et al., 2014, 2017).

Одно из направлений поиска ценных аллелей генов, связанных с устойчивостью к неблагоприятным факторам окружающей среды, – изучение диких форм, а также местных сортов культуры. С целью расширения используемого в селекции аллельного разнообразия генов *Sdr*, что особенно актуально для России, мы решили изучить эти гены у двух локально возделываемых тетраплоидных видов – пшеницы карталинской (*Triticum persicum* Vav.) и пшеницы эфиопской (*Triticum aethiopicum* Jakubz.)[1]. Пшеница карталинская – вид, выращиваемый на Кавказе, включая территории России (Дагестан), Грузии, Армении и северо-восточной Турции. Для него характерна низкая требовательность к теплу, устойчивость к грибным болезням и длительный период послеуборочного дозревания зерна. Однако этот вид пшеницы имеет низкую засухоустойчивость (Дорофеев и др., 1979). Пшеница эфиопская – голозерная тетраплоидная пшеница, встречающаяся в пределах горной Эфиопии и южной части Аравийского полуострова. Эфиопские тетраплоидные пшеницы обладают исключительным разнообразием форм. В то же время им присущи общие признаки, позволяющие выделить их в ранг отдельного вида (Дорофеев и др., 1979). Ввиду горного рельефа местности регионов происхождения пшеницы карталинской и пшеницы эфиопской климатические условия, в которых происходило их формирование, весьма разнообразны. Некоторые районы выращивания этих видов имеют влажный климат, осадки выпадают в том числе в период созревания зерна. Можно предположить, что отдельные формы пшеницы карталинской и пшеницы эфиопской могут нести в себе ценные аллели генов, положительно влияющие на устойчивость к предуборочному прорастанию зерна в колосе.

Материалы и методы

Материал пшеницы эфиопской, собранный в ходе совместной российско-эфиопской биологической экспедиции в 2012 г., был передан нам из Института общей генетики им. Н.И. Вавилова РАН (Бадаева и др., 2018). Образцы пшеницы карталинской предоставлены Всероссийским институтом растениеводства им. Н.И. Вавилова. Информация о стране происхождения всех образцов приведена в табл. 1. Выращивание проводилось в поле на делянках площадью 1 м^2 в течение 2016–2017 гг. Ежегодно осуществлялась оценка образцов по устойчивости к предуборочному прорастанию путем проращивания зерна, собранного в фазу полной спелости и определения индекса прорастания.

Для определения индекса прорастания выделенные из колоса вручную зерновки помещали в чашки Петри на два слоя фильтровальной бумаги, смоченной 8 мл дистиллированной воды. Проращивание осуществляли в термостате при постоянной температуре 24 °C, в темноте. Подсчет и удаление проросших зерен проводили ежедневно. Фильтровальную бумагу увлажняли по мере

[1] В статье используется система рода *Triticum* по В.Ф. Дорофееву (Дорофеев и др., 1979).

Таблица 1. Страна происхождения и ботанические разновидности образцов пшеницы карталинской и пшеницы эфиопской, использованных в исследовании

Номер образца[1]	Номер в каталоге ВИР	Номер образца, код точки сбора[2]	Происхождение	Ботаническая разновидность[4]
		T. persicum		
1	к-13382	–	Грузия	*rubiginosum*
2	к-13989	–	Армения	*rubiginosum*
3	к-13768	–	»	*persicum*
4	к-1694	–	Грузия	*persicum*
5	к-26828	–	Россия[3]	*persicum*
6	к-13938	–	»	*osseticum*
7	к-7106	–	Грузия	*persicum*
8	к-6429	–	»	*stramineum*
9	к-47794	–	Россия[3]	*rubiginosum*, *stramineum*
10	к-49456	–	Канада[3]	*stramineum*
		T. aethiopicum		
11	–	7480, 28а	Эфиопия	*albiviolaceum*
12	–	7610, 47а	»	*nigriviolaceum*, *pseudoschimperi*
13	–	7655, 50б	»	*harraricum*
14	–	7634, 49а	»	*purpureum*
15	–	7567, 40б	»	*rufescens*
16	–	7719, 58б	»	*brownii*
17	–	7417, 21а	»	*nigrimarginatum*
18	–	7589, 42б	»	*pseudorubescens*

[1] Номер образца в данном исследовании; [2] в совместной российско-эфиопской биологической экспедиции в 2012 г.; [3] селекционный или исследовательский материал; [4] по (Дорофеев и др., 1979).

высыхания. Проращивание длилось семь дней. Зерновки, которые не проросли за этот срок, помещали в холодильник на три дня при температуре 4 °C для снятия покоя, после чего жизнеспособные семена подсчитывали. Опыт проводили в шестикратной повторности. В одну чашку Петри помещали зерно с одного колоса. Индекс прорастания подсчитывали по формуле:

$$GI = (7 \times n_1 + 6 \times n_2 + 5 \times n_3 + ... + 1 \times n_7)/(7 \times N),$$

где n_1, n_2 ... n_7 – числа проросших зерен в первый, второй и последующие дни до седьмого дня; N – общее число жизнеспособных зерен, заложенных для проращивания. Индекс прорастания варьировал в пределах от 0 до 1, при этом чем глубже покой семян, тем ниже значение этого индекса (Walker-Simmons, 1988).

В 2016 г. были проведены межвидовые скрещивания некоторых образцов пшеницы карталинской и пшеницы эфиопской, в том числе скрещивания *T. aethiopicum* № 15 × *T. persicum* № 5 и *T. aethiopicum* № 15 × *T. persicum* № 8. В 2017 г. в поле были высеяны гибриды F_1, а зимой 2017–2018 гг. в теплице выращены гибриды F_2. Для оценки индекса прорастания половину зерен, собранных с одного растения F_2, проращивали в одной чашке Петри (однократная повторность).

Выделение ДНК проводили из высушенных листьев растений по методу P.J. Doyle (1991). Праймеры для амплификации генов *Sdr-A1* (SDR-AF, SDR-AR) и *Sdr-B1* (SDR-BF, SDR-BR), включая часть их промоторов, были подобраны специфично для субгеномов A и B пшеницы на основе последовательностей генома пшеницы IWGSC RefSeq v1.0 (Alaux et al., 2018) с применением Primer BLAST (NCBI).

Последовательности всех праймеров, использованных в исследовании, приведены в табл. 2. ПЦР осуществляли на приборе SimpliAmp Thermal Cycler (Applied Biosystems). Концентрация компонентов реакционной смеси составляла: 1× буфер, поставляемый с полимеразой (включая 2.5 ммоль/л $MgCl_2$), 0.2 ммоль/л каждого dNTP, 10 % по объему диметилсульфоксида, по 1 мкмоль/л прямого и обратного праймеров, 0.05 ед./мкл *Taq*-полимеразы (ООО «Силекс»), 0.01 ед./мкл *Pfu*-полимеразы (ООО «Силекс») и 4 нг/мкл матричной ДНК. Добавление диметилсульфоксида в реакционную смесь связано с высоким содержанием гуанина и цитозина в амплифицируемых участках ДНК. Температурные условия ПЦР были следующими: 94 °C – 5 мин; 45 циклов 94 °C – 30 с, 57 °C – 45 с и 72 °C – 2 мин; конечная элонгация 72 °C – 5 мин. Перед секвенированием продукты ПЦР очищали с помощью набора Cleanup Mini («Евроген»).

Секвенирование осуществляли с помощью набора реактивов BrilliantDye Terminator v3.1 (Nimagen), амплификатора SimpliAmp (Applied Biosystems) и прибора 3130xl Genetic Analyzer (Applied Biosystems). Для проведения реакции применяли концевые праймеры, а также дополнительные праймеры SDR-2AF, SDR-2AR (для гена *Sdr-A1*) и SDR-3BF, SDR-3BR (для гена *Sdr-B1*) (Zhang

Таблица 2. ДНК-праймеры, использованные в исследовании

Название праймера	Последовательность 5′-3′	Назначение
SDR-AF	ACAACAGGATGATACAGGGGAC	Амплификация целого гена, секвенирование
SDR-AR	AGGGAGTATAATATAATTTTGCCATCT	
SDR-BF	ACCCCTACGCATTTATACAGACT	
SDR-BR	CTGTCGTGCAATATGAGCTAGAA	
SDR-2AF*	CGTCGGCAGACATCGACTCC	Секвенирование
SDR-2AR*	GAAGCTCACTAGCTCAGAACACGC	
SDR-3BF*	CGTCAGCAGACTTCGACTCGC	
SDR-3BR*	CAAGAAGCTCACTATCTCAGAACACAA	
SDR-SNP-AF	ACCATCCACGTCGAATCCATC	Амплификация фрагментов гена для маркеров
SDR-SNP-AR	CTTGGTTTCGCACGCAGCTC	
SDR-5BF*	CGCCTACGTGTCGGCCC	
SDR-SNP-BR	CACGGCCTGTTCTGCAGGTG	
SDR-B1-InDel-F	ATGTACAAAAGCCTAGCCACAGA	

* Праймеры из (Zhang et al., 2014). Прочие праймеры – оригинальные.

et al., 2014). Для выравнивания нуклеотидных последовательностей использовали программу GeneDoc v.2.7. Анализ нуклеотидной последовательности промотора гена выполняли с помощью онлайн-сервиса PlantCARE (Lescot et al., 2002).

Праймеры для создания новых маркеров подбирали с помощью Primer BLAST (NCBI). ПЦР проводили в объеме 25 мкл. Конечная концентрация компонентов смеси аналогична вышеприведенной, но без добавления *Pfu*-полимеразы. Для выявления однонуклеотидного полиморфизма в гене *Sdr-B1* в позиции (–11) использовали праймеры SDR-5BF и SDR-SNP-BR. Температурные условия ПЦР были следующими: 95 °C – 1 мин; 40 циклов 95 °C – 45 с, 65 °C – 45 с, 72 °C – 45 с; конечная элонгация 72 °C – 2 мин. Продукты ПЦР без предварительной очистки разрезали с помощью эндонуклеазы рестрикции *Psp*CI («СибЭнзим»). В состав реакционной смеси входили буферный раствор, поставляемый с эндонуклеазой (буфер B, «СибЭнзим»), и бычий сывороточный альбумин (BSA) в рекомендуемых производителем концентрациях, а также 1 ед./мкл фермента. Объем реакционной смеси – 10 мкл. Реакция протекала при 37 °C в течение 2 ч.

Для идентификации делеции, расположенной рядом с геном *Sdr-B1* на расстоянии 449–464 пар оснований (п. о.) после стоп-кодона, использовали праймеры SDR-B1-InDel-F и SDR-BR. Температурные условия ПЦР были следующими: 95 °C – 2 мин; 35 циклов 95 °C – 30 с, 58 °C – 30 с, 72 °C – 30 с; конечная элонгация 72 °C – 2 мин.

Для обнаружения однонуклеотидного полиморфизма в гене *Sdr-A1* в позиции 643 применяли праймеры SDR-SNP-AF и SDR-SNP-AR. Температурные условия ПЦР: 95 °C – 1 мин; 10 циклов 95 °C – 45 с, 66 °C с последующим уменьшением температуры на 0.4 °C за цикл – 45 с, 72 °C – 45 с; 35 циклов 95 °C – 45 с, 62 °C – 45 с, 72 °C – 45 с; конечная элонгация 72 °C – 2 мин. Продукты ПЦР без предварительной очистки разрезали помощью эндонуклеазы рестрикции *Ple*19I («СибЭнзим»). В состав реакционной смеси входили буферный раствор, поставляемый с эндонуклеазой (буфер Y, «СибЭнзим»), разбавленный до рекомендуемой концентрации, а также 0.1 ед./мкл фермента. Объем реакционной смеси – 10 мкл. Реакцию проводили при 37 °C в течение 2 ч.

Электрофорез продуктов ПЦР и рестрикции осуществляли в 1.5 % агарозном геле с буфером TBE с добавлением бромистого этидия. Визуализацию результатов проводили с помощью системы гель-документирования GelDoc XR + Bio-Rad.

Статистический анализ данных выполняли с помощью пакета программ Statistica 6.0. Перед дисперсионным анализом значения индекса прорастания преобразовывали по формуле: $\arcsin(\sqrt{x})$. В табл. 1–3 и на рис. 1–4 даны обратно преобразованные средние значения. Для теста Манна–Уитни использовали исходные значения индекса прорастания.

Результаты

Испытуемые образцы пшеницы статистически значимо отличались друг от друга по индексу прорастания как в 2016, так и в 2017 г. (табл. 3). В среднем по годам наиболее низкий индекс прорастания имели образцы пшеницы карталинской № 2, 4, 8, 10 и образец пшеницы эфиопской № 15. Высоким индексом прорастания отличались образцы пшеницы эфиопской № 13, 16 и 18. В среднем пшеница карталинская показала меньший индекс прорастания, чем пшеница эфиопская ($p < 0.05$).

Образцы пшеницы эфиопской № 15 и 16, а также образцы пшеницы карталинской № 5 и 8 были выбраны для секвенирования генов *Sdr-A1* и *Sdr-B1* как наиболее контрастные по индексу прорастания. Полученные последовательности были депонированы в базу данных GenBank под номерами MK396766, MK396767, MK396768, MK396769.

Секвенирование гена *Sdr-A1* у перечисленных выше образцов показало, что они различаются между собой по тому же однонуклеотидному полиморфизму, который ранее был обнаружен у мягкой пшеницы (Zhang et al., 2017). У образцов пшеницы эфиопской № 15 и 16 в позиции 643 от старт-кодона в данном гене был найден нуклеотид A,

Таблица 3. Аллели генов *Sdr-A1*, *Sdr-B1* и индекс прорастания образцов пшеницы карталинской и пшеницы эфиопской

№ образца	Аллели генов		Маркер делеции, п. о.*	Индекс прорастания**		
	Sdr-A1	*Sdr-B1*		2016 г.	2017 г.	Средний
			T. persicum			
1	*a*	*a*	230	0.45^{cd}	0.19^{bc}	0.31^{abc}
2	*a*	*a*	230	0.17^{ab}	0.12^{abc}	0.14^{a}
3	*a*	*b*	214	0.62^{d}	0.35^{cd}	0.48^{abcde}
4	*a*	*b*	214	0.18^{ab}	0.08^{ab}	0.13^{a}
5	*a*	*b*	214	0.63^{de}	0.57^{def}	0.60^{abcde}
6	*a*	*b*	214	0.38^{c}	0.36^{cd}	0.37^{abcd}
7	*a*	*b*	214	0.27^{abc}	0.21^{bc}	0.24^{ab}
8	*a*	*b*	214	0.13^{a}	0.10^{ab}	0.11^{a}
9	*a*	*a*	230	0.80^{ef}	0.19^{bc}	0.49^{abcde}
10	*a*	*b*	214	0.31^{bc}	0.03^{a}	0.14^{a}
			T. aethiopicum			
11	*b*	*a*	230	0.60^{d}	0.68^{efg}	0.64^{abcde}
12	*b*	*a*	230	0.85^{f}	0.75^{efg}	0.80^{bcde}
13	*b*	*b*	230	1.00^{g}	0.83^{fgh}	0.96^{e}
14	*b*	*b*	230	0.44^{cd}	0.35^{cd}	0.39^{abcde}
15	*b*	*b*	214	0.16^{ab}	0.15^{abc}	0.15^{a}
16	*b*	*b*	230	0.90^{f}	0.97^{h}	0.94^{de}
17	*b*	*b*	230	0.45^{cd}	0.55^{de}	0.50^{abcde}
18	*b*	*b*	214	0.89^{f}	0.85^{gh}	0.87^{cde}
p – значение различий между образцами (F-тест)				< 0.001	< 0.001	< 0.001

* Маркер делеции, расположенной рядом с геном *Sdr-B1*; дан размер ПЦР-продукта (п. о.); ** значения индекса прорастания, отмеченные в каждом столбце одинаковыми буквами, не различаются значимо между собой по критерию Тьюки.

свойственный аллелю *Sdr-A1b*, а у образцов пшеницы карталинской № 5 и 8 – нуклеотид G, характерный для аллеля *Sdr-A1a*. В остальном последовательности гена *Sdr-A1* двух исследуемых видов пшеницы не различались и совпадали с референсной последовательностью генома мягкой пшеницы сорта Chinese Spring.

У прочих образцов пшеницы карталинской и пшеницы эфиопской аллели гена *Sdr-A1* были определены с помощью молекулярного маркера, разработанного ранее Y. Zhang с коллегами (2017), основанного на ПЦР и рестрикции. Обнаружено, что у всех исследованных образцов пшеницы карталинской встречается только аллель *Sdr-A1a*, а у всех образцов пшеницы эфиопской – только аллель *Sdr-A1b*.

Влияние аллелей гена *Sdr-A1* на покой семян было изучено в F_2 в двух гибридных комбинациях, полученных от межвидовых скрещиваний образцов пшеницы карталинской и пшеницы эфиопской. Молекулярный маркер, разработанный ранее Y. Zhang с коллегами (2017), предполагает амплификацию фрагмента ДНК размером 1294 п. о. Так как амплификация таких крупных фрагментов ДНК, особенно содержащих GC-богатые участки, с использованием *Taq*-полимеразы часто вызывает затруднения, нами повторно были подобраны праймеры для маркирования этого однонуклеотидного полиморфизма таким образом, чтобы минимизировать размер ПЦР-продукта. В результате были найдены специфичные для субгенома A праймеры SDR-SNP-AF и SDR-SNP-AR (см. табл. 2), дающие продукт ПЦР размером 499 п. о. Рестрикция ПЦР-продукта осуществлялась с помощью фермента *Ple*19I. При присутствии аллеля *Sdr-A1a* разрезание не происходит, а при наличии аллеля *Sdr-A1b* образуются фрагменты размером 387 и 112 п. о. (рис. 1, *a*). Испытание нового маркера на изучаемых образцах пшеницы карталинской и пшеницы эфиопской показало полное совпадение результатов анализа с данными, полученными на основе ранее разработанного маркера. Новые праймеры позволили значительно повысить надежность амплификации.С помощью нового маркера в гибридной комбинации *T. aethiopicum* № 15 × *T. persicum* № 5 было проанализировано 66 растений F_2. Из них 15 оказались гомозиготами по аллелю *Sdr-A1a*, 11 были гомозиготами по аллелю *Sdr-A1b*, а остальные 40 – гетерозиготами. Расщепление в F_2 соответствовало теоретически ожидаемому 1:2:1 ($\chi^2 = 4.125$; $p = 0.178$). В связи с тем, что в F_2 статистическое распределение индекса прорастания было далеко от нормального, мы использовали непараметрический U-критерий Манна–Уитни для оценки связи индекса с генотипами растений. При этом в расчетах использовали только гомозиготные формы. По результатам статистического теста ген *Sdr-A1* не оказал значимого влияния на индекс прорастания в F_2 *T. aethiopicum* № 15 × *T. persicum* № 5 ($p = 0.09$), но была тенденция к более низкому индексу у растений с аллелем *Sdr-A1b*.

В гибридной комбинации *T. aethiopicum* № 15 × *T. persicum* № 8 было проанализировано 83 растения F_2. Из них у 25 выявлен аллель *Sdr-A1a*, у 15 – аллель *Sdr-A1b,* а 43 были гетерозиготами. Расщепление в целом соответствовало менделевскому 1:2:1 ($\chi^2 = 2.98$; $p = 0.284$), однако, как и в описанной выше комбинации, доля гомозигот по аллелю *Sdr-A1b* была несколько ниже ожидаемой. В этой гибридной комбинации связь индекса прорастания с маркером гена *Sdr-A1* оказалась существенной ($p = 0.035$ по U-критерию Манна–Уитни). При этом растения с аллелем *Sdr-A1b* имели меньший индекс прорастания зерна, чем растения с аллелем *Sdr-A1a*. Объединение данных, полученных в двух популяциях F_2, в одну выборку подтвердило ту же тенденцию ($p = 0.012$ по U-критерию Манна–Уитни) (рис. 2).

Секвенирование гена *Sdr-B1* у образцов № 5 и 8 пшеницы карталинской и образцов № 15 и 16 пшеницы эфиопской показало, что эти образцы не отличаются друг от друга в кодирующей части данного гена. В промоторе гена у образца № 16 имеются две однонуклеотидные замены, в то время как у образцов № 5, 8 и 15 присутствует динуклеотидная инсерция CT относительно референсной последовательности Chinese Spring (рис. 3, *а*). В то же время 5′-некодирующая область гена у всех четырех образцов пшеницы несет известный ранее вариант однонуклеотидного полиморфизма в позиции (–11) от старт-кодона, характерный для аллеля *Sdr-B1b* (Zhang et al., 2014).

Маркер для идентификации аллелей гена *Sdr-B1* был также оптимизирован нами путем подбора праймеров для получения более короткого продукта ПЦР (см. рис. 1, *б*). Аллели *Sdr-B1*, идентифицированные по однонуклеотидному полиморфизму в позиции (–11), не показали статистически значимой связи с индексом прорастания в изучаемой коллекции образцов двух видов пшеницы ($p = 0.9$ по критерию Фишера).

Рядом с 3′-концом гена *Sdr-B1* секвенированная нами последовательность частично захватывает 3′-нетранслируемую область лежащего близко на комплементарной цепи ДНК гена *TraesCS2B02G215200*, предположительно кодирующего субъединицу 18 кДа НАДН-убихинон-оксидоредуктазы. В 3′-нетранслируемой области гена *TraesCS2B02G215200* у образцов № 5, 8 и 15 присутствует 16-нуклеотидная делеция, которая отсутствует у образца № 16 и в референсной последовательности мягкой пшеницы Chinese Spring (см. рис. 3, *б*).

Ввиду того, что эта 16-нуклеотидная делеция может оказаться более простым в применении сцепленным маркером аллелей гена *Sdr-B1*, для ее идентификации нами были подобраны фланкирующие праймеры (SDR-B1-InDel-F и SDR-BR). Электрофорез продуктов ПЦР с указанными праймерами позволяет идентифицировать два варианта фрагментов размером 230 и 214 п. о. Присутствует также дополнительный фрагмент большего размера (примерно 266 п. о.), вероятно, получаемый с хромосомы 2A (рис. 4). В результате испытания маркера на образцах коллекции обнаружено, что у пшеницы карталинской наличие фрагмента размером 230 п. о. совпадает с присутствием аллеля *Sdr-B1a*, который в предыдущих исследованиях был связан с низким индексом прорастания. Среди образцов эфиопской пшеницы такая закономерность нарушается. В целом по двум видам пшеницы образцы с коротким маркерным фрагментом (214 п. о.) имели в среднем низкий (0.35), а с длинным фрагментом (230 п. о.) – более высокий индекс прорастания (0.57). Тем не менее статистически значимой связи между наличием этой делеции и индексом прорастания в испытуемой коллекции не было обнаружено ($p = 0.1$ по критерию Фишера для средних за два года испытания).

Обсуждение

Ранее исследователи отмечали, что пшеница карталинская обладает хорошей устойчивостью к предуборочному прорастанию (Дорофеев и др., 1979). Это подтвердили и наши результаты. В течение двух лет в среднем она показывала меньший индекс прорастания, чем пшеница эфиопская.

В предыдущих исследованиях аллели *Sdr-A1a* и *Sdr-B1a* были ассоциированы с устойчивостью, в то время как аллели *Sdr-A1b* и *Sdr-B1b* – с неустойчивостью к предуборочному прорастанию (Zhang, 2014, 2017). У всех образ-

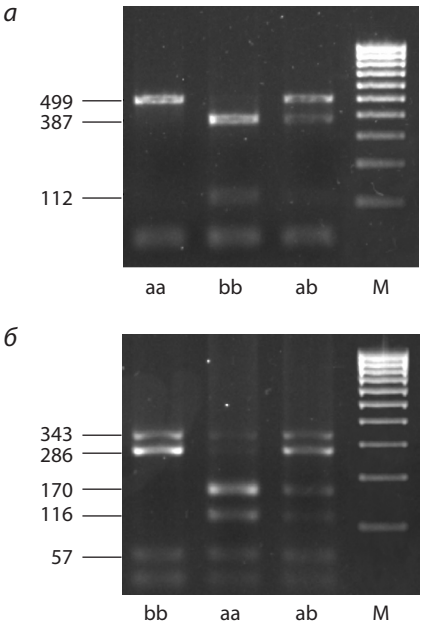


Рис. 1. Новые маркеры однонуклеотидных полиморфизмов (SNP) в генах *Sdr* пшеницы.

а – маркер SNP A643G гена *Sdr-A1*. Пример электрофореза продуктов ПЦР, полученных с использованием праймеров SDR-SNP-AF/SDR-SNP-AR и разрезанных эндонуклеазой *Ple*19I: aa – гомозигота по аллелю *Sdr-A1a*, bb – гомозигота по аллелю *Sdr-A1b*, ab – гетерозигота. Указаны размеры продуктов (п. о.); *б* – маркер SNP A(-11)G в 5′-нетранслируемой области гена *Sdr-B1*. Пример электрофореза продуктов ПЦР, полученных с использованием праймеров SDR-5BF и SDR-SNP-BR и разрезанных эндонуклеазой *Psp*CI: bb – гомозигота по аллелю *Sdr-B1b*, aa – гомозигота по аллелю *Sdr-B1a*, ab – гетерозигота. Фрагмент 343 п. о. соответствует неразрезанному продукту ПЦР, который остается в результате низкой эффективности отрезания короткого фрагмента размером 57 п. о. M – маркер длин ДНК M-100 («Синтол»).

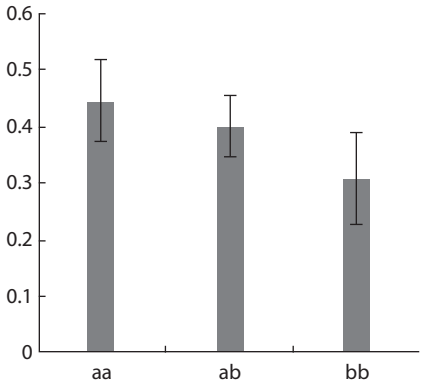


Рис. 2. Индекс прорастания зерен, собранных с растений F_2 с различным генотипом *Sdr-A1*: aa – гомозиготы по аллелю *Sdr-A1a*, bb – гомозиготы по аллелю *Sdr-A1b*, ab – гетерозиготы.

Средние рассчитаны по объединенным данным двух гибридных комбинаций № 15 × № 5 и № 15 × № 8. Показаны 95 % доверительные интервалы.

а

```
                                  -454
T. carth_5     CTGCCAACGTGACCCCACCCCACGCCCCATGAATGGAGCGGCAGCCAGCCAGTGACCGCG    240
T. aeth_16     CTGCCAACGTCACCCCACCCCACGCCCCATGAATGGAGCGGCAGCCAGCCAGTGACCGCG    240
T. aest_CS     CTGCCAACGTGACCCCACCCCACGCCCCATGAATGGAGCGGCAGCCAGCCAGTGACCGCG    240

T. carth_5     GCGGGACGCGCCACGGGGGACGAGGCGACACGAAGCGAGCAGGCGCGGTGCGGCGAACGC    300
T. aeth_16     GCGGGACGCGCCACGGGGGACGAGGCGACACGAAGCGAGCAGGCGCGGTGCGGCGAACGC    300
T. aest_CS     GCGGGACGCGCCACGGGGGACGAGGCGACACGAAGCGAGCAGGCGCGGTGCGGCGAACGC    300
                                                  -317
T. carth_5     GATCGCGCCGCGCATGCAACGGGCCTCTCGCGCCTCCCCGTGCCCGCCCGCCCGTCCGGT    360
T. aeth_16     GATCGCGCCGCGCATGCAACGGGCCTC--GCGCCTCCCCGTGCCCGCCCGCCCGTCCGGT    358
T. aest_CS     GATCGCGCCGCGCATGCAACGGGCCTC--GCGCCTCCCCGTGCCCGCCCGCCCGTCCGGT    358

T. carth_5     GCCTCCCCCCACTCCGTCACACGCCTGACGCCCCTCTCACTGGACGCCACTGGAATCCAC    420
T. aeth_16     GCCTCCCCCCACTCCGTCACACGCCTGACGCCCCTCTCACTGGACGCCACTGGAATCCAC    418
T. aest_CS     GCCTCCCCCCACTCCGTCACACGCCTGACGCCCCTCTCACTGGACGCCACTGGAATCCAC    418
           -226
T. carth_5     AGTCCTCTCCCTCCAAAGCAGCGCGCCCCGCGACTCGCCTCCGCCTACGTGTCGGCCCCG    480
T. aeth_16     GGTCCTCTCCCTCCAAAGCAGCGCGCCCCGCGACTCGCCTCCGCCTACGTGTCGGCCCCG    480
T. aest_CS     AGTCCTCTCCCTCCAAAGCAGCGCGCCCCGCGACTCGCCTCCGCCTACGTGTCGGCCCCG    478

T. carth_5     TCCCCGCCCGCTCGCCCACGTACCCCGCGCCTCGTTCCCACGTGCCCCTCCCTCTGCGCG    540
T. aeth_16     TCCCCGCCCGCTCGCCCACGTACCCCGCGCCTCGTTCCCACGTGCCCCTCCCTCTGCGCG    538
T. aest_CS     TCCCCGCCCGCTCGCCCACGTACCCCGCGCCTCGTTCCCACGTGCCCCTCCCTCTGCGCG    540
                                                                  5'UTR
T. carth_5     CATCCGATTGGCCGCCCACGCCTTCTTAAGCCGGCACGGCACCGGGACCCAACGCCGTGC    600
T. aeth_16     CATCCGATTGGCCGCCCACGCCTTCTTAAGCCGGCACGGCACCGGGACCCAACGCCGTGC    598
T. aest_CS     CATCCGATTGGCCGCCCACGCCTTCTTAAGCCGGCACGGCACCGGGACCCAACGCCGTGC    598
                                                -11          CSD (Sdr-B1)
T. carth_5     ACTCCGTCCACCCCCGTCAGCAGACTTCGACTCGCGCGTGCACGCAATGGCCATGGTGCA    660
T. aeth_16     ACTCCGTCCACCCCCGTCAGCAGACTTCGACTCGCGCGTGCACGCAATGGCCATGGTGCA    658
T. aest_CS     ACTCCGTCCACCCCCGTCAGCAGACTTCGACTCGCACGTGCACGCAATGGCCATGGTGCA    658
```

б

```
                                 3'UTR (TraesCS2B02G215200)
T. carth_5     TAGCCAAGAAATTCTGTGTGAAGGGGAATTTTACA----------------AGCATTTTG    2084
T. aeth_16     TAGCCAAGAAATTCTGTGTGAAGGGGAATTTTACAGGCAAGGCTAATAACAAGCATTTTG    2098
T. aest_CS     TAGCCAAGAAATTCTGTGTGAAGGGGAATTTTACAGGCAAGGCTAATAACAAGCATTTTG    2098
```

Рис. 3. Выравнивание нуклеотидных последовательностей, содержащих полиморфизмы, выявленные в окружении гена *Sdr-B1*.

а – промотор и 5′-конец гена *Sdr-B1* образцов пшеницы эфиопской № 16 и пшеницы карталинской № 5, выровненные против референсной последовательности хромосомы 2B пшеницы мягкой сорта Chinese Spring (CS). Числа над строками – расстояние от старт-кодона. UTR – нетранслируемая область; CSD – белок-кодирующая последовательность гена; *б* – выравнивание в области 16-нуклеотидной делеции в 3′-нетранслируемой области гена *TraesCS2B02G215200*, расположенного на комплементарной цепи ДНК рядом с геном *Sdr-B1* (*TraesCS2B02G215300*).

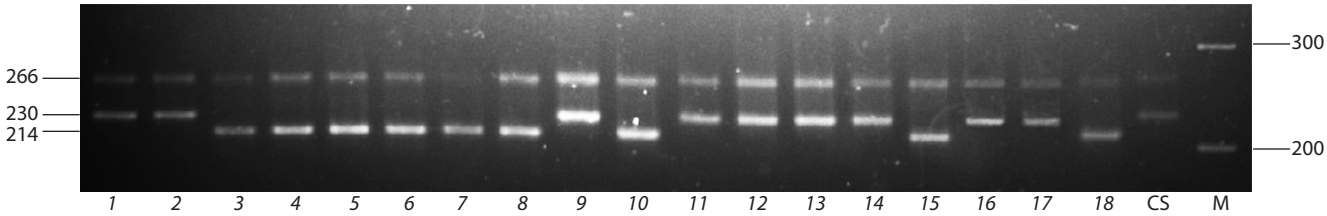


Рис. 4. Маркер делеции в некодирующей области рядом с геном *SDR-B1* у образцов *Triticum persicum* (*1–10*), *T. aethiopicum* (*11–18*) и *T. aestivum* сорта Chinese Spring (CS).

M – маркер длин ДНК M-100 («Синтол»). Размеры маркерных фрагментов – 230 и 214 п. о. Фрагмент 266 п. о. – побочный продукт с хромосомы 2A.

цов пшеницы карталинской нами был обнаружен аллель *Sdr-A1a*, а у всех образцов менее устойчивой пшеницы эфиопской – аллель *Sdr-A1b*. Исходя из этого, можно было бы предположить, что аллель *Sdr-A1a* вносит некий вклад в устойчивость к предуборочному прорастанию пшеницы карталинской.

Однако при анализе связи индекса прорастания с генотипами растений в F_2 получились результаты, противоречащие выдвинутому выше предположению, – растения с аллелем *Sdr-A1b* дали зерно с меньшим индексом прорастания, чем растения с аллелем *Sdr-A1a*. Гетерозиготные растения показали промежуточное значение индекса прорастания между гомозиготами. Иное направление связи генотипов растений с индексом прорастания, наблюдаемое в этом случае, может быть вызвано несколькими причинами. Во-первых, возможно, что влияние аллелей гена *Sdr-A1* на покой семян значительно зависит от условий выращивания растений. Растения F_2 росли в теплице, в то время как исходные образцы пшеницы давали урожай в поле. Во-вторых, взаимодействие гена *Sdr-A1* с другими генами могло поменять эффекты аллелей на противоположные в данных гибридных комбинациях. В-третьих, ве-

роятно, что такая перемена связана с мутацией в какой-то тесно сцепленной регуляторной последовательности ДНК (энхансере или сайленсере), которая не была захвачена при секвенировании гена *Sdr-A1*.

Анализ промотора гена *Sdr-B1* с помощью онлайн-сервиса PlantCARE выявил, что в нем преобладают цис-регуляторные элементы, связанные с реакцией на свет, абсцизовую кислоту, метилжасмонат и аноксию, а также ответственные за экспрессию в меристеме. Наличие этих регуляторных элементов в промоторе гена *Sdr-B1* в некоторой степени объясняет известное усиление покоя семян пшеницы и родственных злаков в связи с действием света, абсцизовой кислоты и недостатка кислорода и ослабление покоя семян под действием метилжасмоната (Walker-Simmons, 1988; Hoang et al., 2014; Xu et al., 2016). Динуклеотидная инсерция CT у образцов № 5, 8 и 15, вероятно, не влияет на структуру цис-действующих элементов. Вместе с тем однонуклеотидные замены у образца № 16 вызывают изменения двух регуляторных последовательностей. Замена G(–454)C обусловливает появление дополнительного элемента, ответственного за реакцию на метилжасмонат, и исчезновение G-box светочувствительного элемента, а замена A(–226)G – исчезновение GT-1 связывающего светочувствительного элемента. Все эти изменения в промоторе могут ослаблять, хотя незначительно, экспрессию гена *Sdr-B1* и, соответственно, быть одной из причин слабого покоя семян у образца № 16.

Заключение

Идентифицируемый с помощью CAPS-маркера однонуклеотидный полиморфизм A(-11)G в 5′-некодирующей области гена *Sdr-B1* не показал статистически значимой связи с индексом прорастания у изучаемых образцов пшеницы. Однако на том же материале нами была обнаружена слабая связь ($p = 0.1$) индекса прорастания с делецией в 3′-нетранслируемой области другого гена, лежащего близко на комплементарной цепи ДНК. Предположительно, эта делеция может влиять на экспрессию гена *Sdr-B1* через механизм РНК-интерференции, так как гены *TraesCS2B02G215200* и *Sdr-B1* (*TraesCS2B02G215300*) находятся на комплементарных друг другу цепях ДНК и разделены очень небольшим промежутком, примерно 200 п. о. В случае, если делеция 16 нуклеотидов рядом с геном *SDR-B1* фактически не оказывает влияния на его экспрессию, ее можно использовать как удобный вспомогательный маркер для определенных аллелей этого гена в некоторых гибридных комбинациях без проведения рестрикции.

Список литературы / References

Бадаева Е.Д., Шишкина А.А., Гончаров Н.П., Зуев Е.В., Лысенко Н.С., Митрофанова О.П., Драгович А.Ю., Кудрявцев А.М. Эволюция *Triticum aethiopicum* Jakubz. с позиции хромосомного анализа. Генетика. 2018;54(6):613-628. DOI 10.7868/S0016675818060048.
[Badaeva E.D., Shishkina A.A., Dragovich A.Y., Kudriavtsev A.M., Goncharov N.P., Zuev E.V., Lysenko N.S., Mitrofanova O.P. Evolution of *Triticum aethiopicum* Jakubz. from the position of chromosome analysis. Russ. J. Genet. 2018;54(6):629-642. DOI 10.1134/S1022795418060029.]

Дорофеев В.Ф., Филатенко А.А., Мигушова Э.Ф., Удачин Р.А., Якубцинер М.М. Культурная флора СССР. Т. 1. Пшеница. Л.: Колос, 1979.
[Dorofeev V.F., Filatenko A.A., Migushova E.F., Udachin R.A., Jakubtsiner M.M. Cultural Flora of the USSR. Vol. 1. Wheat. Leningrad: Kolos, 1979. (in Russian)]

Alaux M., Rogers J., Letellier T., Flores R., Alfama F., Pommier C., Mohellibi N., Durand S., Kimmel E., Michotey C., Guerche C., Loaec M., Laine M., Steinbach D., Choulet F., Rimbert H., Leroy P., Guilhot N., Salse J., Feuillet C. Linking the International Wheat Genome Sequencing Consortium bread wheat reference genome sequence to wheat genetic and phenomic data. Genome Biol. 2018; 19(1):111. DOI 10.1186/s13059-018-1491-4.

Doyle P.J. DNA Protocols for Plants. Molecular Techniques in Taxonomy. 1991;57(1):283-293. DOI 10.1007/978-3-642-83962-7.

Hoang H.H., Sechet J., Bailly C., Leymarie J., Corbineau F. Inhibition of germination of dormant barley (*Hordeum vulgare* L.) grains by blue light as related to oxygen and hormonal regulation. Plant Cell Environ. 2014;37(6):1393-1403. DOI 10.1111/pce.12239.

Lescot M., Dehais P., Thijs G., Marchal K., Moreau Y., Van de Peer Y., Rouze P., Rombauts S. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325-327.

Nakamura S. Grain dormancy genes responsible for preventing pre-harvest sprouting in barley and wheat. Breed. Sci. 2018;68(3):295-304. DOI 10.1270/jsbbs.17138.

Olaerts H., Courtin C.M. Impact of preharvest sprouting on endogenous hydrolases and technological quality of wheat and bread: a review. Compr. Rev. Food Sci. Food Safety. 2018;17(3):698-713. DOI 10.1111/1541-4337.12347.

Sugimoto K., Takeuchi Y., Ebana K., Miyao A., Hirochika H., Hara N., Ishiyama K., Kobayashi M., Ban Y., Hattori T., Yano M. Molecular cloning of *Sdr4*, a regulator involved in seed dormancy and domestication of rice. Proc. Natl. Acad. Sci. USA. 2010;107(13):5792-5797. DOI 10.1073/pnas.0911965107.

Vetch J.M., Stougaard R.N., Martin J.M., Giroux M. Allelic impacts of *TaPHS1*, *TaMKK3*, and *Vp1B3* on preharvest sprouting of northern great plains winter wheats. Crop Sci. 2018;59(1):140-150. DOI 10.2135/cropsci2018.05.0341.

Walker-Simmons M. Enhancement of ABA responsiveness in wheat embryos by high temperature. Plant Cell Environ. 1988;11(8):769-775. DOI 10.1111/j.1365-3040.1988.tb01161.x.

Xu Q., Truong T.T., Barrero J.M., Jacobsen J.V., Hocart C.H., Gubler F. A role for jasmonates in the release of dormancy by cold stratification in wheat. J. Exp. Bot. 2016;67(11):3497-3508. DOI 10.1093/jxb/erw172.

Zhang Y., Miao X., Xia X., He Z. Cloning of seed dormancy genes (*TaSdr*) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. Theor. Appl. Genet. 2014;127(4):855-866. DOI 10.1007/s00122-014-2262-6.

Zhang Y., Xia X., He Z. The seed dormancy allele *TaSdr-A1a* associated with pre-harvest sprouting tolerance is mainly present in Chinese wheat landraces. Theor. Appl. Genet. 2017;130(1):81-89. DOI 10.1007/s00122-016-2793-0.

ORCID ID
M. Bazhenov orcid.org/0000-0002-7301-1363
E. Guseva orcid.org/0000-0001-5157-0386
V. Rubets orcid.org/0000-0003-1233-8837

Благодарности. Исследование выполнено за счет гранта Российского научного фонда (проект № 17-76-10054).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 18.01.2019. После доработки 19.06.2019. Принята к публикации 19.06.2019.

ГЕНЕТИКА И СЕЛЕКЦИЯ РАСТЕНИЙ
Оригинальное исследование / Original article
УДК 631.523.13:582.736

Вавиловский журнал генетики и селекции. 2019;23(8):972-980
DOI 10.18699/VJ19.574

The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera

N.V. Shatskaya[1#], V.S. Bogdanova[1#], O.E. Kosterin[1, 2] ✉, G.V. Vasiliev[1], A.K. Kimeklis[3], E.E. Andronov[3], N.A. Provorov[3]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] All-Russia Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia
✉ e-mail: kosterin@bionet.nsc.ru

The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. were assembled on the base of the data of high-throughput sequencing of DNA isolated from a sample from North Osetia, Russia, using Illumina and PacBio platforms. The long PacBio reads were sufficient for reliable assembling organellar genomes while the short Illumina reads obtained from total DNA were unacceptable for this purpose because of substantial contamination by nuclear sequences. The organellar genomes were circular DNA molecules; the genome of mitochondria was represented by two circular chromosomes. A phylogenetic analysis on the basis of plastid genomes available in public databases was performed for some representatives of the tribes Fabeae, Trifolieae and Cicereae. As was expected, the *V. formosa* branch proved to be sister to the *Pisum* branch, and the tribe Fabeae was monophyletic. The position of *Trifolium* L. appeared sensitive to the phylogeny reconstruction method, either clustering with Fabeae or with the genera *Medicago* L., *Trigonella* L. and *Melilotus* Mill., but the internodes between successive divergences were short in all cases, suggesting that the radiation of *Trifolium*, other Trifolieae and Fabeae was fast, occurring within a small time interval as compared to further evolution of these lineages. The data on the relatedness of the plastid genomes of *Trifolium* and Fabeae correlate with the similarity of N_2-fixing symbionts in these legumes represented by *Rhizobium leguminosarum* biovars *trifolii* and *viciae*, while the symbionts of *Medicago*, *Melilotus* and *Trigonella* belong to the *Sinorhizobium meliloti* and *S. medicae* species, which are distant from *Rhizobium*.
Key words: *Vavilovia formosa* (Stev.) Fed.; *Vavilovia* A. Fedorov; *Lathyrus* L.; *Vicia* L.; *Pisum* L.; *Lens* L.; *Trifolium* L.; *Medicago* L.; *Trigonella* L.; *Melilotus* Mill.; *Cicer* L.; Fabeae; Trifolieae; Cicereae; crop wild relatives; pea; plastid genome; phylogeny; paraphyly; monophyly.

For citation: Shatskaya N.V., Bogdanova V.S., Kosterin O.E., Vasiliev G.V., Kimeklis A.K., Andronov E.E., Provorov N.A. The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):972-980. DOI 10.18699/VJ19.574

Пластидный и митохондриальный геномы *Vavilovia formosa* (Stev.) Fed. и филогения родственных родов бобовых

Н.В. Шацкая[1#], В.С. Богданова[1#], О.Э. Костерин[1, 2] ✉, Г.В. Васильев[1], А.К. Кимеклис[3], Е.Е. Андронов[3], Н.А. Проворов[3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии, Пушкин, Санкт-Петербург, Россия
✉ e-mail: kosterin@bionet.nsc.ru

На основании данных высокопроизводительного секвенирования на платформах Illumina и PacBio тотальной ДНК, выделенной из образца *Vavilovia formosa* (Stev.) Fed. из Северной Осетии (Россия), собраны пластидный и митохондриальный геномы этого вида. Длинные риды, получаемые на платформе PacBio, оказались достаточными для надежной сборки геномов органелл, тогда как короткие риды, получаемые на платформе Illumina, – непригодными для этой цели ввиду значительной контаминации последовательностями ядерного происхождения. Геномы органелл представляют собой кольцевые молекулы ДНК, причем митохондриальный геном состоит из двух кольцевых хромосом. На основе имеющихся в публичных базах данных последовательностей пластидных геномов и пластидного генома *Vavilovia* предпринят филогенетический анализ, вовлекший некоторых представителей триб Fabeae, Trifolieae и Cicereae. Как и ожидалось, ветвь *V. formosa* оказалась сестринской к ветви *Pisum* L. (горох), а триба Fabeae – монофилетична. Позиция рода *Trifolium* L. (клевер) зависела от метода реконструкции филогении – он кластеризовался либо с Fabeae, либо с родами *Medicago* L. (люцерна), *Trigonella* L. (пажитник) и *Melilotus* Mill. (донник). Вне зависимости от метода реконструкции длина ветвей между последовательными дивергенциями была незначительной, что свидетельствует о быстрой радиации *Trifolium*, других представителей триб Trifolieae и Fabeae в течение короткого времени по сравнению с дальнейшей эволюцией

© Shatskaya N.V., Bogdanova V.S., Kosterin O.E., Vasiliev G.V., Kimeklis A.K., Andronov E.E., Provorov N.A., 2019
[#] N.V. Shatskaya, V.S. Bogdanova contributed equally to this work.

соответствующих линий. Данные о родстве пластидных геномов рода *Trifolium* и трибы Fabeae коррелируют со сходством N_2-фиксирующих симбионтов этих бобовых, представленных *Rhizobium leguminosarum* штаммов *trifolii* и *viciae*, тогда как симбионты *Medicago*, *Melilotus* и *Trigonella* принадлежат к видам *Sinorhizobium meliloti* и *S. medicae*, эволюционно отдаленным от *Rhizobium*.
Ключевые слова: *Vavilovia formosa* (Stev.) Fed.; *Vavilovia* A. Fedorov; *Lathyrus* L.; *Vicia* L.; *Pisum* L.; *Lens* L.; *Trifolium* L.; *Medicago* L.; *Trigonella* L.; *Melilotus* Mill.; *Cicer* L.; Fabeae; Trifolieae; Cicereae; дикие сородичи культурных растений; горох; пластидный геном; филогения; парафилия; монофилия.

Introduction

Vavilovia formosa (Stev.) Fed. is a small perennial herbaceous legume confined to highlands of the Caucasus and Anterior Asia (Davis, 1970; Vishnyakova et al., 2016). Although morphologically variable, it is traditionally considered the only member of the monotypic genus *Vavilovia* A. Fedorov. Morphological and molecular data suggest it to be the closest genus to *Pisum* L. (peas, annual plants), to which the important crop *Pisum sativum* L. belongs. Both genera belong to the tribe Fabeae Rchb. Recently, H. Schaefer et al. (2012) reconstructed a phylogeny of this tribe and showed that *Pisum* and *Vavilovia* form a clade inside the speciose genus *Lathyrus* L., making it paraphyletic. They propose to subsume *Pisum* and *Vavilovia* to *Lathyrus* but the corresponding nomenclatorial change for *Pisum* is not yet broadly accepted and that for *Vavilovia* has not been formally made.

At present, genomic research is extensively conducted in both fundamental (e. g., phylogenetics, phylogeography, evolutionary theory and taxonomy, comparative and functional genomics) and applied (e. g., QTL analysis, association mapping, and marker-assisted selection) aspects. Organellar genomes are among the most popular research objects, since their relatively small size allows their sequencing through the high-throughput approach rather easily. Since mitochondria and plastids have the apparently symbiotic origin (from α-proteobacteria and cyanobacteria, respectively), their research may shed light on the coevolution of plants with microorganisms that genetically and functionally interact with these organelles. Specifically, the N_2-fixing symbionts of legumes, rhizobia, form organelle-like compartments, symbiosomes, inside the plant cells. Being metabolically integrated, they apparently coevolve with plastids and mitochondria (de la Peña et al., 2018).

Thus far, plastid genomes have been sequenced in many legume genera, including *Vicia*, *Lathyrus*, *Pisum*, *Lens*, *Cicer*, *Trifolium*, *Medicago* etc., whereas complete mitochondrial genomes are available only for 14 species of Fabales (ncbi.nlm.nih.gov, accessed on March 5, 2018). This is due to the nature of plant mitochondrial DNA, which generally occurs as a set of interconverting subgenomic molecules as a result of homologous recombination between repeted regions (reviewed in Gualberto et al., 2014). For this reason, plant mitochondrial genomes are more difficult to assemble than those of plastids (Smith, Keeling, 2015).

The plastid and mitochondrial genomes of *V. formosa* are not available yet, in spite of explosive interest to this species in the recent decade (Akopian, Gabrielyan, 2008; Mikić et al., 2009, 2013, 2014; Sinjushin et al., 2009; Akopian et al., 2010, 2014; Atlagić et al., 2010; Oskoueiyan et al., 2010; Sinjushin, Belyakova, 2010; Zemerski-Škorić et al., 2010; Zorić et al., 2010; Vishnyakova et al., 2013, 2016; Safronova et al., 2014, 2015). This interest was motivated by *V. formosa* being although the most distant but still a pea crop wild relative, which may harbor some genes useful for pea pre-breeding and somehow transferrable to pea.

The phylogenetic tree of most species of the tribe Fabeae has been extensively and reliably reconstructed by H. Schaefer et al. (2012), but the positions of the genera evolutionary closest to this tribe are problematic. According to the traditional taxonomy, the tribes most related to Fabeae are Cicereae Alefeld, with the only genus *Cicer* L., and then Trifolieae (Bronn) Benth., with the genera *Trifolium* s.l., *Medicago* L. s.l., *Trigonella* L., *Melilotus* Mill., *Parochetus* Buch.-Ham. ex D. Don. and *Ononis* L. (Yakovlev, 1991). These genera comprise the so-called 'vicioid clade' in the sense by M.F. Wojciechowski et al. (2004). However, in the phylogenetic tree reconstructed by M.F. Wojciechowski et al. (2004) based on the plastid gene *matK* there was a highly supported divergence between the branches including (i) the genera *Ononis*, *Medicago*, *Trigonella*, and *Melilotus* and (ii) the genera *Trifolium*, *Vicia*, *Pisum*, *Lathyrus*, and *Lens*, the last four genera belonging to Fabeae. Interestingly, this classification is correlated to the plant symbiotic affinities to N_2-fixing nodule bacteria, since the plants from the second branch are inoculated by closely related symbionts: *Trifolium* by *Rhizobium leguminosarum* bv. *trifolii*, *Vicia*, *Pisum*, *Lathyrus*, *Lens* by *R. leguminosarum* bv. *vicieae*), while *Medicago*, *Trigonella*, and *Melilotus* have distant symbionts from the *Sinorhizobium* genus (Biondi et al., 2003; Dudeja, Nidhi, 2013). The genus *Cicer* (traditionally attributed to the monophyletic tribe Cicereae) appeared sister to branches (i) and (ii) taken together. The more proximal branch contained the genus *Galega* L. (traditionally attributed to the large tribe Galegae (Bronn) Torr. ex Gray), and the first divergence again was formed by a Trifolieae representative *Parochetus*.

The phylogenetic analysis of Fabeae by H. Schaefer et al. (2012), based on five plastid and one nuclear (ribosomal ITS) sequence, included an outgroup consisting of three *Trifolium* species, three *Medicago* species, one species *Melilotus*, and one *Ononis* species. Although the entire tree including the outgroup is inevitably unrooted, there was a node with bootstrap support 100 uniting the genus *Trifolium* with the tribe Fabeae, which agrees with the result by M.F. Wojciechowski et al. (2004).

Their coherent results meant that the evolutionary lineage including Fabeae and *Triolium* is sister to the lineage including most of the rest of the tribe Trifolieae in the traditional sense, thus making Trifolieae paraphyletic. It is noteworthy that the phylogenetic marker used by M.F. Wojciechowski et al. (2004) and five of six markers used by H. Schaefer et al. (2012) were plastid sequences. However, a Maximum Likelihood phylogenetic analysis based on 28 nuclear sequences showed *Trifolium* to cluster with *Medicago* and form a branch opposed to *Pisum* (Kreplak et al., 2019, Fig. 2, *b*), which is a pattern corresponding to the traditional systematics.

In view of this controversy of the phylogenetic position of *Trifolium*, it was interesting to consider a phylogenetic tree reconstructed from complete or nearly complete plastid genomes.

In this work we (i) for the first time report the complete DNA sequence of both plastid and mitochondrial genomes of *Vavilovia formosa* and (ii) use the plastid genome to reconstruct the phylogeny of several legume genera. The obtained data allow us to address the correlation between the plastid-based phylogeny of legumes and their symbiotic affinities presumably reflecting the tight functional and coevolutionary interactions of plastids with temporal N_2-fixing organelles, symbiosomes (de la Peña et al., 2018).

Materials and methods

Plant material. *Vavilovia* seeds were provided by the Gorsky State Agrarian University in Vladikavkaz. They represent a *V. formosa* population in the North Ossetian State Natural Reserve, North Ossetia, the Caucasus, Russia.

DNA isolation and high throughput sequencing. DNA from *Vavilovia* plant tissues was isolated with AxyPrep™ Multisource Genomic DNA Miniprep kit according to manufacturer's recommendations. Whole genome sequencing was done on the Illumina and PacBio platforms in the Macrogen genome sequencing company (Korea).

Organellar genome assembly. To assemble the plastid genome of *Vavilovia*, the reads were filtered with the Mirabait utility of the MIRA4.0 package (Chevreux et al., 1999) using the sequence of the *Pisum sativum* chloroplast genome (NC_014057) as a probe. A subset of sequences longer than 10 kb was searched to find a read containing the starting point of the assembly, the *trnH* gene. Then a read overlapping the initial read was selected, the reads were merged, and the dataset was searched for a next read to elongate the assembly. The assembly was elongated in such a manner, until it closed into a circle. It was used as a reference sequence for mapping the *Vavilovia* plastid genome with MIRA4.0 (Chevreux et al., 1999).

Two assemblies were made, one starting from long PacBio reads and the other from short Illumina reads. It is commonly accepted that the best results are gained by combination of these two types of reads (see e. g. Gnerre et al., 2011). The comparison of the two assemblies of *Vavilovia* plastid genome revealed that there appeared various regions with a lot of discrepancies. While the assembly of long reads corresponded well to the reference sequence, the assembly of short reads had a number of mismatches, such as nucleotide substitutions and short indels.

To understand the origin of the discrepancy, some of such regions were checked more carefully. For example, the region corresponding to nucleotide positions 39,400–40,000 of the reference sequence contained 17 mismatches within about 120 bp of alignment, possibly due to the nuclear origin of the reads involved into the assembly. Since short Illumina reads (up to 150 bp) do not permit to investigate their genomic environment, all long reads of our dataset that shared homology with that region were checked whether they belonged to the chloroplast genome indeed. A sample of 939 PacBio reads longer than 10 kb was filtered with the Mirabait utility (Chevreux et al., 1999) using the above-mentioned stretch of 600 bp of the reference sequence. In total, 89 reads were filtered out. Of them, 80 lay entirely in the plastid genome, while the other 9 matched the assembly partially, sharing with it DNA stretches of variable lengths, 300 to 16,000 bp. These 9 reads were used as a query for a BLAST search of the nonredundant database at ncbi.nlm.nih.gov (Altschul et al., 1990). Three of them appeared to correspond entirely to the mitochondrial genome. This is quite natural, since the mitochondrial genome shares about 2.5 kb of homologous DNA stretches with the chloroplast genome, as evidenced from the comparison of the *Vicia faba* L. mitochondrial genome (KC189947) with the *Pisum sativum* plastid genome (NC_014057).

The remaining six reads contained a 300–500 bp region of homology with plastid/mitochondrial DNA, but the rest part had either no homology in the nonredundant database or 1000–1500 bp stretches of homology with genomic clones of some leguminous plants. Most probably, these reads represented nuclear copies of plastid genes. The DNA stretch corresponding to the region 54,400–55,200 of the reference sequence had 15 mismatches per 600 bp of the assembly. A total of 88 reads (longer than 10 kb) that had homology to this region were filtered out. Four of them matched the plastid genome partially, with 8–16 kb corresponding to the plastid genome and 600–2700 bp with no significant similarity in the nonredundant database. Other two randomly taken regions had no obvious discrepancies in the assembly made of short reads with the reference sequence. Seventy-eight reads (longer than 10 kb) were filtered out that passed across the region 60,000–60,500. One of them contained a stretch of 2700 bp that did not match the plastid DNA. All of the 119 long reads passing across the region 80,000–80,500 entirely matched the plastid DNA.

Based on the above observations, a conclusion was inferred that discrepancies in the assemblies made from long vs. short reads arose due to the presence of nuclear copies of plastid DNA of various lengths, from about 300 to 16,000 bp, with the mean of about 7,000 bp. Therefore, the assembly of long PacBio reads was considered more appropriate for plastid genome reconstruction.

The resulting assembly was reasonably consistent. The total amount of mismatches was 0.25 %, and the average coverage depth was 78. These values suggested that the PacBio reads were sufficient for reliable assembling an organellar genome, while the short Illumina reads obtained from total DNA were unacceptable for this purpose because of substantial contamination by nuclear sequences.

The mitochondrial genome was assembled in a similar manner, with the original filtering of reads using the *V. faba* mitochondrial genome (KC189947). The assembly consisted of two ring chromosomes with average coverage depth 84 and 59, and the total number of mismatches was 0.37 %.

The plastid genome of *V. formosa* was assigned the accession number MK604478, and the two chromosomes of its mitochondrial genome got the accession numbers MK48602 and MK48603 in public databases.

Alignment of plastid genomes for phylogenetic analysis. We undertook phylogenetic analysis of the plastid genomes available in public databases of some representatives of the tribes Fabeae, Trifolieae and Cicereae. The plastid genome sequence of *Vavilovia formosa* in general was not collinear

Accession numbers, coverage information, and percentages of similarity to the *V. formosa* plastid genome in the plastid genome reconstructions studied

Accession number	Species	Tribe	Representation of the reconstruction in the original plastid genome, %	Coverage of the *V. formosa* plastid genome, %	Identity to the *V. formosa* plastid genome, %
HG966674	*Pisum sativum*, voucher WL1238	Fabeae	92.9	93	95.5
MG458702	*Pisum fulvum*, voucher WL2140	Fabeae	93.0	92	95.5
KJ850235	*Lathyrus clymenum*	Fabeae	90.0	89	95.7
KJ850239	*Lens culinaris*	Fabeae	84.2	85	92.7
KF042344	*Vicia faba*	Fabeae	86.8	87	95.4
KJ850242	*Vicia sativa*	Fabeae	84.7	84	93.8
KJ788292	*Trifolium strictum*	Trifolieae	81.0	83	93.9
EU849487	*Trifolium subterraneum*	Trifolieae	66.2	77	93.4
KC894706	*Trifolium repens*	Trifolieae	76.8	83	94.0
KU321683	*Medicago sativa*	Trifolieae	77.7	81	94.0
JX512024	*Medicago truncatula* cultivar Paraggio	Trifolieae	78.3	78	93.9
NC_041419	*Melilotus albus*	Trifolieae	79.2	82	94.1
MK460508	*Trigonella foenum-graecum* voucher I.S. Choi MD025	Trifolieae	81.9	84	94.0
EU835853	*Cicer arietinum* voucher ICCV 10	Cicereae	70.2	79	92.1
DQ317523	*Glycine max* cultivar PI 437654	Phaseoleae	57.2	70	87.2

to those of other Fabaceae, differing from them by a large number of structural rearrangements. To make alignment, homologous DNA stretches were found by Blastn software at ncbi.nlm.nih.gov and manually put in the order and orientation corresponding to the *Vavilovia* plastid genome. Each next stretch of homology was sought in the portion of the plastid genome of a species to be aligned that was not yet included in the reconstruction. Then the plastid genome of *V. formosa* and reconstructions of the plastid genomes of the other species were aligned with ClustalW (Larkin et al., 2007) incorporated into the MEGA6 package (Tamura et al., 2013).

The obtained plastid genome reconstructions of the genera in question covered 77 to 93 % of the *Vavilovia* plastid genome, whereas that of *Glycine max*, used as an outgroup, covered 70 % (Table).

Phylogenetic analysis. Bayesian MCMC analysis was performed with the use of BEAST 2.4.3 software (Drummond, Rambaut, 2007). The GTR+I+G model was chosen using jModelTest 2.1.10 (Guindon, Gascuel, 2003; Darriba et al., 2012). An uncorrelated lognormal relaxed clock model and a Yule process of speciation were applied. One MCMC analysis was run for 100 million generations. Trees were visualized using the program FigTree 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/) by A. Rambaut.

The Maximum Likelihood reconstruction of the phylogeny was made with the aid of the MEGA6 package (Tamura et al., 2013) using the Kimura 2-p parameter, GTR+I+G model of mutation rates, and bootstrap test with 100 replications.

Results and discussion

The structure of the *Vavilovia formosa* mitochondrial genome. The mitochondrial genome of *Vavilovia* was assembled into two non-overlapping circles, 264,766 bp and 88,581 bp, totaling 353,347 bp. This is close to the mitochondrial genome size of *Lotus japonicus* L. (JN872551), 380,861 bp; or *Arabidopsis thaliana* (L.) Heynh. (NC_037304), 367,808 bp. It is larger than in *Medicago truncatula* (KT971339), 271,618 bp, and smaller than in *Vicia faba* (KC189947), 588,000 bp.

Interestingly, it appeared impossible to construct a single master molecule of the *Vavilovia* mitochondrial genome. Instead, two 'chromosomes' were obtained (Fig. 1). However, this is quite consistent with the dynamic nature of plant mitochondrial genomes (Gualberto et al., 2014). Another curious fact concerns the *Nad5* gene, which appeared to belong to both 'chromosomes', since its exons 1–3 reside in the first, larger 'chromosome', whereas exons 4–5 are in the second 'chromosome', thus requiring trans-splicing to produce the entire coding sequence. A similar situation has been described in *Silene* L., where some species possess up to 128 mitochondrial 'chromosomes', with exons of many genes present in more than one 'chromosome' (Sloan et al., 2012).

Phylogenetic analysis of the mitochondrial genomes of *Vavilovia* and related genera is impossible yet, as of the studied genera (see Table) complete mitochondrial genomes are presently available only for *M. truncatula*, *V. faba* and *G. max* (ncbi.nlm.nih.gov accessed on 22 August 2019).

The structure of the plastid genome of *Vavilovia formosa.* The content of the plastid genome of *V. formosa* is shown schematically in Fig. 2.

The total length is 122,196 bp, which is similar to the plastid genome size of *Pisum*, 122,180 bp in *P. sativum* (HG966674) and 120,837 bp in *P. fulvum* (MG458702). Expectedly, the gene content appeared very similar to that of *Pisum.* A notable difference is that the *Vavilovia* plastid genome has a tandem triplication of the tRNA gene for methionine. The three copies differ by nucleotide substitutions and a 5 bp long insertion/

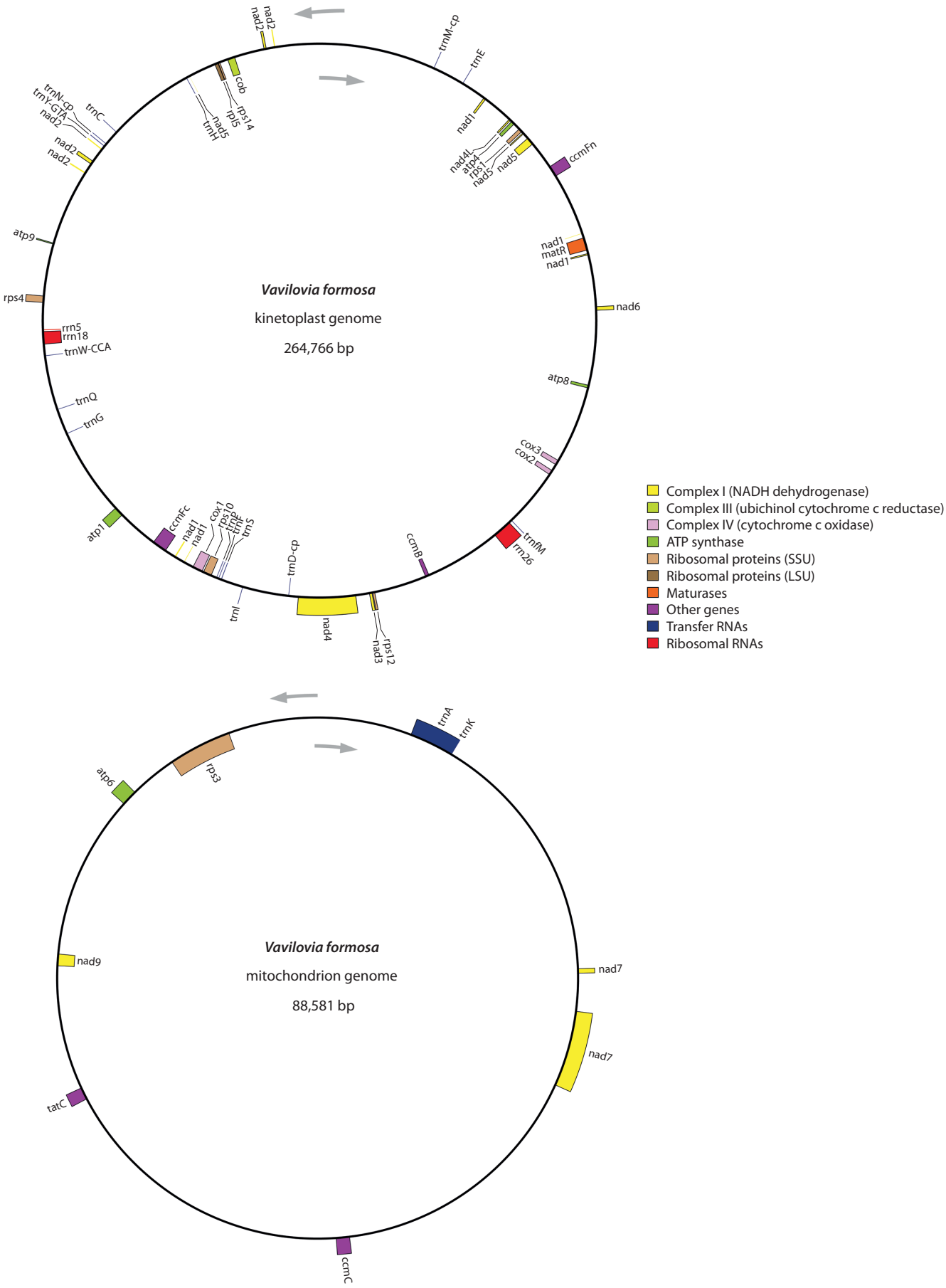


Fig. 1. Schematic presentation of *V. formosa* mitochondrial genome (assembled as two circles) drawn with OGDRAW (Lohse et al., 2013).

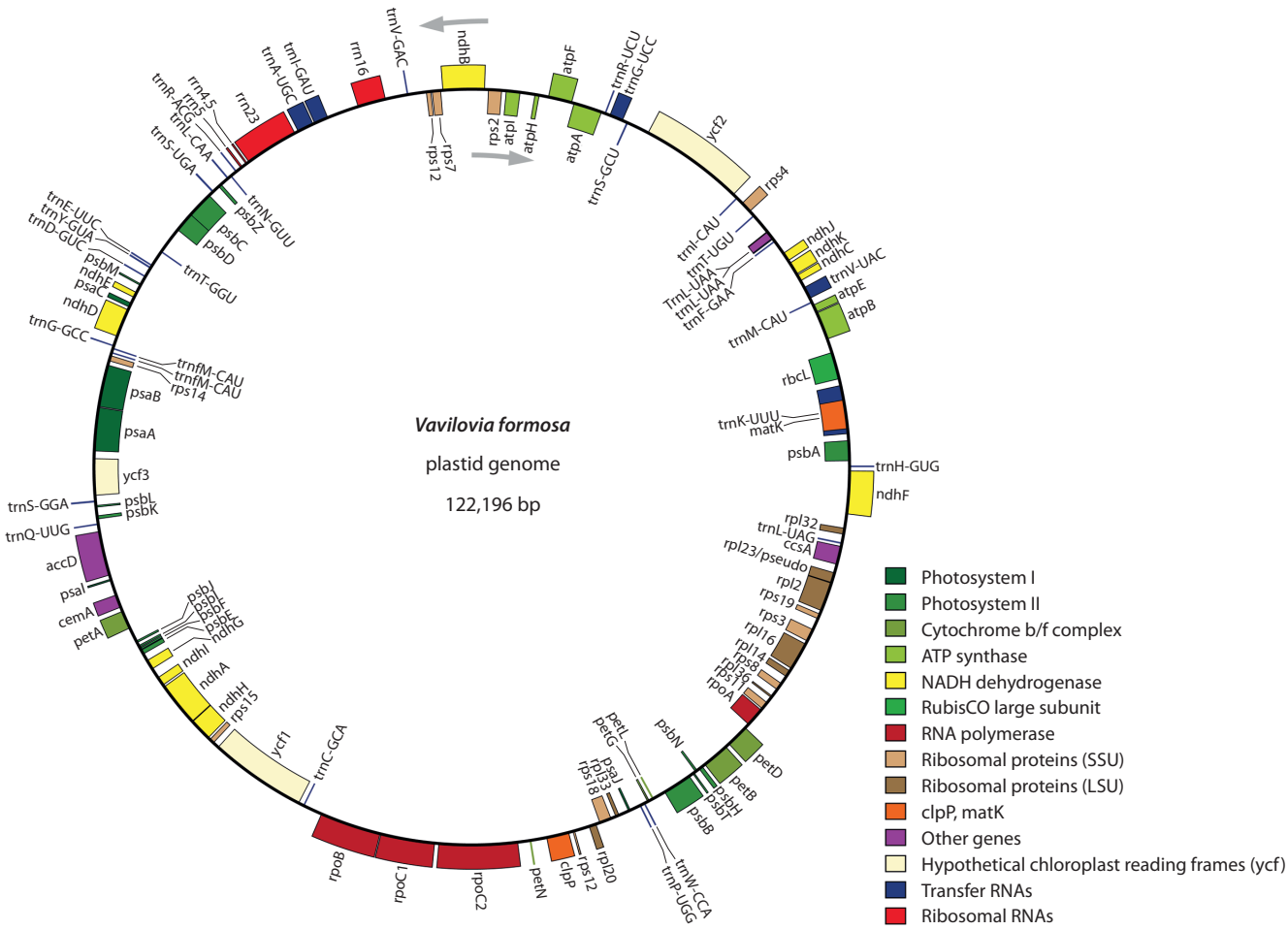


Fig. 2. Schematic presentation of *V. formosa* plastid genome drawn with OGDRAW (Lohse et al., 2013).

deletion. It is not known whether all the copies are functional. In addition, the gene order in *Vavilovia* differs from that of *Pisum* by 10 rearrangements.

Phylogenetic analysis involving plastid genomes of some related legume genera including *Vavilovia*. Figure 3 shows the obtained Bayesian phylogeny reconstruction for representatives of the tribes Fabeae, Trifolieae and Cicereae based on the plastid genome reconstructions and using the soybean plastid genome reconstruction as an outgroup. As expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the phylogenetic reconstructions based on shorter sequences: the *Vavilovia* branch is sister to the *Pisum* branch, as in (Sinjushin et al., 2009; Oskoueiyan et al., 2010; Schaefer et al., 2012), and their united branch is sister to *Lathyrus clymenum* as in (Schaefer et al., 2012); *Lens culinaris* L. is sister to the two involved *Vicia* spp.; and the tribe Fabeae is monophyletic, as in (Wojciechowski et al., 2004; Schaefer et al., 2012). Thus, the phylogenetic position of *Vavilovia* is unequivocal.

As already mentioned, the positions of *Medicago* and *Trifolium* in the phylogenetic reconstructions by M.F. Wojciechowski et al. (2004) (not involving *Vavilovia*) and H. Schaefer et al. (2012) contradicted the traditional taxonomy as showing *Medicago*, *Trigonella* and *Melilotus* to be a sister branch to that uniting *Trifolium* and Fabeae, thus making the traditional tribe Trifolieae paraphyletic. The phylogeny reconstructed here by the Bayesian analysis of the complete (or nearly complete) plastid genomes is expected to be more reliable, and it is consistent with the aforementioned results by M.F. Wojciechowski et al. (2004) and H. Schaefer et al. (2012). However, one can notice that although the node uniting *Trifolium* with Fabeae has a robust support of the posterior probability of 0.86 (see Fig. 3), the branch leading to it after the divergence from *Medicago* is very short. The same is seen in the trees by H. Schaefer et al. (2012).

At the same time *Trifolium* and other representatives of the traditional Trifolieae – *Medicago*, *Melilotus*, *Trigonella*, formed a united branch in the Maximum Likelihood tree with the highest possible bootstrap support (100), which is sister to Fabeae (Fig. 4). A similar pattern, with *Medicago* and *Trifolium* forming a branch sister to *Pisum*, was constructed by K. Kreplak et al. (2019), who made a phylogenetic reconstruction based on 28 nuclear sequences using the same Maximum Likelihood method. However, the branch leading to the traditional Trifolieae, including *Trifolium*, is again very short, both in our tree (see Fig. 4) and in the tree by (Kreplak et al., 2019, Fig. 2, *b*).

The fact that the positions of Fabeae, *Trifolium*, and other Trifolieae in the tree depend on the method of phylogeny re-

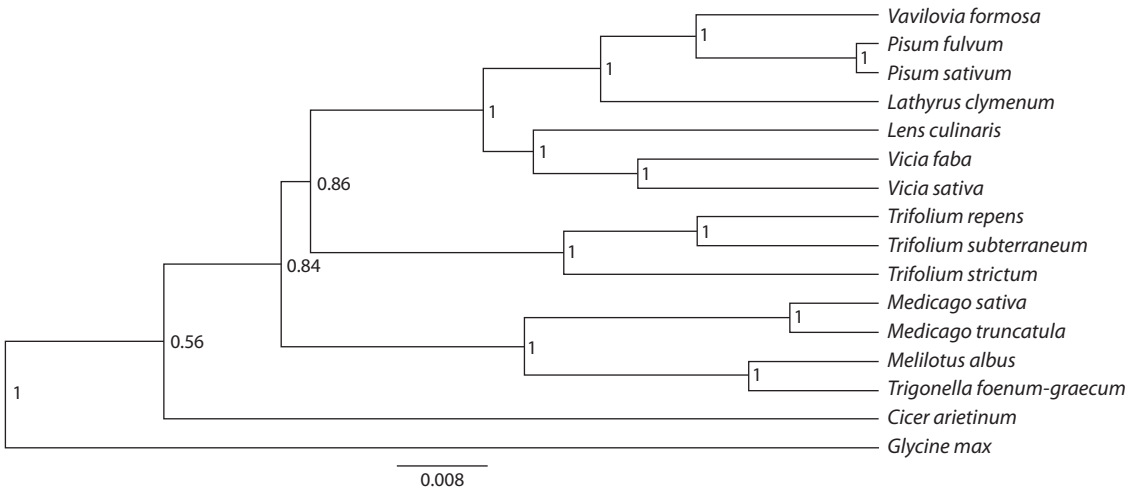


Fig. 3. The Bayesian phylogeny reconstruction (GTR+I+G model) with posterior probabilities for representatives of the tribes Fabeae, Trifolieae, and Cicereae based on the plastid genome reconstructions.

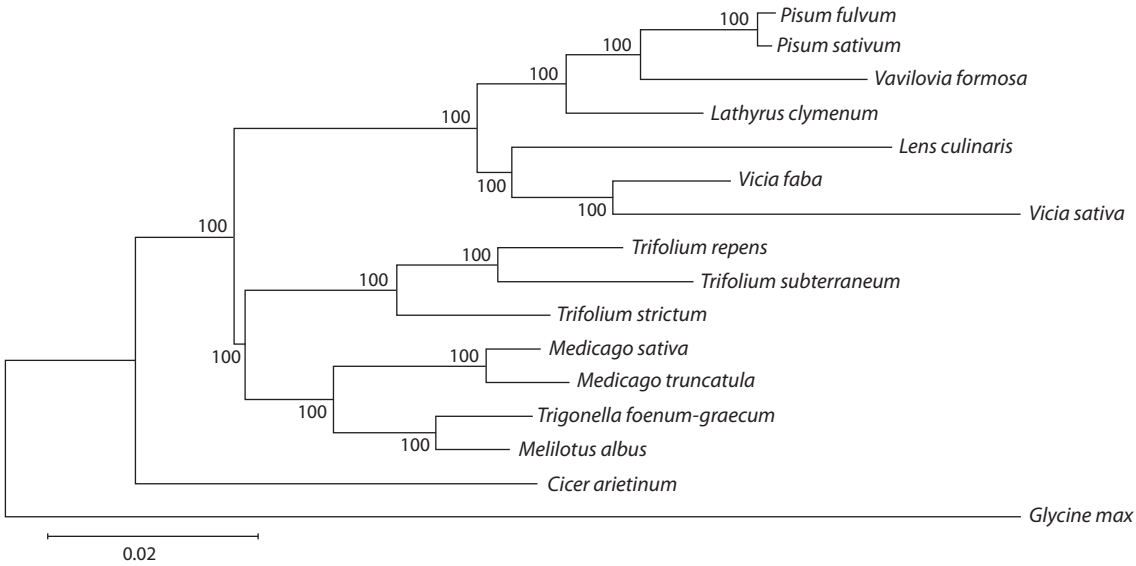


Fig. 4. Phylogenetic relationships of some Fabaceae obtained with the Maximum Likelihood algorithm based on the reconstructed plastid genome sequences.

Bootstrap values are given.

construction but each time reveal a short branch suggests that the radiation of *Medicago*, *Melilotus*, *Trigonella*, *Trifolium*, and Fabeae was fast. Its duration was short as compared to further evolution of these lineages; that is, the last common ancestors of Fabeae and *Trifolum*, of Fabeae and the traditional Trifolieae, and of *Trifolium* and the rest of traditional Trifolieae existed at close times. H. Schaefer et al. (2012) reconstructed the crown age of Fabeae as Middle Myocene (23–16 mya). The relative lengths of branches in the reconstructed phylogenetic trees suggest that the radiation to Fabeae and Trifolieae took place ca 1.5–1.8 myr earlier, that is in the Oligocene. For the time being there is no unequivocal reason to reconsider the traditional taxonomy uniting *Trifolium* and *Medicago* as opposed to Fabeae.

The data on relatedness of the plastid genomes of *Trifolium* and Fabeae correlate to the similarity of N_2-fixing symbionts in these legumes represented by *Rhizobium leguminosarum* biovars *trifolii* and *viciae* (Dudeja, Nidhi, 2013), while the symbionts of *Medicago*, *Melilotus* and *Trigonella* belong to the *Sinorhizobium meliloti* and *S. medicae* species, which are distant from *Rhizobium* (Biondi et al., 2003). This might arise from the functional relationship between rhizobial bacteroids and host plant plastids, where nitrogen fixed by the bacteroids is included into the amino acids and amides synthesized inside plastids of the infected cells of the nodules formed by the host plant (de la Peña et al., 2018). In view of this interaction, one can suggest that the related rhizobial symbionts were acquired by *Trifolium* and Fabeae plants due to compatibility with the similar plastid genomes.

Conclusion

Thus, phylogenetic analysis of a sample of the available plastid genomes of representatives of related legume genera, including *Vavilovia*, reported here, confirmed the expected

phylogenetic position of *Vavilovia* itself but challenged the presumed position of *Trifolium* and conjectured a certain coevolution between the plastids and bacterial symbionts of legumes, possibly because of their functional interaction.

References

Akopian J.A., Gabrielyan I.G. On high-mountain pea, *Vavilovia formosa* (Stev.) Fed. (Fabaceae) in Armenia. Crop Wild Relative. 2008;6: 26-27.

Akopian J., Sarukhanyan N., Gabrielyan I., Vanyan A., Mikić A., Smýkal P., Kenicer G., Vishnyakova M., Ambrose M. Reports on establishing an ex situ site for ''beautiful'' vavilovia (*Vavilovia formosa*) in Armenia. Genet. Resour. Crop Evol. 2010;57:1127-1134. DOI 10.1007/s10722-010-9606-0.

Akopian J.A., Sinjushin A.A., Gabrielyan I.G., Shaboyan G. On some biomorphological peculiarities of seedlings of *Vavilovia formosa* (Stev.) Fed. (Fabaceae). Legum Perspect. 2014;5:34-35.

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

Atlagić J., Mikić A., Sarukhanyan N., Vanyan A., Akopian J., Gabrielyan I., Smýkal P., Kenicer G., Vishnyakova M. Ambrose M. Contributions to the characterization of *Vavilovia formosa* (syn. *Pisum formosum*). II. Morphology of androecium and gynoecium and mitosis. Pisum Genet. 2010;41:21-24.

Biondi E.G., Giuntini P.E., Roumiantseva M.L., Andronov E.E., Onichtchouk O.P., Kurchak O.N., Simarov B.V., Dzyubenko N.I., Mengoni A., Bazzicalupo M. Genetic relationship of *Sinorhizobium meliloti* and *Sinorhizobium medicae* strains isolated from Caucasian region. FEMS Microbiol. Lett. 2003;220(2):207-213. DOI 10.1016/S0378-1097(03)00098-3.

Chevreux B., Wetter T., Suhai S. Genome sequence assembly using trace signals and additional sequence information. In: Computer Science and Biology: Proc. German Conference on Bioinformatics (GCB). 1999;99:45-56.

Darriba D., Taboada G.L., Doallo R., Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods. 2012; 9(8):772. DOI 10.1038/nmeth.2109.

Davis P.H. *Vavilovia* A. Fed. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Pueyo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahlawat S.K., Siwach P., Duhan J.S. (Eds.). Biotechnology: Prospects and Application. New Dehli: Springer, India, 2013;113-125. DOI 10.1007/978-81-322-1683-4_10.

Gnerre S., Maccallum I., Przybylski D., Ribeiro F.J., Burton J.N., Walker B.J., Sharpe T., Hall G., Shea T.P., Sykes S., Berlin A.M., Aird D., Costello M., Daza R., Williams L., Nicol R., Gnirke A., Nusbaum C., Lander E.S., Jaffe D.B. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc. Natl. Acad. Sci. USA. 2011;108:1513-1508. DOI 10.1073/pnas.1017351108.

Gualberto J.M., Mileshina D., Wallet C., Niazi A.K., Weber-Lotfi F., Dietrich A. The plant mitochondrial genome: dynamics and maintenance. Biochimie. 2014;100:107-120. DOI 10.1016/j.biochi.2013.09.016. PMID: 24075874.

Guindon S., Gascuel O. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst. Biol. 2003;52: 696-704. DOI 10.1080/10635150390235520.

Kreplak K., Madoui M.-A., Cápal P., Novák P., Labadie K., Aubert G., Bayer P.E., Gali K.K., Syme R.A., Main D., Klein A., Bérard A., Vrbová I., Fournier C., d´Agata L., Belser C., Berrabah W., Toegelová H., Milec Z., Vrána J., Lee H., Kougbeadjo A., Térézol M., Huneau C., Turo C.J., Mohellibi N., Neumann P., Falque M., Gallardo K., McGee R., Tar'an B., Bendahmane A., Aury J.M., Batley J., Le Paslier M.C., Ellis N., Warkentin T.D., Coyne C.J., Salse J., Edwards D., Lichtenzveig J., Macas J., Doležel J., Wincker P., Burstin J. A reference genome for pea provides insight into legume genome evolution. Nat. Genet. 2019;51:1411-1422. DOI 10.1038/s41588-019-0480-1.

Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947-2948. DOI 10.1093/bioinformatics/btm404.

Lohse M., Drechsel O., Kahlau S., Bock R. OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 2013;41(Web Server issue):W575-W581. DOI 10.1093/nar/gkt289.

Mikić A., Smýkal P., Kenicer G., Vishnyakova M., Akopian J., Sarukhanyan N., Gabrielyan I., Vanyan A., Toker C., Ćupina B., Ambrose M., Mihailović V., Ellis N. A revival of the research on beautiful vavilovia (*Vavilovia formosa* syn. *Pisum formosum*). Pisum Genet. 2009;41:34-39.

Mikić A., Smýkal P., Kenicer G., Vishnyakova M., Sarukhanyan N., Akopian J., Vanyan A., Gabrielyan I., Smýkalová I., Sherbakova E., Zorić L., Atlagić J., Škorić T., Cupina B., Krstić P., Jajić I., Antanasović S., Dorđević V., Mihailović V., Ivanov A., Ochatt S., Ambrose M. The bicentenary of the research on 'beautiful' Vavilovia (*Vavilovia formosa*), a legume crop wild relative with taxonomic and agronomic potential. Bot. J. Linn. Soc. 2013;172:524-531. DOI 10.1111/boj.12060.

Mikić A., Smýkal P., Kenicer G., Vishnyakova M., Sarukhanyan N., Akopian J.A., Vanyan A., Gabrielyan I., Smýkalová I., Sherbakova E., Zorić L., Atlagić J., Zeremski-Škorić T., Cupina B., Krstić P., Jajić I., Antanasović S., Dorđević V., Mihailović V., Ivanov A., Ochatt S., Toker C., Zlatković B., Ambrose M. Beauty will save the world, but will the world save beauty? The case of the highly endangered *Vavilovia formosa* (Stev.). Fed. Planta. 2014;240(5):1139-1146. DOI 10.1007/s00425-014-2136-9.

Oskoueiyan R., Kazempour-Osaloo S., Maassoumi A.A., Nejadsattari T., Mozaffarian V. Phylogenetic status of *Vavilovia formosa* (Fabaceae-Fabeae) based on nrDNA ITS and cpDNA sequences. Biochem. Syst. Ecol. 2010;38:313-319. DOI 10.1016/j.bse.2010.01.011.

Safronova V.I., Kimeklis A.K., Chizhevskaya E.P., Belimov A.A., Andronov E.E., Pinaev A.G., Tikhonovich I.A., Pukhaev A.R., Popov K.P. Genetic diversity of rhizobia isolated from nodules of the relic species *Vavilovia formosa* (Stev.) Fed. Antonie van Leeuwenhoek. 2014;105:389-399. DOI 10.1007/s10482-013-0089-9.

Safronova V.I., Kuznetsova I.G., Sazanova A.L., Kimeklis A.K., Belimov A.A., Andronov E.E., Pinaev A.G., Chizhevskaya E.P., Pukhaev A.R., Popov K.P., Willems A., Tikhonovich I.A. *Bosea vaviloviae* sp. nov., a new species of slow-growing rhizobia isolated from nodules of the relict species *Vavilovia formosa* (Stev.) Fed. Antonie van Leeuwenhoek. 2015;107:911-920. DOI 10.1007/s10482-015-0383-9.

Schaefer H., Hechenleitner P., Santos-Guerra A., Menezes de Sequeira M., Pennington R.T., Kenicer G., Carine M.A. Systematics, biogeography, and character evolution of the legume tribe Fabeae with special focus on the middle-Atlantic island lineages. BMC Evol. Biol. 2012;12:250. DOI 10.1186/1471-2148-12-250.

Sinjushin A.A., Belyakova A.S. On intraspecific variation of *Vavilovia formosa* (Stev.) Fed. (= *Pisum formosum* (Stev.) Alef.: Fabeae). Pisum Genet. 2010;41:31-34.

Sinjushin A.A., Demidenko N.V., Gostimskii S.A. Preliminary report on taxonomical position of *Vavilovia formosa* (Stev.) Fed. evidenced

from morphological and molecular data. Pisum Genet. 2009;41: 15-20.

Sloan D.B., Alverson A.J., Chuckalovcak J.P., Wu M., McCauley D.E., Palmer J.D., Taylor D.R. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10(1):e1001241. DOI 10.1371/journal.pbio.100124.

Smith D.R., Keeling P.J. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proc. Natl. Acad. Sci. USA. 2015;112:10177-10184. DOI 10.1073/pnas.1422049112.

Tamura K., Stecher G., Peterson D., Kumar S. MEGA6: molecular evolutionary genetics analysisversion 6.0. Mol. Biol. Evol. 2013;30: 2725-2729. DOI 10.1093/molbev/mst197.

Vishnyakova M., Burlyaeva M., Akopian J., Murtazaliev R., Mikić A. Reviewing and updating the detected locations of beautiful vavilovia (*Vavilovia formosa*) on the Caucasus sensu stricto. Genet. Res. Crop Evol. 2016;63:1085-1102. DOI 10.1007/s10722-016-0440-x.

Vishnyakova M.A., Burlyaeva M.O., Seferova I.V., Bagmet L.V., Semenov V.A. Expedition collections of the tribe Vicieae representatives in the Russian Federation and the adjacent territories of the North Caucasus. Bull. Appl. Bot. Genet. Plant Breed. 2013;172: 82-86.

Wojciechowski M.F., Lavin M., Sanderson M.J. A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. Am. J. Bot. 2004;91(11):1846-1862. DOI 10.3732/ajb.91.11.1846.

Yakovlev G.P. Bobovye Zemnogo Shara = Legumes of the Globe. Leningrad: Nauka Publ., 1991. (in Russian)

Zemerski-Škorić T., Mikić A., Sarukhanyan N., Vanyan A., Akopian J., Gabrielyan I., Smýkal P., Kenicer G., Vishnyakova M., Ambrose M. Contributions to the characterization of *Vavilovia formosa* (syn. *Pisum formosum*). III. Contents of macro- and microelements. Pisum Genet. 2010;41:28-30.

Zorić L., Luković J., Mikić A., Akopian J., Gabrielyan I., Sarukhanyan N., Vanyan A., Smýkal P., Kenicer G., Vishnyakova M., Ambrose M. Contributions to the characterization of *Vavilovia formosa* (syn. *Pisum formosum*). I. Anatomy of stem, leaf and calyx. Pisum Genet. 2010;41:21-24.

ORCID ID

O.E. Kosterin orcid.org/0000-0001-5955-4057

Acknowledgements. The work by N.V. Shatskaya, V.S. Bogdanova, O.E. Kosterin and G.V. Vasiliev on the assembly of the plastid and mitochondrial genomes and phylogenetic analysis was supported by State Budgeted Project 0324-2019-0041 for the Institute of Cytology and Genetics SB RAS, Novosibirsk, and the Russian Foundation for Basic Research, project 19-04-00162. The work by E.A. Andronov, A.K. Kimeklis, and N.A. Provorov on DNA isolation and high-throughput sequencing was supported by the Russian Science Foundation, project 19-16-00081.
The assembly of organellar genomes and phylogenetic analysis were performed with the use of the Computational Facility of the Siberian Supercomputer Center SB RAS and the Computational Facility of the Novosibirsk State University.

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

Received April 15, 2019. Revised September 20, 2019. Accepted September 20, 2019.

ГЕНЕТИКА И СЕЛЕКЦИЯ РАСТЕНИЙ
Оригинальное исследование / Original article
УДК 633.353:631.559.2

Вавиловский журнал генетики и селекции. 2019;23(8):981-992
DOI 10.18699/VJ19.36-o

Ecological stability of broad bean (*Vicia faba* L.) in organic farming conditions

N.A. Georgieva ✉, V.I. Kosev

Institute of Forage Crops, Pleven, Bulgaria
✉ e-mail: imnatalia@abv.bg

The main direction in the breeding of legumes is the development of high-productive cultivars with stable yields. In order to assess the ecological stability of 17 broad bean accessions (Fb 1896, Fb 1903, Fb 1929, Fb 2481, Fb 2486, Fb 3270, BGE 002106, BGE 029055, BGE 032012, BGE 041470, BGE 043776, BGE 046721, FbH 13, FbH 14, FbH 15, FbH 16, BGP) with regard to key quantitative traits, a field experiment was conducted in the Institute of Forage Crops (Pleven) in 2016–2018. Plants were grown under organic farming conditions without the use of fertilizers or pesticides. Three types of stability parameters were calculated based on regression, variance, and nonparametric analysis. The results of the variance analysis showed a significant genotype × environment interaction for all quantitative traits except pod width. The factor 'environment' had the greatest impact on the phenotypic manifestation of the traits, followed by the factors 'genotype' and 'genotype × environment interaction'. In terms of plant height and 1st pod height, accessions FbH 16 and FbH 13 can be classified as high (79 cm, 35 cm) and ecologically stable (bi = 0.76, bi = 0.79). BGE 029055 was little variable and had high numbers of pods (15) and seeds (41) per plant. Accessions FbH 14, FbH 16, FbH 15, and BGP were distinguished by high seed weight per plant (from 28.36 to 34.93 g), but they exhibited instability (bi > 1) under unfavorable environmental conditions. In contrast, Fb 1903, BGE 043776, and Fb 3270 were very stable (bi < 1) but low-productive. Accessions Fb 1896, Fb 1929, Fb 2481, Fb 2486, BGE 002106, and BGE 029055 showed intermediate parameters, as they had the coefficient of linear regression close to 1, but they were also low-productive. BGE 041470 appeared to be of special interest for breeding. It had high values of 100 seed mass (101.38 g) and seed weight per plant (32.14 g), being relatively stable (bi = 1.10). GGE biplot analysis determined accessions BGE 046721, BGE 032012, FbH 15 and FbH 16 as a promising breeding material combining high and stable seed yield.
Key words: *Vicia faba* L.; stability parameter; productivity; traits; accessions; regression analysis; nonparametric analysis.

For citation: Georgieva N.A., Kosev V.I. Ecological stability of broad bean (*Vicia faba* L.) in organic farming conditions. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):981-992. DOI 10.18699/VJ19.36-o

Экологическая стабильность боба овощного (*Vicia faba* L.) в условиях органического хозяйства

Н.А. Георгиева ✉, В.И. Косев

Институт кормовых культур, Плевен, Болгария
✉ e-mail: imnatalia@abv.bg

Важное направление в селекции бобовых – создание высокоурожайных сортов со стабильной продуктивностью. Для оценки экологической стабильности 17 образцов *Vicia faba* (Fb 1896, Fb 1903, Fb 1929, Fb 2481, Fb 2486, Fb 3270, BGE 002106, BGE 029055, BGE 032012, BGE 041470, BGE 043776, BGE 046721, FbH 13, FbH 14, FbH 15, FbH 16, BGP) по основным количественным признакам в Институте кормовых культур (г. Плевен, Болгария) в 2016–2018 гг. был проведен полевой эксперимент. Растения выращивали в условиях органического земледелия, без использования удобрений и пестицидов. На основе регрессионного, дисперсионного и непараметрического анализа были рассчитаны три типа параметров стабильности. Результаты дисперсионного анализа доказывают взаимодействие генотип × среда для всех количественных признаков, кроме ширины бобов. Фактор среды оказывает наибольшее влияние на фенотипическое проявление признаков, далее следуют факторы «генотип» и «генотип × среда». По признакам «высота растения» и «высота бобов» образцы FbH 16 и FbH 13 можно отнести к высоким (79 и 35 см соответственно) и экологически стабильным (bi = 0.76, bi = 0.79). BGE 029055 имеет низкую вариабельность, с высокими значениями количества бобов (15) и количества семян с растения (41). Образцы FbH 14, FbH 16, FbH 15 и BGP характеризуются высокой массой семян (от 28.36 до 34.93 г), но проявляют нестабильность (bi > 1) в неблагоприятных условиях среды. Fb 1903, BGE 043776 и Fb 3270, напротив, очень стабильны (bi < 1), однако низкопродуктивны. Промежуточные позиции занимают образцы Fb 1896, Fb 1929, Fb 2481, Fb 2486, BGE 002106 и BGE 029055, у которых коэффициент линейной регрессии приближается к единице, но они также являются низкопроизводительными. Интересен с позиций селекции образец BGE 041470, отличающийся высокими значениями массы 100 семян (101.38 г)

© Georgieva N.A., Kosev V.I., 2019

и массы семян с растения (32.14 г) и относительной стабильностью (bi = 1.10). Результаты анализа двойных диаграмм GGE позволяют расценивать образцы BGE 046721, BGE 032012, FbH 15 и FbH 16 как перспективный материал для селекции, сочетающий высокую и стабильную продуктивность семян.
Ключевые слова: *Vicia faba* L.; параметры стабильности; продуктивность; признаки; образцы; регрессионный анализ; непараметрический анализ.

Introduction

Broad bean (*Vicia faba* L.) is an important legume crop that is primarily used as a protein source in human diet, as forage for animals and as an available source of nitrogen for the biosphere (Rubiales, 2010). Enlargement of broad bean growing is desirable, but it is hampered due to unstable yields (Link et al., 2010).

Yield is a complex quantitative characteristic that is usually controlled by several genes and affected by environmental conditions. The importance of genotype-environment interaction in national programs for cultivar evaluation and breeding programs has been found in most major crops (Kendal et al., 2016; Pouresmael et al., 2018). Researchers investigate and describe stable genotypes by use of different parametric and nonparametric methods or single-variate and multivariate statistical methods (Hamayoon et al., 2011; Imtiaz et al., 2013; Kendal, Sayar, 2016).

In recent years, breeders' interest in legume crops, and especially in broad beans, has increased. Great attention is paid to the study of the starting material, both for the development of new varieties and for the improvement of the existing ones (Kurkina, 2013). In order to identify the best donors and promising candidate cultivars, statistical analysis of as many quantitative traits as possible can be used (Kurkina, 2003; Buravtseva et al., 2009).

The main direction in the breeding of legumes is the development of high-productive cultivars with stable yields over the years (Kazydub, Shamanin, 2003). The terms "plasticity" and "stability" are used to characterize the potential of modification variability and genotypic variability of the individual traits (or their groups). Plasticity (the ability of traits to change), as well as stability under different environmental conditions, are considered to be the basic adaptive properties of living organisms. Ecological plasticity of a genotype (cultivar) is the ability to a sustainable formation of high (compared to other genotypes) yield over a wide range of meteorological and agro-technical conditions (Zhuchenko, 2012; Marakaeva, Kazydub, 2016). Stability is a particularly important characteristic of cultivars in organic production conditions (Konvalina et al., 2009; Georgieva, 2017).

The aim of this study was to evaluate the ecological stability of broad bean accessions regarding main quantitative traits in broad bean in organic farming conditions.

Material and methods

The experimental activity was conducted at the Institute of Forage Crops (Pleven, Bulgaria) during the period 2016–2018. The objects of the study were 17 broad bean (*Vicia faba* L.) accessions of different origins: Portugal (Fb 1896, Fb 1903, Fb 1929, Fb 2481, Fb 2486, Fb 3270), Spain (BGE 002106, BGE 029055, BGE 032012, BGE 041470, BGE 043776, BGE 046721) and Bulgaria (FbH 13, FbH 14, FbH 15, FbH 16, BGP).The randomized block method (Barov, 1982) was used, with a plot size of 4 m^2 and three replications. The sowing was done by hand, at a depth of 8 cm and with a rate of 30 seeds per m^2. Plants were grown under organic farming conditions without the use of fertilizers and pesticides. A biometric estimation was performed, including the following traits: plant height (cm), 1st pod height (cm), pods number per plant, pod lenght (cm), pod width (cm), seeds number per plant, seed weight per plant (g), 100 seeds mass (g).

The obtained data were processed by two-factor analysis of variance for each trait for determine of the influence of genotypes (accessions) (G), environments (E) and their interaction (G × E). The evaluation of the ecological stability of the studied accessions was performed using the following analyzes and parameters: *regression analysis* – according to (Eberhart, Russell, 1966) where the regression coefficient (bi) and the variance of the deviations from regression (Si^2) were calculated; according to (Tai, 1979) in which the parameters a_i (linear response to the environmental effects) and λ_i (deviation from the linear response) were determined; *variance analysis* – mean variance component (PP) according to (Plaisted, Peterson, 1979), ecovalence (W^2) according to (Wricke, 1965), stability variances (σ^2 and W_i) according to (Shukla, 1972) and (Annicchiarico, 1992) respectively; and *non-parametric analysis* – by the rank-parameter (Si_2) on the model of (Huehn, 1990a, b). Plaisted and Peterson's (1979) mean variance component (PP) was a measure of a cultivar's contribution to the G × E interaction and was determined from a total through "pair-wise" analysis. According to Annicchiarico's method (1992), a reliability index (W_i) was calculated which estimates the probability of a particular genotype (cultivar) to show a performance below the environmental average or below any standard used. GGE biplot model was done, which uses decomposition of the singular value of first two principal components (Yan, 2002). All experimental data were processed statistically with using the computer software GENES 2009.7.0 for Windows XP (Cruz, 2009).

Results

Analysis of variance

The results of the two-factor analysis of variance of the traits (Table 1) confirmed the differences in the environmental conditions during the experimental period. With regard to the trait of pod width, the variances of factors were insignificant, on the basis of which it can be considered that genotype differences between the accessions were not essential. For all other traits, the influence of the factors was statistically significant, and the environment had the greatest share. The effect of genotype was stronger pronounced than that of GE interaction for the 1st pod height, pod length and 100 seeds mass. In terms of plant height, the GE interaction was more pronounced. Consequently, for improving this trait, further research will be needed. The statistical significance of the factors genotype, environment and their interactions for almost all traits was an objective prerequisite for determining the stability parameters.

Table 1. Mean squares in analysis of variance of broad bean accessions regarding main quantitative traits

Source of variation	DF	Plant height	1st pod height	Pods number	Seeds number	Pod lenght	Pod width	100 seeds mass	Seed weight
Environment	2	12851.1*	690.94*	286.32*	2926.02*	113.83*	1.20	35356.36*	9731.29*
Genotype	16	177.13*	147.74*	34.32*	289.38*	17.72*	0.24	2292.23*	310.92*
G×E	32	182.05*	72.12*	23.55*	263.85*	3.18***	0.08	709.42*	239.80*
Env/Gen	34	927.29*	108.52*	39.012*	420.45*	9.69*	0.14	2747.48*	848.94*
Env/Gen 1	2	605.38*	92.03*	23.24*	58.24*	1.30	0.13	5874.38*	608.41*
Env/Gen 2	2	1122.58*	156.85*	32.18*	273.06*	2.65*	0.05	452.619*	343.72*
Env/Gen 3	2	732.47*	205.65*	23.94*	136.60*	6.28**	0.10	3476.60*	446.80*
Env/Gen 4	2	1051.84*	283.73*	23.23*	149.73*	1.78	0.01	3253.31*	436.17*
Env/Gen 5	2	903.05*	109.54*	17.44*	70.98*	9.37***	0.16	4837.00*	831.81*
Env/Gen 6	2	3035.11*	516.04*	12.72*	419.25*	2.57*	0.04	597.64*	59.35*
Env/Gen 7	2	903.00**	3.64**	20.17*	600.13*	17.01*	0.19	225.16*	455.60*
Env/Gen 8	2	951.12***	9.35***	29.57*	610.76*	2.38	0.11	295.67*	563.18*
Env/Gen 9	2	1326.65*	81.75*	17.65*	376.83*	15.97*	0.18	2068.95*	694.54*
Env/Gen 10	2	1073.02*	83.75*	19.26*	266.96*	17.48*	0.05	1285.91*	794.75*
Env/Gen 11	2	1380.87*	70.74*	7.68**	36.51*	1.40	0.43	1209.42*	93.13*
Env/Gen 12	2	700.98*	44.45*	5.21**	3.46**	7.83**	0.21	3452.99*	383.41*
Env/Gen 13	2	469.91*	47.73*	6.51**	306.33*	14.00*	0.22	3273.78*	1014.93*
Env/Gen 14	2	660.35*	11.43**	345.16*	2023.68*	17.02*	0.08	3474.64*	2971.63*
Env/Gen 15	2	142.95*	2.34*	20.34*	710.91*	14.30*	0.17	2949.75*	1652.84*
Env/Gen 16	2	477.29*	14.25*	24.62*	658.21*	16.51*	0.01	5105.18*	1505.32*
Env/Gen 17	2	227.38*	111.63*	34.21*	446.05*	16.83*	0.33	4874.15*	1576.42*
Total	50	1							

Note: Genotypes: 1, Fb1896; 2, Fb1903; 3, Fb 1929; 4, Fb 2481; 5, Fb 2486; 6, Fb 3270; 7, BGE 002106; 8, BGE 029055; 9, BGE 032012; 10, BGE 041470; 11, BGE 043776; 12, BGE 046721; 13, FbH 13; 14, FbH 14; 15, FbH 15; 16, FbH 16; 17, BGP.
* Significant at $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Evaluation of stability

The influence of environmental conditions on the trait formation was expressed by the values of the stability parameters calculated by the different methods (Table 2 and Table 3). Regarding the plant height, when comparing the accessions, it was found that the FbH 13 and FbH 16 (79 cm) had the greatest height, followed by Fb 3270 and BGE 043776 (78 cm) (Fig. 1). For the experimental conditions, BGP and FbH 15 were the most stable. They were characterized by very low values of bi (bi = 0.24; bi = 0.31), as in FbH 15, the value of Si^2 was considerably lower (Si^2 = 43.97).

The accessions Fb1896, Fb 1929, BGE 046721 and FbH 14 were distinguished by mean stability for bi (bi = 0.86–0.95) and low variance. From the highest plant accessions, particular attention should be paid to FbH 16 and FbH 13, which were definitely stable (bi = 0.76; bi = 0.79) and weakly affected by the environmental conditions, unlike Fb 3270 and BGE 043776, whose plants were also high but they exhibited strong variability. According to the values of the regression coefficient, BGE 002106, BGE 029055, Fb 2481 and Fb 2486 were stable, with height below the average for the group of tested accessions, and BGE 032012 and BGE 041470 exhibited plasticity with respect to this trait due to its high value (bi = 1.18–1.32).

The trait of 1st pod height had high values in FbH 16, Fb 2481 and FbH 13, followed by Fb 1929 and BGE 046721. According to stability data, FbH 16 exhibited high stability (bi = 0.06; σ^2 = 4.17; W_i = 103.31), while Fb 2481 and Fb 1929 (considering the coefficient of linear regression bi as well as the parameter PP) were characterized by very well-expressed responsiveness (bi > 1; PP = 33.65–36.85). Similarly was the behavior of Fb 1896, Fb 1903, Fb 2486, Fb 3270, BGE 032012, BGE 041470 and BGE 043776, but they set their 1st pods at a relatively lower height. Plants of FbH 13 formed their first pods highly and, according to most stability parameters, showed very low variability.

The largest number (15) of pods per plant was obtained from accessions BGE 029055 and FbH 14 (see Fig. 1). The following positions were occupied by Fb 2486, BGE 043776, BGP (12 pods per plant). According to the stability data of the considered trait (see Table 2), all parameters defined FbH 14 as unstable. It is notable that accessions that formed a larger number of pods (11–15) were highly variable and responsive to improvement in growing conditions. An exception was

Table 2. Parameters of phenotypic stability in broad bean accessions regarding main quantitative traits

Accessions	bi	Si^2	a_i	λ_i	σ^2	PP	W^2	W_i
	Eberhart, Russell, 1966		Tai, 1979		Shukla, 1972	Plaisted, Peterson, 1979	Wricke, 1965	Annicchiarico, 1992
				Plant height				
Fb 1896	0.86**	24.51**	0.8669	39.6022	14.69	38.9799	101.3134	96.6814
Fb 1903	1.06	184.23**	1.0578	294.1470	936.34	79.4805	558.7322	91.1648
Fb 1929	0.93	52.96**	0.9291	84.9413	39.15	44.8390	167.4867	91.5588
Fb 2481	1.09	100.43**	1.0916	160.6005	142.12	57.8982	314.9792	91.9984
Fb 2486	1.09	3.10**	1.0899	5.4691	71.45	32.0026	22.5111	86.7744
Fb 3270	1.94**	116.75**	1.9450	186.5475	–8.33	180.6308	1701.1350	88.8600
BGE 002106	1.08	14.99**	1.0789	24.4243	38.59	34.9145	55.3980	84.1335
BGE 029055	1.10*	14.60**	1.1085	23.7940	351.60	35.5501	62.5767	90.7561
BGE 032012	1.32**	0.74	1.3240	1.7015	47.84	44.3418	161.8715	93.4103
BGE 041470	1.18**	8.91**	1.1838	14.7143	54.03	36.9819	78.7482	90.5619
BGE 043776	1.35**	0.27	1.3512	0.9337	245.44	46.6715	188.1837	102.2922
BGE 046721	0.95	12.36**	0.9498	20.2371	205.40	33.7193	41.9000	97.6550
FbH 13	0.76**	20.34**	0.7619	32.9499	46.44	43.0864	147.6932	104.3693
FbH 14	0.90*	31.41**	0.9003	50.6029	23.40	39.7729	110.2691	95.6382
FbH 15	0.31**	43.97**	0.3181	70.6230	19.58	104.0226	835.9135	88.7764
FbH 16	0.79**	0.03	0.7946	0.0476	118.85	35.6631	63.8529	105.7174
BGP	0.24**	119.96**	0.2491	191.7374	0.33	137.4324	1213.2480	90.0438
				1st pod height				
Fb 1896	1.29	15.77**	1.2928	25.6599	24.01	16.5806	55.2583	86.9871
Fb 1903	1.92**	4.24**	1.9224	7.2371	186.74	19.0091	82.6858	89.5109
Fb 1929	1.50*	75.50**	1.5045	120.8449	203.76	33.6578	248.1301	97.6634
Fb 2481	2.24**	52.64**	2.2437	84.3306	25.86	36.8590	284.2840	100.6840
Fb 2486	1.58**	4.93**	1.5823	8.3652	32.82	15.5190	43.2677	72.4388
Fb 3270	3.25**	57.62**	3.2526	92.0538	7.25	63.4964	585.1307	83.5793
BGE 002106	0.23**	0.69	0.2262	1.5986	70.32	16.2578	51.6120	81.6485
BGE 029055	0.26**	4.01**	0.2658	6.8280	110.24	24.3403	142.8964	63.9726
BGE 032012	1.22	13.71**	1.2223	22.3807	47.56	15.7731	46.1379	95.1409
BGE 041470	1.26	12.67**	1.2577	20.7200	6.71	15.6185	44.3912	76.0487
BGE 043776	1.32	0.33	1.3196	0.0067	0.67	12.4233	8.3049	91.5196
BGE 046721	0.78	12.80**	0.7802	20.9221	3.97	15.5222	43.3041	105.1743
FbH 13	1.08	0.32	1.0838	0.0164	161.84	11.7412	0.6005	115.3304
FbH 14	0.18**	6.39**	0.1811	10.6762	13.94	18.2874	74.5346	101.0185
FbH 15	0.23	0.23	0.2309	0.1270	36.65	15.9604	48.2529	98.1128
FbH 16	0.06**	0.81	0.5578	1.6707	4.17	29.4062	200.1110	103.3181
BGP	0.07**	72.05**	0.2763	115.2605	66.35	42.6060	349.1913	79.2728
				Pods number				
Fb 1896	1.13	0.74	1.1336	1.7139	11.73	3.9314	3.8258	99.3485
Fb 1903	0.58	17.32**	0.5807	28.1234	28.86	8.8026	58.8420	70.7805
Fb 1929	1.19	0.21	1.1884	0.1928	28.67	3.7303	1.5543	67.0757
Fb 2481	1.14	0.50	1.1430	1.3265	1.64	3.8745	3.1832	82.9719
Fb 2486	0.99	0.35	0.9874	1.0895	228.57	3.7748	2.0565	109.5617
Fb 3270	0.07**	1.51*	0.7752	2.7405	14.35	13.4150	110.9343	72.4880
BGE 002106	0.98	2.30**	0.9812	4.2002	21.31	4.2938	7.9188	84.8447
BGE 029055	1.27	1.19*	1.2739	2.4223	0.55	4.2193	7.0777	131.0650

Table 2 (continued)

Accessions	bi	Si^2	a_i	λ_i	σ^2	PP	W^2	W_i
	Eberhart, Russell, 1966		Tai, 1979		Shukla, 1972	Plaisted, Peterson, 1979	Wricke, 1965	Annicchiarico, 1992
BGE 032012	1.02	0.31	1.0226	0.0429	41.00	3.6013	0.0978	85.2169
BGE 041470	1.07	0.25	1.0662	0.1360	28.64	3.6284	0.4031	72.4842
BGE 043776	0.40	2.99**	0.3974	5.2887	23.82	5.5536	22.1467	106.2217
BGE 046721	0.38	1.49	0.3816	2.8798	1.79	5.2091	18.2556	74.5156
FbH 13	0.53	0.87	0.5271	1.9108	19.93	4.5757	11.1025	88.4313
FbH 14	4.27**	25.27**	4.2791	40.1479	21.47	42.2413	436.5020	83.2290
FbH 15	0.77	6.54**	0.7715	10.9442	28.78	5.5718	22.3528	82.5891
FbH 16	0.95	5.89**	0.9521	9.9328	0.15	5.2550	18.7748	86.7098
BGP	1.09	9.17**	1.0891	15.1406	17.43	6.1396	28.7652	99.9964
				Seeds number				
Fb 1896	0.33**	26.11**	0.3284	42.1215	42.46	64.4062	234.4991	78.2605
Fb 1903	0.65**	133.84**	0.6458	213.8359	113.48	83.1058	445.6945	60.3994
Fb 1929	0.89	0.33	0.8908	–0.0018	175.39	44.0063	4.1004	70.1788
Fb 2481	0.93	0.14	0.9318	0.3142	212.15	43.8375	2.1940	68.6338
Fb 2486	0.23**	41.06**	0.2269	65.9392	737.04	72.8423	329.7770	93.4250
Fb 3270	0.31**	83.03**	0.3082	132.5242	248.78	228.0640	2082.8700	66.4390
BGE 002106	1.77**	42.03**	1.7659	67.4807	13.19	72.7638	328.8903	85.7961
BGE 029055	1.86**	9.86**	1.8603	16.2036	59.02	68.8939	285.1836	121.1575
BGE 032012	1.43**	16.88**	1.4282	27.4285	182.01	53.8015	114.7286	92.6802
BGE 041470	1.17	19.89**	1.1726	32.2421	93.13	49.9243	70.9385	94.6878
BGE 043776	0.03**	23.89**	0.0336	38.5313	83.77	82.6130	440.1294	66.6970
BGE 046721	0.13**	0.08	0.1282	0.6190	708.10	66.9059	262.7313	92.4008
FbH 13	0.92	106.55**	0.9210	170.3375	317.61	72.2229	322.7817	74.6468
FbH 14	2.86**	415.14**	2.8531	661.9547	327.45	258.6003	2427.7500	67.6624
FbH 15	1.67**	155.14**	1.6662	247.7587	42.03	98.4585	619.0894	87.1802
FbH 16	1.95**	1.42*	1.9520	2.7444	123.01	71.7127	317.0196	62.5495
BGP	1.57**	14.03**	1.5706	22.8795	218.19	57.3785	155.1271	126.6104
				Pod length				
Fb 1896	0.44	0.31	0.4310	0.0130	81.62	0.5798	4.3228	86.9333
Fb 1903	0.59	0.12	0.5872	0.3292	467.37	0.4522	2.8821	93.9952
Fb 1929	0.82	0.84	0.8200	1.8710	81.77	0.5470	3.9528	89.4074
Fb 2481	0.47	0.13	0.4660	0.3045	566.13	0.5830	4.3589	82.3804
Fb 2486	1.05	1.01	1.0495	2.1340	722.30	0.5556	4.0493	92.5431
Fb 3270	0.12	1.32*	0.1138	2.5855	212.64	1.5511	15.2931	62.7822
BGE 002106	1.59	0.21	1.5903	0.1739	38.76	0.6356	4.9527	94.2557
BGE 029055	0.56	0.12	0.5511	0.3297	641.46	0.4889	3.2962	80.6047
BGE 032012	1.52	0.02	1.5261	0.4861	10.04	0.6034	4.5894	104.2784
BGE 041470	1.55	0.67	1.5498	1.5747	261.03	0.8148	6.9767	129.3576
BGE 043776	0.43	0.23	0.4260	0.1475	112.51	0.6089	4.6519	77.4786
BGE 046721	0.84	1.74*	0.8391	3.2989	18.40	0.7774	6.5539	127.0036
FbH 13	1.02	4.36**	1.0200	7.4793	88.38	1.4441	14.0837	82.2417
FbH 14	1.54	0.47	1.5422	1.2566	75.16	0.7519	6.2667	89.6194
FbH 15	1.37	0.81	1.3744	1.8112	84.77	0.6636	5.2693	90.2566
FbH 16	1.57	0.31	1.5738	0.0197	8.28	0.5873	4.4073	99.0947
BGP	1.54	0.36	1.5402	1.0922	19.64	0.7221	5.9298	91.1550

Table 2 (end)

Accessions	bi	Si^2	a_i	λ_i	σ^2	PP	W^2	W_i
	Eberhart, Russell, 1966		Tai, 1979		Shukla, 1972	Plaisted, Peterson, 1979	Wricke, 1965	Annicchiarico, 1992
				100 seeds mass				
Fb 1896	1.68**	3.02**	1.6800	5.2232	0.09	289.0503	1932.9440	112.0244
Fb 1903	0.47**	4.59**	0.4627	7.8220	6.34	225.5393	1215.6430	108.1333
Fb 1929	1.29**	0.33	1.2930	–0.0869	7.36	149.5031	356.8807	95.0492
Fb 2481	1.24**	28.79**	1.2423	46.3377	29.13	147.2577	331.5209	93.5870
Fb 2486	1.35**	685.76**	1.3531	1093.4540	–0.42	346.0680	2576.9080	90.3787
Fb 3270	0.52**	26.50**	0.5177	42.8251	1.50	210.7062	1048.1160	38.9100
BGE 002106	0.29**	33.31**	0.2898	53.5880	2.41	312.5915	2198.8200	78.9987
BGE 029055	0.27**	99.07**	0.2655	158.6011	10.40	343.0164	2542.4430	75.0066
BGE 032012	0.95	121.02**	0.9525	193.4120	23.81	150.9705	373.4545	92.5282
BGE 041470	0.73**	108.76**	0.7346	173.9030	38.61	172.8278	620.3135	110.6985
BGE 043776	0.76**	3.12**	0.7609	5.5056	2.60	139.8718	248.1039	91.2702
BGE 046721	1.21**	260.45**	1.2134	415.5642	3.81	203.9377	971.6716	86.1608
FbH 13	1.21**	164.59**	1.2063	262.8200	–0.31	177.3867	671.8013	87.0916
FbH 14	1.26**	120.22**	1.2585	192.1197	–0.01	174.5284	639.5201	96.0861
FbH 15	1.18**	27.04**	1.1827	43.5746	5.37	137.4564	220.8246	90.3847
FbH 16	1.56**	38.68**	1.5578	62.0823	17.32	242.8314	1410.9410	99.7928
BGP	1.03	1779.01**	1.0298	2835.8290	0.18	590.8681	5341.7100	60.5632
				Seed weight				
Fb 1896	0.96	50.19**	0.9646	80.5250	0.01	62.1813	153.0098	99.9013
Fb 1903	0.45**	151.91**	0.4489	242.6463	0.19	119.8607	804.4483	77.7036
Fb 1929	0.88*	0.20	0.8832	0.2065	0.08	50.0498	15.9948	71.2676
Fb 2481	0.87*	3.43	0.8672	5.9933	0.02	51.4193	31.4623	68.6158
Fb 2486	0.84**	286.87**	0.8370	457.7392	0.01	127.6183	892.0627	97.1530
Fb 3270	0.05**	38.11**	0.0542	61.1723	0.01	149.4923	1139.1100	30.5370
BGE 002106	0.87*	12.46**	0.8732	20.3965	0.01	53.6636	56.8105	93.8915
BGE 029055	0.89	66.65**	0.8991	106.7460	0.44	67.4573	212.5978	105.2589
BGE 032012	1.10	0.31	1.1015	0.0368	0.04	49.6832	11.8554	99.4171
BGE 041470	1.10*	55.42**	1.1146	88.8561	0.20	64.7745	182.2975	112.5134
BGE 043776	0.30**	27.19**	0.3008	43.8128	0.01	105.4840	642.0755	61.6103
BGE 046721	0.77**	31.03**	0.7666	49.9871	0.12	62.4878	156.4716	96.1664
FbH 13	1.15*	170.29**	1.1515	271.9284	0.05	96.2819	538.1455	63.2004
FbH 14	2.10**	301.70**	2.0977	481.2721	0.34	250.9772	2285.2930	75.4486
FbH 15	1.56**	176.57**	1.5569	281.9522	0.11	127.0596	885.7535	76.3140
FbH 16	1.62**	0.90	1.6207	1.9782	0.20	88.0090	444.7108	68.7624
BGP	1.46**	234.63**	1.4623	374.4689	0.16	132.7076	949.5417	78.5726

BGE 029055, in which most of the stability parameters characterized it as relatively stable.

With respect to the number of seeds per plant, BGE 029055, BGP and FbH 14 can be distinguished, which formed 35–41 seeds per plant. As in the previous trait, and in this one, accessions with an increased number of seeds showed clearly pronounced instability. However, the values of the parameters Si^2 and PP gave reason to consider that BGE 029055 (Si^2 = 9.85; PP = 68.89) and BGP (Si^2 = 14.03; PP = 57.37) had a breeding value. They had a high adaptive ability, which in conditions above the averages, could provide them with a high realization of the number of seeds per plant. Accessions Fb 2486 and BGE 046721 were determined as very stable, with the coefficient of linear regression significantly smaller than 1, parameter W_i close to 100 and number of seeds above the average for the group of tested accessions. Their low adap-

Table 3. Rank-parameter (Si_2) of the studied traits according to the model of Huehn (1990a, b)

Accessions	Plant height	1st pod height	Pods number	Seeds number	Pod lenght	100 seeds mass	Seed weight
Fb 1896	9	7	8	8	11	12	6
Fb 1903	11	9	14	11	7	11	12
Fb 1929	11	11	4	3	10	6	3
Fb 2481	10	15	7	2	11	6	5
Fb 2486	3	6	7	10	10	12	12
Fb 3270	14	14	14	12	13	9	12
BGE 002106	8	8	8	9	11	14	5
BGE 029055	7	12	8	10	7	12	7
BGE 032012	9	8	4	7	7	7	4
BGE 041470	8	8	4	6	11	8	6
BGE 043776	10	4	11	10	7	5	11
BGE 046721	5	8	11	9	10	9	9
FbH 13	10	2	11	10	11	8	9
FbH 14	9	10	17	17	11	8	16
FbH 15	11	7	12	12	10	5	14
FbH 16	8	12	11	11	12	10	10
BGP	13	15	12	9	11	13	14

Note: Si_2 = 1: better performance; Si_2 = 17: worst performance.

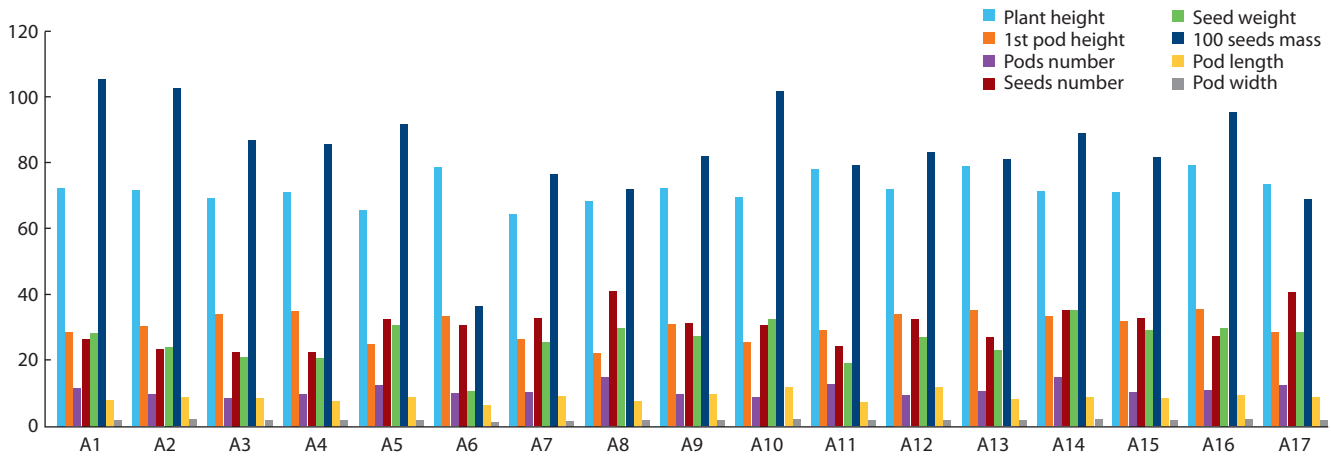


Fig. 1. Main quantitative traits in broad bean accessions (2016–2018).

Accessions: 1, Fb1896; 2, Fb1903; 3, Fb 1929; 4, Fb 2481; 5, Fb 2486; 6, Fb 3270; 7, BGE 002106; 8, BGE 029055; 9, BGE 032012; 10, BGE 041470; 11, BGE 043776; 12, BGE 046721; 13, FbH 13; 14, FbH 14; 15, FbH 15; 16, FbH 16; 17, BGP.

tive ability implied that they were not suitable for growing at a high level of agro-technology, but they can exhibit well in unfavorable environmental conditions.

Because of the lack of significance in the stability parameters of Eberhart and Russell (1966) regarding the trait of pod length for all accessions, the complex evaluation of ecological stability was made on the basis of the remaining parameters. According to the parameter values based on variance analysis and nonparametric method of Huehn (1990a, b), Fb 1903 and BGE 029055 outlined as the most stable, although they were one of the accessions with the shortest pods (see Table 3). A breeding compromise between the value of the trait and its stability (according to the indicator σ^2) can be found in BGE 046721 (σ^2 = 18.4) and BGE 032012 (σ^2 = 10.04). The ambiguous assessments according to the different types of parameters show that longer-term researches are needed to determine the environmental sustainability of this trait.

In terms of 100 seeds mass, Fb 1896 (105.48 g), Fb 1903 (102.34 g) and BGE 041470 (101.38 g) were clearly pronounced favorites, characterized by a high value of the trait (see Fig. 1). A comparison of the three types of stability parameters showed that Fb 1896 was unstable and largely variable (see Table 2). Its reaction to the environment is highly predictable. It is suitable for growing in favorable environ-

mental conditions. If conditions deteriorate, it can give way to the other accessions and to form seeds with smaller mass. Accession Fb 1903 also had a high mass of 100 seeds, but its coefficient of linear regression bi was less than 1, so it can be defined as stable.

BGE 041470 was characterized by the fact that the coefficient bi (bi = 0.73) was closer to 1 (compared to Fb 1903), which determined it as closest to the "ideal" genotype for this trait. The stability parameters of Plaisted and Peterson (1979) and Annicchiarico (1992) confirmed the evaluation according to Eberhart and Russell (1966) and Tai (1979). This accession is of interest from a breeding point of view, combining a high mass of 100 seeds with ecological stability. It is suitable for growing in a wide range of environmental conditions.

Table 2 also presents data on the stability parameters of the seed weight per plant. FbH 14, FbH 16, Fb H 15 and BGP had high values of the regression coefficient (bi ≥ 1.5), which determined them as unstable. With close average productivity (see Fig. 1) and similar values for bi (bi = 1.10–1.15) were accessions FbH 13 and BGE 032012. More plastic, i. e. more responsive to favorable growing conditions was FbH 13. Plants of BGE 041470 formed seeds with relatively large mass and were characterized by a coefficient of linear regression slightly above 1. Because of these characteristics, this accession was of interest from a breeding point of view.

Accessions Fb 1896, Fb 1929, Fb 2481, Fb 2486, BGE 002106 and BGE 029055 had coefficients close to 1, but they were distinguished by low productivity. Fb 1903 and BGE 043776 were stable under changing environmental conditions, with regression coefficient values of bi = 0.45 and bi = 0.30, respectively. In addition, BGE 043776 was also characterized by a low variance, i. e. its empirical values approximate to the theoretical ones. The most stable (bi = 0.05), but with the lowest value of the considered trait was Fb 3270.

GGE biplot analysis

The differences in stability of the accessions and their adaptability to the environment are presented in a two-dimensional coordinate system (Fig. 2). The axial axis shows the main effects (environment and genotype), and the ordinate axis – the effects of their interaction (PC1 or PC2). On the vertices of the polygon are situated the accessions, which are furthest from the center. All accessions fall into this polygon. The lines separating the polygon of sectors represent a set of hypothetical environments. The accession, forming the polygon angle for each sector, has the highest value regarding the studied trait for the environment that falls within this sector.

In terms of the plant height, from the location of the accessions and environments (years), it is obvious that there is variability due to the influence of the environment and especially the climatic difference between the second (2017) and third (2018) experimental years. This result is in full compliance with the data from the variance analysis (see Table 1). The accessions, whose projections on the ordinate were close to the origin of the coordinate system, were stable and very weakly interact with the environment. These were Fb 1896, Fb 3270, BGE 002106, Fb 1929, BGP, FbH 15. By contrast, accessions Fb 1903 and BGE 043776 were the most distant therefore they were characterized by a specific reaction to the environment and low adaptation.

The result regarding the 1st pod height indicated that the first two components (PC1 and PC2) determined 90.3 % of the total variation of the trait as a result of GE interaction. The vertices of the polygon were occupied by Fb 1929, Fb 1903, BGE 002106, BGP, FbH 13, FbH 14 and FbH 15. Accessions FbH 16, BGE 043776 and BGE 002106 exhibited their biological potential the best during the third year. They were among the genotypes with the greatest environmental stability. Fb 2486, Fb 3270 and BGE 041470 developed relatively well in the first and second years. The location of the first year (closer to the center of the coordinate system) determined it as the most suitable environment for the broad bean development with respect to the considered trait.

The accessions, which were characterized by a number of pods above the average, had positive values of the first principal component (PC1 > 0) and a value of PC2, tending to zero, were suitable for growing in different environments. In this trait, the accessions reacted differently to changing environmental conditions. In the most favorable position were those located in the quadrant, limited by the positive part of the abscissa (PC1) and the ordinate (PC2). From the accessions with a larger number of pods (FbH 14, BGP and BGE 029055), BGE 029055 can be preferred because of its lower variability. Fb 1903 and Fb 3270 were in the group of stable genotypes, however, they formed relatively few pods (10). A considerable part of the accessions had high (positive or negative) values of PC2, so they can be identified as having specific adaptability to the environment.

The graphical analysis of the stability in terms of the number of seeds per plant gave an analogous assessment obtained from the stability parameters. BGE 029055, BGP and FbH 14 had an increased number of seeds and occupied an extreme right position of the coordinate system, which determined them as the most productive and at the same time the most variable. Accessions Fb 1896 and Fb 1929 were one of the lowest productive, and their location close to the center of the coordinate system suggests a slight influence of the environment on the trait manifestation. In close proximity to them, with almost the same stability, were accessions Fb 2486 and BGE 046721. In breeding terms, they were more valuable because they formed a larger number of seeds.

GGE biplot analysis regarding 100 seeds mass showed that Fb 1929, BGE 032012 and BGE 043776 were located close to the origin of the coordinate system. They had low values on PC2 and PC1. This placement defined them as very stable, but they had a small mass of 100 seeds. BGE 041470 also had a small projection on PC2, but it was with a much higher value on the axial axis due to its larger mass of 100 seeds. Accessions Fb 1896 and Fb 1903 were distinguished with a high mass of 100 seeds and specific adaptability.

With regard to the trait of seed weight per plant, the first position was occupied by BGE 041470, characterized by a very good combination between stability and level of the trait. Accessions Fb 2486, FbH 16, FbH 14, Fb 1896 and BGE 029055 were distinguished by relatively good productivity and stability, and can be recommended as suitable for growing conditions. The most pronounced variability of the considered trait had FbH 13 and FbH 15, and BGE 032012 was characterized by the best stability but with very low productivity.

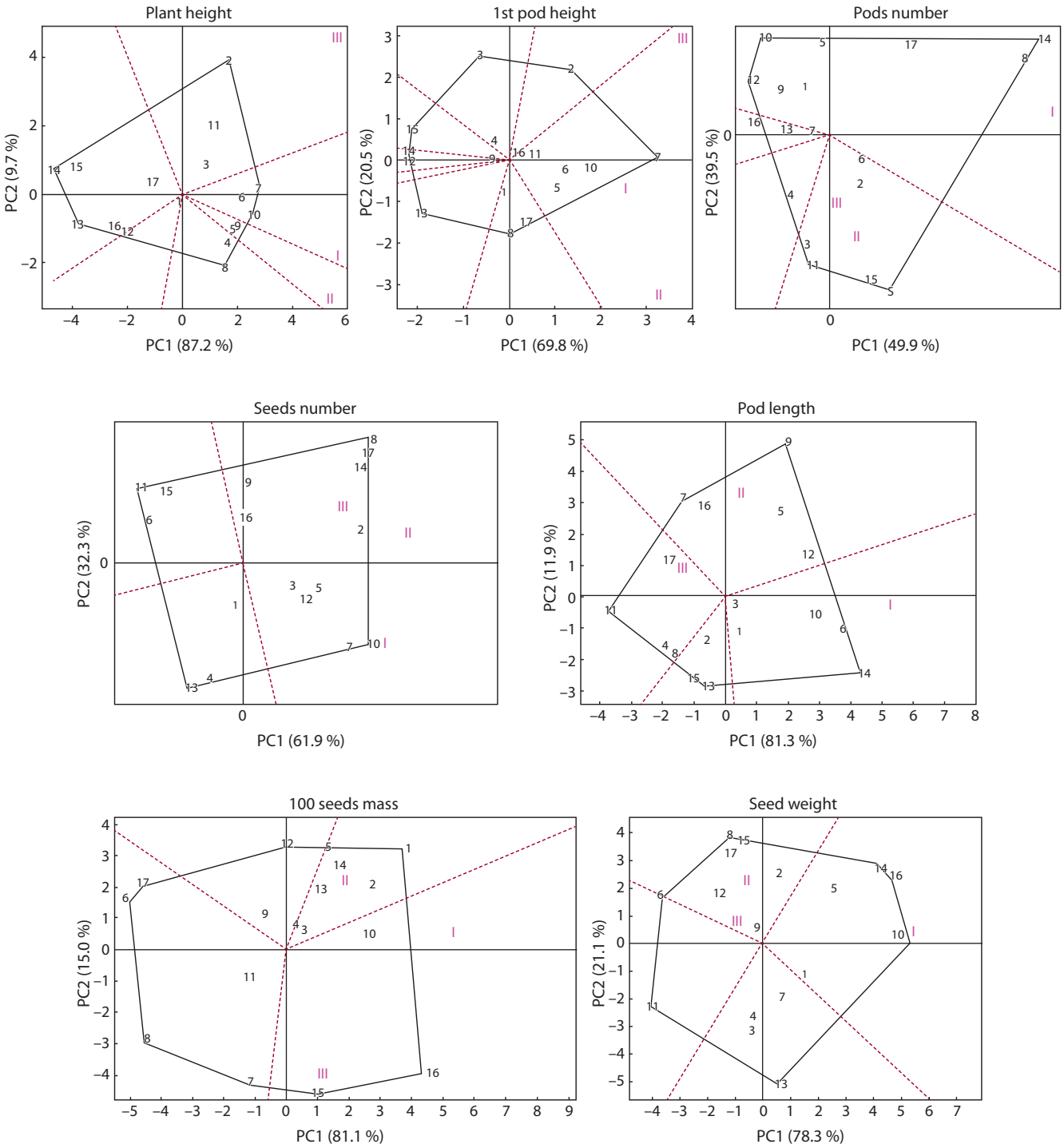


Fig. 2. GGE biplot analysis regarding the studied traits in broad bean accessions.

Accessions: 1, Fb1896; 2, Fb1903; 3, Fb 1929; 4, Fb 2481; 5, Fb 2486; 6, Fb 3270; 7, BGE 002106; 8, BGE 029055; 9, BGE 032012; 10, BGE 041470; 11, BGE 043776; 12, BGE 046721; 13, FbH 13; 14, FbH 14; 15, FbH 15; 16, FbH 16; 17, BGP.

In Figure 3, *a*, the X axis (or the line of the average yield of the broad bean accessions) passes through the beginning of the coordinate system with an arrow indicating the positive end of the axis. The axis Y of the coordinate system (axis of stability) also passes through the beginning of the coordinate system perpendicular to the X axis. Thus, the average yield of the accessions can be estimated from the projection of their markers along the X axis, and the stability – from the projection along the Y axis. Accessions FbH 14, BGE 041470, Fb 2486 had the highest average yield, and Fb 3270, BGE 043776, Fb 2481 and Fb 1929 – the lowest one. The yield in FbH 14, BGE 041470, BGE 029055, Fb 1896 and Fb 1929 was the most variable, whereas Fb 2481, FbH 16 and BGE 046721 were distinguished with high stability.

The graphical analysis also enables to be determined accessions that combine high yield and stability. The center of

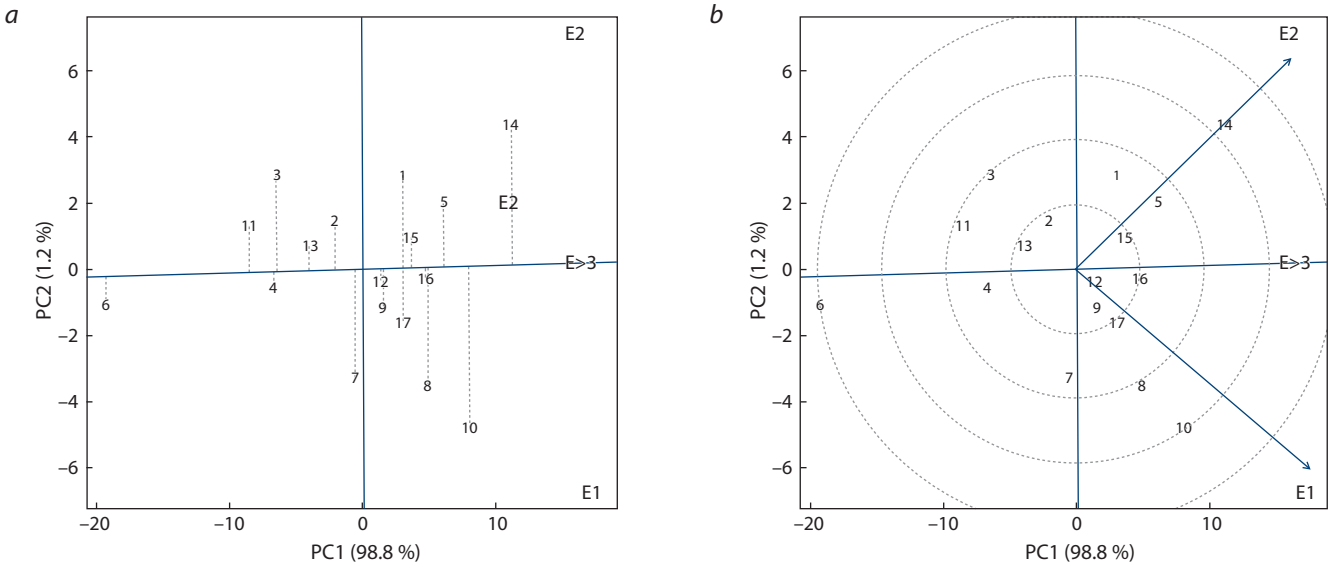


Fig. 3. GGE biplot analysis regarding seed yield in broad bean accessions.

Accessions: 1, Fb1896; 2, Fb1903; 3, Fb 1929; 4, Fb 2481; 5, Fb 2486; 6, Fb 3270; 7, BGE 002106; 8, BGE 029055; 9, BGE 032012; 10, BGE 041470; 11, BGE 043776; 12, BGE 046721; 13, FbH 13; 14, FbH 14; 15, FbH 15; 16, FbH 16; 17, BGP.

Table 4. Correlation coefficients between stability parameters and seed weight per plant in broad bean accessions

Parameter	bi	Si^2	a_i	λ_i	σ^2	PP	W^2	W_i	Si_2
Si^2	0.442								
a_i	0.99**	0.442							
λ_i	0.442	0.99**	0.442						
σ^2	0.465	0.273	0.465	0.273					
PP	0.348	0.769**	0.348	0.769**	0.309				
W^2	0.348	0.769**	0.348	0.769**	0.309	0.99**			
W_i	0.248	0.016	0.248	0.016	0.290	–0.395	–0.395		
Si_2	0.118	–0.299	0.263	0.759**	0.258	0.882**	0.882**	–0.371	
Seed weight per plant	0.530*	0.226	0.530*	0.226	0.369	0.111	0.111	0.335	0.162

Note: bi, Si^2 – Eberhart and Russell (1966); a_i, λ_i – Tai (1979); σ^2 – Shukla (1972); PP – Plaisted and Peterson (1979); W^2 – Wricke (1965); W_i – Annicchiarico (1992); Si_2 – Huehn (1990a, b). Significant at ** $p < 0.01$, * $p < 0.05$.

the concentric circles (Fig. 3, *b*) presents the position of the "ideal" genotype, which is defined by the projection of the environment axis determined by the longest vector of genotypes with yield above the average and lower projection of the perpendicular line (low variability in all environments). More preferred is this one, which is closer to the "ideal" genotype. In our study, BGE 046721, BGE 032012, FbH 13, Fb 1903, FbH 15 and FbH 16 were located closer to the center of the concentric circles, so they can be defined as close to the ideal type with regard to the magnitude and variability of seed yield compared to the rest accessions.

Correlation analysis

The results of the correlation dependences between seed weight per plant and the stability parameters were presented in Table 4. The average productivity of the accessions positively, strongly and significant correlated with the parameters bi ($r = 0.530$) and a_i ($R = 0.530$). For the remaining parameters (especially PP, W^2 and Si^2), the correlation was also positive but statistically insignificant and weaker.

The coefficient of regression bi interacted positively with the other parameters, especially with the parameter a_i ($r = 0.990$). The parameter Si^2 (Eberhart, Russel, 1966) was in positive dependence with λ_i ($r = 0.99$), PP ($r = 0.769$) and W^2 ($r = 0.769$). Strong positive correlations at a high level of statistical significance were found between W^2 and the parameters λ_i ($r = 0.769$) and PP ($r = 0.990$).

The only negative correlations were established between the parameter Si_2 (Huehn, 1990a, b) and Si^2 (Eberhart, Russel, 1966) ($r = –0.299$) and W_i ($r = –0.371$), as well as between W_i and respectively PP ($r = –0.395$) and W^2 ($r = –0.395$).

Discussion

According to (Ilchovska, 2008), for a more objective assessment of the genotype-environment interaction, the studied cultivars (genotypes) should be studied in a wider range of en-

vironmental conditions. Some researchers (Yan, Tinker, 2006; Farshadfar et al., 2011) considered that in order to monitor the model of interaction between genotype and environment, and to better interpret the results, it is appropriate to use the biplot analysis. This polygonal graphic model represented the behavior and the interaction of genotypes in different environments, the differentiation of environment and also the specific environment for each of the genotypes studied.

The data obtained in this study regarding the interaction of genotype and environment confirmed the results of researches in other crops. In spring vetch, Sayar (2017) found that the factor 'environment' had the greatest impact on the yield, followed by the GE interaction and the factor 'genotype'. The presence of a considerable share of the factor 'environment' on the manifestation of main quantitative traits was reported by Tilahun et al. (2015) and Kanouni et al. (2015) in chickpeas, and by Sayar and Han (2016) in peas. Mortazavian et al. (2014) expressed an opinion that the stability analysis of the traits acquires important meaning when the effect of GE interaction is considerable.

The results of the correlation analysis of the data in our study were in support of the conclusions reached by other authors (Paula et al., 2014; Bornhofen et al., 2017) regarding the presence of positive relationships between the stability parameters and some main quantitative characteristics. According to Annicchiarico (2002), one of the possible explanations of this fact is that these methods (to determine the parameters mentioned) are related to the concept of static stability, which is obtained independently of the average expression of a given trait. The choice (made by these methods) may lead to the selection of stable but not necessarily productive genotypes.

Cargnelutti et al. (2009) concluded that the significant positive relationships between the parameters of the different methods showed that, in terms of stability, genotypes were evaluated in a similar way. The authors stated that the methods can provide much information and that only one method for stability evaluation would be sufficient to select the best genotypes. However, even in a strong relationship between the methods, the genotype position could be different in the different methods.

According to Balashova et al. (2013), one of the main tasks in the breeding of leguminous crops, including broad bean, was the seed yield stabilization. The authors considered that this problem can be solved by changing the idiotype of plants. According to them, one of the reasons for reducing the grain yield was the excessive increase in height and the lodging of plants. This was due to the fact that cultivars continue their vegetative growth after the formation of reproductive organs, especially in humid climatic conditions. For this reason, breeders' efforts should be directed to changing the plant habitus and development of highly productive determinant forms.

Conclusions

The results of the variance analysis in 17 broad bean accessions showed a significant genotype × environment interaction for all quantitative traits (excluding pod width). The factor environment had the greatest impact on the phenotypic manifestation of the traits, followed by the factors genotype and genotype × environment interaction.

In terms of plant height and 1st pod height, accessions FbH 16 and FbH 13 can be determined as high (79 cm, 35 cm) and ecologically stable (bi = 0.76, bi = 0.79). BGE 029055 was low variable, with high values of the number of pods (15) and seeds (41) per plant.

Accessions FbH 14, FbH 16, FbH 15 and BGP were distinguished by high seed weight per plant (from 28.36 to 34.93 g), but they exhibited instability (bi > 1) under unfavorable environmental conditions. In contrast, Fb 1903, BGE 043776 and Fb 3270 were very stable (bi < 1) but low-productive. Intermediate positions occupied accessions Fb 1896, Fb 1929, Fb 2481, Fb 2486, BGE 002106 and BGE 029055, which had the coefficient of linear regression close to 1, but they were also low-productive. Interesting from a breeding view point was BGE 041470, which is characterized by high values of 100 seeds mass (101.38 g) and seed weight per plant (32.14 g), as well as with relative stability (bi = 1.10).

GGE biplot analysis determined accessions BGE 046721, BGE 032012, FbH 15 and FbH 16 as a promising breeding material combining high and stable seed yield.

References

Annicchiarico P. Cultivar adaptation and recomendation from alfafa trials in Northern Italy. J. Genet. Plant Breed. 1992;4:269-278.

Annicchiarico P. Genotype × environment interaction: challenges and opportunities for plant breeding and cultivar recommendation. FAO Plant Production and Protection Paper. Rome, 2002.

Balashova I.T., Pronina E.P., Sirota S.M., Gordeev D.K. Bean culture in Chernozem zone of Russia. Ovoschi Rossii = Vegetable Crops of Russia. 2013;1:60-62. (in Russian)

Barov V. Analysis and schemas of the field trial. Sofia, 1982. (in Bulgarian)

Bornhofen E., Benin G., Storck L., Woyann L.G., Duarte T., Stoco M.G. Statistical methods to study adaptability and stability of wheat genotypes. Bragantia. 2017;76(1):1-10.

Buravtseva T.V., Lagutina L.V., Gurkina M.V. Assessment of new beans source material from VIR collection and allocation of sources of economically important features. In: The Role of Genetic Resources and Breeding Achievements in Ensuring the Dynamic Development of Agricultural Production: Proc. of N. Int. Sci. Conf. (July 8–9, 2009, Orel). Orel, 2009;219-233.

Cargnelutti Filho A., Storck L., Riboldi J., Guadagnin J.P. Associação entre métodos de adaptabilidade e estabilidadeem em milho. Ciência Rural. 2009;39(2):340-347.

Cruz C.D. Programa Genes: Biometria. version 7.0. Viçosa, Brazil: Univ. of Federal Viçosa, 2009.

Eberhart S.A., Russel W.A. Stability parameters for comparing varieties. Crop Sci. 1966;6:36-40.

Farshadfar E., Zali H., Mohammadi R. Evaluation of phenotypic stability in chickpea genotypes using GGE-Biplot. Ann. Biol. Res. 2011; 2:282-292.

Georgieva N. Suitability of pea cultivars for organic farming conditions. Biol. Agric. Hortic. 2017;33(4):225-234.

Hamayoon R., Khan H., Naz S.L., Munir I., Arif M., Khalil I., Khan A. Performance of chickpea genotypes under two different environmental conditions. Afr. J. Biotechnol. 2011;10:1534-1544.

Huehn M. Non-parametric measures of phenotypic stability: Part 1. Theory. Euphytica. 1990a;47:189-194.

Huehn M. Non-parametric measures of phenotypic stability: Part 2. Application. Euphytica. 1990b;47:195-201.

Ilchovska M. Ecological assessment of common and special maize hybrids. In: Proc. of Int. Sci. Conf. of the Union of Scientists. Stara Zagora, 2008;56-62.

Imtiaz M., Malhotra R.S., Singh M., Arslan S. Identifying high yielding, stable chickpea genotypes for spring sowing: specific adaptation

to location and sowing seasons in the Mediterranean region. Crop Sci. 2013;53:1472-1480.

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;7:219-230.

Kazydub N.G., Shamanin V.P. Course of Lectures on Selection and Genetics of Legumes (peas, soybean, common beans, vetch, legumes). Omsk: OmGAU Publ., 2003. (in Russian)

Kendal E., Sayar M.S. The stability of some spring triticale genotypes using biplot analysis. J. Anim. Plant Sci. 2016;26:754-765.

Kendal E., Sayar M.S., Tekdal S., Aktaş H., Karaman M. Assessment of the impact of ecological factors on yield and quality parameters in triticale using GGE biplot and AMMI analysis. Pak. J. Bot. 2016; 48:1903-1913.

Konvalina P., Stehno Z., Moudry J. The critical point of conventionally bread soft wheat varieties in organic farming systems. Agron. Res. 2009;7(2):801-810.

Kurkina Yu. Biological dependence among legumes and their selection value: PhDissertation. Voronezh, 2003. (in Russian)

Kurkina Yu. Breeding of broad bean for seed yield. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2013;16(4/2):922-925. (in Russian)

Link W., Balko C., Stoddard F.L. Winter hardiness in faba bean: physiology and breeding. Field Crops Res. 2010;115:297-307.

Marakaeva T.V., Kazydub N.G. Assessment of ecological plasticity and stability of samples of haricot of Western Siberia vegetable in the conditions of the southern forest-steppe. Int. Res. J. 2016;48(5):181-184. (in Russian)

Mortazavian S.M.M., Nikkhah H.R., Hassani F.A., Sharif-Al-Hosseini M., Taheri M., Mahlooji M. GGE biplot and AMMI analysis of yield performance of barley genotypes across different environments in Iran. J. Agr. Sci. Tech. 2014;16:609-622.

Paula T.M., Marinho C.D., Souza V., Barbosa M.H.P., Peternelli L.A., Kimbeng C.A. Relationships between methods of variety adaptability and stability in sugarcane. Genet. Mol. Res. 2014;13(2):4216-4225.

Plaisted R.L., Peterson L.C. A technique for evaluating the ability of selection to yield consistently in different location and seasons. Am. Potato J. 1979;36:381-385.

Pouresmael M., Kanouni H., Hajihasani M., Astraki H., Mirakhorli A., Nasrollahi M., Mozaffari J. Stability of chickpea (*Cicer arietinum* L.) landraces in National Plant Gene Bank of Iran for drylands. J. Agr. Sci. Tech. 2018;20:387-400.

Rubiales D. Faba beans in sustainable agriculture. Field Crops Res. 2010;115:201-202.

Sayar M.S. Additive main effects and multiplicative interactions (AMMI) analysis for fresh forage yield in common vetch (*Vicia sativa* L.) genotypes. Agric. For. 2017;63:119-127.

Sayar M.S., Han Y. Forage yield performance of forage pea (*Pisum sativum* spp. *arvense* L.) genotypes and assessments using GGE biplot analysis. J. Agr. Sci. Tech. 2016;18:1621-1634.

Shukla G.K. Some aspects of partitioning genotype-environmental components of variability. Heredity. 1972;28:237-245.

Tai G.C.C. Analysis of genotype – environment interactions of potato yield. Crop Sci. 1979;19:434-438.

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;14:1564-1575.

Wricke G. Zur berechnung der ökovalenz bei sommerweizen und hafer. Pflanzenzuchtung. 1965;52:127-138.

Yan W. Singular-value partitioning in biplot analysis of multienvironment trial data. Agron. J. 2002;94:990-996.

Yan W., Tinker N.A. Biplot analysis of multi-environment trial data: principles and applications. Can. J. Plant Sci. 2006;86:623-645.

Zhuchenko A.A. Mobilization of Genetic Resources of Flowering Plants on the Basis of their Identification Systematization. Moscow, 2012. (in Russian)

ORCID ID
N. Georgieva orcid.org/0000-0002-6364-1310
V.I. Kosev orcid.org/0000-0003-3292-4200

Acknowledgements. The authors thank the Agricultural Academy in Bulgaria for funding the study.

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

Received March 20, 2019. Revised July 1, 2019. Accepted July 8, 2019. Published online August 30, 2019.

ГЕНЕТИКА И СЕЛЕКЦИЯ ЖИВОТНЫХ
Оригинальное исследование / Original article
УДК 636.52:575.174.015.3

Вавиловский журнал генетики и селекции. 2019;23(8):993-998
DOI 10.18699/VJ19.575

Эффективность использования SNP-маркеров в гене *MSTN* в селекции кур пушкинской породы

Н.В. Дементьева, А.Б. Вахрамеев, Т.А. Ларкина, О.В. Митрофанова ✉

Всероссийский научно-исследовательский институт генетики и разведения сельскохозяйственных животных – филиал Федерального научного центра животноводства – ВИЖ им. академика Л.К. Эрнста, Пушкин, Санкт-Петербург, Россия
✉ e-mail: mo1969@mail.ru

В птицеводстве показатели, отражающие интенсивность роста молодняка и экстерьерные характеристики кур, – важные ориентиры для селекции. Традиционный селекционный отбор, основанный на фенотипической оценке, отличается невысокой эффективностью при низком коэффициенте наследуемости признака, и в малочисленных группах животных и птиц, разводимых в биоресурсных коллекциях, его сложно применять. Использование молекулярно-генетических маркеров, связанных с экономически значимыми признаками, позволяет проводить ранний отбор птицы. Это влечет за собой повышение рентабельности птицеводства. В последнее время удобными для селекционных целей маркерами служат однонуклеотидные полиморфизмы (SNP). В опытной популяции кур пушкинской породы в течение пяти поколений (P1–P5) проводили маркерную селекцию на увеличение живой массы. В ее основе был отбор по однонуклеотидной замене rs313744840 в гене *MSTN*. В результате отбора за этот период произошло значительное увеличение частоты аллеля A в этом гене: с 0.11 до 0.50. Ассоциация SNP-маркера с мясными качествами в опытной группе кур привела к изменениям экстерьерного профиля взрослой птицы в возрасте 330 дней. Наибольшей живой массой и длинным корпусом отличались особи с генотипами AG и AA. В результате селекции птица в среднем стала крупнее за счет увеличения количества гетерозиготных особей с длинным корпусом и большим обхватом груди. Глубина груди и ширина таза увеличились параллельно с ростом частоты аллеля A в опытной популяции. Обнаружена тенденция к увеличению этих показателей при замене G на A в генотипе. Насыщение популяции желательными аллелями привело к росту средних показателей живой массы кур. Анализ экстерьерных параметров взрослой птицы показал, что этот рост достигается за счет увеличения глубины и объема корпуса птицы, а не роста длины конечностей. Таким образом, в опытной популяции кур пушкинской породы на протяжении пяти поколений проведена маркерная селекция на увеличение живой массы, которая достоверно ($p < 0.001$) изменила экстерьерный профиль взрослой птицы.
Ключевые слова: ПЦР-ПДРФ; полиморфизм; миостатин; куры; маркерная селекция; экстерьер.

Для цитирования: Дементьева Н.В., Вахрамеев А.Б., Ларкина Т.А., Митрофанова О.В. Эффективность использования SNP-маркеров в гене *MSTN* в селекции кур пушкинской породы. Вавиловский журнал генетики и селекции. 2019;23(8):993-998. DOI 10.18699/VJ19.575

Efficiency of using SNP markers in the *MSTN* gene in the selection of the Pushkin breed chickens

N.V. Dementeva, A.B. Vakhrameev, T.A. Larkina, O.V. Mitrofanova ✉

Russian Research Institute of Farm Animal Genetics and Breeding – Branch of the L.K. Ernst Federal Science Center for Animal Husbandry, Pushkin, St. Petersburg, Russia
✉ e-mail: mo1969@mail.ru

In the poultry industry, indicators reflecting the growth rate of young stock and the exterior characteristics of chickens are important benchmarks for breeding. Traditional selection based on phenotypic evaluation is characterized by low efficiency with a low character inheritance ratio and is difficult to apply in small groups of animals and birds bred in bioresource collections. The use of molecular genetic markers associated with economically important traits makes it possible to carry out early selection of birds. This entails an increase in the profitability of the poultry industry. Recently, single nucleotide polymorphisms (SNPs) have served as convenient markers for selection purposes. For five generations (P1–P5), an experimental selection of hens of the Pushkin breed was carried out for live weight. It was based on selection for single nucleotide polymorphism rs313744840 in the *MSTN* gene. As a result, a significant increase in the frequency of allele A in this gene, from 0.11 to 0.50, took place. The association of SNP markers with meat qualities in the experimental group led to changes in the exterior profile of an adult bird at 330 days of age. The individuals with the AA and AG genotypes had the greatest live weight and longest body. As a result of selection, the bird on average became larger due to an increase in the number of heterozygous individuals with long bodies and large chest girths. The depth of the chest and the width of the pelvis increased due to an increase in the frequency of allele A in the experimental population.

© Дементьева Н.В., Вахрамеев А.Б., Ларкина Т.А., Митрофанова О.В., 2019

A tendency towards an increase in these indicators with the substitution of G with A in the genotype was found. Saturation of the population with desirable alleles led to an increase in the average live weight of the chickens. Analysis of the exterior parameters of adult birds showed that this growth is achieved by increasing the depth and volume of the bird body, and not by increasing the length of the limbs. Thus, marker selection carried out for five generations in the experimental population of Pushkin breed chickens to increase body weight has reliably ($p < 0.001$) changed the exterior profile of adult birds.
Key words: PCR-RFLP; polymorphism; myostatin; chickens; marker selection; exterior.

For citation: Dementeva N.V., Vakhrameev A.B., Larkina T.A., Mitrofanova O.V. Efficiency of using SNP markers in the *MSTN* gene in the selection of the Pushkin breed chickens. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):993-998. DOI 10.18699/VJ19.575 (in Russian)

Введение

Интенсивность роста молодняка и экстерьерные характеристики птицы – важные показатели, влияющие на продуктивность. Применение молекулярно-генетических маркеров, связанных с экономически значимыми признаками, позволяет проводить ранний отбор птицы, что обуславливает повышение рентабельности птицеводства. В последнее время широко используются в качестве маркеров однонуклеотидные полиморфизмы (single nucleotide polymorphism, SNP). В геноме курицы выявлено около 13.93 млн таких полиморфизмов (Boschiero et al., 2018). Поскольку генотипирование SNP несложно и поиск ориентирован только на интересующие генетические последовательности, это уменьшает время и стоимость оценки (Smaragdov, 2009).

В настоящее время в мире проводится множество исследований, направленных на поиск геномных ассоциаций между отдельными SNP и селекционными признаками у кур (Fornari et al., 2014; Cruz et al., 2015; Grupioni et al., 2017). Найденные связи помогают в картировании полиморфизмов, лежащих в основе сложных признаков у сельскохозяйственной птицы. Современные технологии, основанные на работе с SNP-чипами, дают возможность выявлять полногеномные ассоциации для показателей интенсивности роста у кур (Zhang et al., 2015). Но и селекция по отдельным маркерам продолжает оставаться удобным инструментом достижения генетического прогресса в популяциях сельскохозяйственных животных и птиц. Особенно она актуальна при разведении малочисленных и генофондных популяций кур.

Традиционный селекционный отбор основан на фенотипической оценке, отличается невысокой эффективностью и сложностью. Его не всегда возможно использовать для количественных признаков, а также для малочисленных групп животных и птиц, какими обычно являются популяции, разводимые в рамках биоресурсных коллекций. Применение методов отбора по маркерам (MAS-селекция – marker-assisted selection) позволяет селекционерам выбирать сельскохозяйственных животных с высокими показателями породного соответствия и воспроизводства в раннем возрасте и более эффективно (Khlestkina, 2014). Процесс отбора, фокусирующийся на связи интересующего признака и генетических маркеров, приводит к накоплению в популяции индивидуумов с желательными комбинациями генов (Collard et al., 2005).

При использовании данного метода на птице с малой численностью популяций необходим контроль генетического разнообразия. Подобные работы встречаются на аборигенных породах кур (Li et al., 2008; Guo et al., 2016). До сих пор генетические исследования экстерьерных особенностей кур были направлены на качественные показатели внешнего вида птицы, такие как окраска и форма (Sun et al., 2015; Guo et al., 2016). В то же время живая масса и экстерьерные характеристики – важные показатели отбора в птицеводстве. Следовательно, нельзя упускать из внимания и поиск связей в геноме, ответственных за рост и накопление мышечной массы.

Одним из таких генов, непосредственно влияющих на массу тела различных животных и человека, на протяжении последних двадцати лет является ген миостатина *MSTN* (Zhang et al., 2012). Миостатин – белок, подавляющий в организме рост и дифференцировку мышечной ткани, играет роль негативного регулятора массы скелетных мышц. Этот белок секретируется мышечными клетками и действует по принципу обратной связи. При возрастании мышечной массы увеличивается секреция миостатина, что тормозит дальнейший рост мышц.

Природные мутации, приводящие к снижению количества миостатина и (или) подавляющие его функции, были выявлены у различных видов животных и птиц, в том числе кур. У них найдено большое количество SNP в этом гене и показано их влияние на скорость роста, репродуктивные показатели и качество мяса (Baron et al., 2002; Dementeva et al., 2017). Кроме того, у различных пород овец, свиней, собак и кур определены мутации в некодирующих регуляторных областях, что влияет на уровень экспрессии *MSTN*, а, следовательно, на рост и объем мышечной массы (Hu et al., 2013). Проведено изучение влияния SNP в этом гене на живую массу в различных условиях выращивания и смертность цыплят (Ye et al., 2007).

Целью нашего исследования было определение влияния проводимой MAS-селекции по SNP rs313744840 в гене миостатина на экстерьерные характеристики кур пушкинской породы из биоресурсной коллекции Всероссийского научно-исследовательского института генетики и разведения сельскохозяйственных животных (ВНИИГРЖ). В задачи работы входило: проследить динамику частоты аллелей rs313744840 в гене миостатина в популяции пушкинской породы кур в течение ряда лет, проанализировать ассоциации замены rs313744840 в гене миостатина с экстерьерными показателями птицы.

Материалы и методы

Материалом для исследования послужила ДНК, выделенная из крови кур и петухов экспериментальной популяции пушкинской породы, которую в течение пяти поколений (P1–P5) отбирали по замене rs313744840 в гене *MSTN* и живой массе. Популяция была сформирована на

базе биоресурсной коллекции ВНИИГРЖ «Генетическая коллекция редких и исчезающих пород кур» (Пушкин, Санкт-Петербург).

При рождении цыплят индивидуально помечали. В ходе выращивания проводили их регулярное взвешивание в возрасте 7, 33, 80 и 90 дней. Одновременно вели выбраковку больных или травмированных цыплят. Птицу с желательным генотипом по замене в гене миостатина отбирали в племенное ядро, представители которого в возрасте 330 дней были взвешены и оценены по экстерьерным показателям. Кроме того, в поколениях P2–P4 не применяли отбор кур по массе. В поколении P5 была проведена жесткая выбраковка кур, и в племенное ядро были взяты особи с повышенной интенсивностью роста, высокой живой массой и хорошей яйценоскостью матерей, даже если генотип по замене rs313744840 в гене *MSTN* отличался от желательного. Измеряли следующие параметры: длину корпуса, обхват груди, обхват плюсны, длину киля, длину плюсны, длину бедра, глубину груди, ширину таза, ширину груди в ключицах, угол груди. В качестве контроля была взята популяция кур пушкинской породы из этой же биоресурсной коллекции, которую содержали в аналогичных условиях при групповом разведении.

Кровь для выделения ДНК отбирали у птиц в возрасте 33 дней из вены крыла в микропробирку, содержащую в качестве антикоагулянта 50 мкл 0.5 мМ ЭДТА pH 8.0 (ЭДТА – динатриевая соль этилендиаминтетрауксусной кислоты). До использования образцы хранили при температуре –20 °C. Геномную ДНК выделяли стандартным фенольно-детергентным способом.

Генотипирование одиночного нуклеотидного полиморфизма проводили с помощью метода ПЦР-ПДРФ (метод полимеразной цепной реакции с последующим анализом ресткриктных фрагментов). Инструменты для выявления SNP показаны в табл. 1. Полимеразную цепную реакцию выполняли по методике, описанной ранее (Mitrofanova et al., 2017).

Для электрофореза использовали 1.5 % агарозные гели, содержащие флуоресцентный краситель бромистый этидий и TBE-буфер (45 мМ трис-борат, 1 мМ ЭДТА). Смесь после рестрикции вносили в кармашки геля. Электрофорез проводили в течение 1 ч при рабочем напряжении 150 В. Сигнал флуоресценции фотографировали в системе гель-документации (Kodak).

На основании генотипирования проводили однофакторный дисперсионный анализ для выявления различий по продуктивным признакам в пакете программ SigmaPlot (версия 12.0.5) и использования критериев Шапиро–Уилка и *t*-критерия Стьюдента.

Различия по экстерьерным показателям между генотипами кур анализировали с применением коррекции дисперсионного анализа на множественное сравнение (All Pairwise Multiple Comparison Procedures).

Результаты

В течение пяти поколений (2012–2016 гг.) в опытной популяции пушкинской породы кур проводили отбор на увеличение частоты аллеля A по замене rs313744840 в гене *MSTN*. В результате отбора за этот период произошло значительное увеличение частоты аллеля A: с 0.11 до 0.50 (табл. 2).

Поскольку критическое значение χ^2 для составляет 3.84, то нарушение равновесия Харди–Вайнберга наблюдалось в поколении P3 из-за избытка гомозигот GG, а в поколении P5 – из-за избытка гетерозигот.

В рамках изучения ассоциации SNP-маркера с мясными качествами у птиц в опытной группе была проведена экстерьерная оценка взрослой птицы в возрасте 330 дней. По результатам оценки поколения P5 проводимая селекция достоверно ($p < 0.001$) увеличила длину корпуса, обхват и глубину груди, длину киля, ширину груди, дли-

Таблица 1. Инструменты для выявления одиночного нуклеотидного полиморфизма

Номер замены	Замена	Температура отжига, °C	Праймеры (3′–5′)	Эндонуклеаза	Размер фрагментов, п. н.
rs313744840	A/G	60	aaccaatcgtcggttttgac cgttctctgtgggctgacta	*MspI*	297, 260, 37

Таблица 2. Динамика изменения частот генотипов и аллелей по замене rs313744840 в гене *MSTN* при направленном отборе у кур пушкинской породы биоресурсной коллекции

Поколение кур	Показатель	Генотип			Аллель		χ^2
		AA	AG	GG	A	G	
P2, 2013 г.	Поголовье	4	22	113			
	Частота	0.03	0.16	0.81	0.11	0.89	2.41
P3, 2014 г.	Поголовье	32	71	143			
	Частота	0.13	0.29	0.58	0.28	0.72	18.87
P4, 2015 г.	Поголовье	64	135	69			
	Частота	0.24	0.50	0.26	0.49	0.51	0.02
P5, 2016 г.	Поголовье	87	243	90			
	Частота	0.21	0.58	0.21	0.5	0.5	10.34

Таблица 3. Экстерьерная оценка опытного поголовья кур пушкинской породы (популяций P4 и P5) в сравнении с контрольной популяцией

Показатель	Популяция			Уровень достоверности
	P4 (n = 98)	P5 (n = 97)	Контрольная (n = 20)	
Живая масса в 330 дней, г	2430 ± 29[1a]	2707 ± 34[1b]	2500 ± 80[1c]	1a-1b; 1b-1c – $p < 0.05$
Длина корпуса, см	18.71 ± 0.09[2a]	19.49 ± 0.07[2b]	18.23 ± 0.20[2c]	2a-2b; 2b-2c; 2a-2c – $p < 0.001$
Обхват груди, см	34.61 ± 0.52[3a]	35.24 ± 0.15[3b]	31.55 ± 0.34[3c]	3a-3c; 3b-3c – $p < 0.001$
Обхват плюсны, см	4.02 ± 0.03	4.08 ± 0.02	3.85 ± 0.6	нд
Длина плюсны, см	10.33 ± 0.09[4a]	10.64 ± 0.07[4b]	10.15 ± 0.10[4c]	4a-4b; 4b-4c – $p < 0.05$
Длина бедра, см	10.66 ± 0.80	10.10 ± 0.07	9.77 ± 0.34	нд
Глубина груди, см	11.39 ± 0.08[5a]	12.29 ± 0.08[5b]	11.38 ± 0.12[5c]	5a-5b; 5b-5c – $p < 0.001$
Ширина таза, см	8.72 ± 0.06[6a]	9.36 ± 0.06[6b]	8.66 ± 0.25[6c]	6a-6b; 6b-6c – $p < 0.001$
Ширина груди в ключицах, см	7.07 ± 0.06[7a]	7.56 ± 0.06[7b]	6.72 ± 0.08[7c]	7a-7b; 7b-7c – $p < 0.001$
Угол груди, град.	77.04 ± 0.34[8a]	80.69 ± 0.47[8b]	81.21 ± 1.46[8c]	8a-8b; 8a-8c – $p < 0.001$
Длина киля, см	11.21 ± 0.07	11.85 ± 0.07[9b]	10.95 ± 0.16[9c]	9b-9c – $p < 0.001$

Примечание. нд – достоверных различий между группами не обнаружено.

Таблица 4. Экстерьерная оценка опытной популяции кур P5 различных генотипов по rs313744840 в гене *MSTN*

Показатель	Генотип			Уровень достоверности
	AA (n = 34)	AG (n = 48)	GG (n = 9)	
Живая масса в 330 дней, г	2599 ± 44[1a]	2731 ± 59[1b]	2436 ± 117[1c]	1a-1b; 1b-1c – $p < 0.05$
Длина корпуса, см	19.10 ± 0.11[2a]	19.58 ± 0.10[2b]	19.12 ± 0.16[2c]	2a-2b; 2b-2c – $p < 0.05$
Длина корпуса лентой, см	21.65 ± 0.13[3a]	22.34 ± 0.14[3b]	21.56 ± 0.28[3c]	3a-3b; 3b-3c – $p < 0.05$
Косая длина туловища, см	22.01 ± 0.15	22.62 ± 0.18[4b]	21.64 ± 0.43[4c]	4a-4c – $p < 0.05$
Обхват груди, см	34.71 ± 0.22	35.25 ± 0.26	34.28 ± 0.45	нд
Ширина груди в ключицах, см	7.53 ± 0.09	7.43 ± 0.08	7.01 ± 0.26	нд
Глубина груди, см	12.42 ± 0.17	12.10 ± 0.12	12.03 ± 0.20	нд
Обхват плюсны, см	3.97 ± 0.04	4.02 ± 0.02	3.97 ± 0.6	нд
Длина киля, см	11.66 ± 0.09	11.72 ± 0.11	11.83 ± 0.25	нд
Длина плюсны, см	10.68 ± 0.11	10.56 ± 0.11	10.53 ± 0.31	нд
Длина бедра, см	9.97 ± 0.13	10.06 ± 0.10	9.83 ± 0.19	нд
Длина голени, см	14.38 ± 0.12	14.68 ± 0.10	14.67 ± 0.19	нд
Обхват голени, см	11.89 ± 0.20	12.03 ± 0.13	11.61 ± 0.33	нд
Ширина таза, см	9.34 ± 0.10	9.25 ± 0.10	9.09 ± 0.18	нд
Угол груди, град.	81.06 ± 0.67	80.62 ± 0.69	79.33 ± 1.9	нд

Примечание. нд – достоверных различий между группами не обнаружено.

ну плюсны и ширину таза у кур опытной популяции по сравнению с курами из контрольной группы (табл. 3). Угол груди у кур поколения P5 увеличился по сравнению с курами предыдущего поколения P4, но не отличался от угла груди кур из контрольной популяции.

Анализ различий между генотипами по замене rs313744840 в гене *MSTN* у кур в поколении P5 опытной популяции приведен в табл. 4. Гетерозиготы AG отличались от остальных особей по живой массе и в среднем весили 2731 ± 59 г. Для них показано достоверное увеличение размеров корпуса по трем показателям: длине корпуса циркулем, длине корпуса лентой, косой длине туловища. По признакам глубины груди, ширины груди в ключицах, ширины таза отмечена тенденция к увеличению этих показателей при замене G на A в генотипе.

Обсуждение

В нашей работе ген миостатина у кур был выбран для использования в качестве критерия для отбора. Интерес к его изучению не угасает. Выявляется множество различных аллельных вариантов, определяются взаимосвязи однонуклеотидных замен в этом гене с продуктивными по-

казателями у различных животных (Ye et al., 2007; He et al., 2013; Hope et al., 2013; Tu et al., 2013; Dementeva et al., 2017; Mitrofanova et al., 2017; Rooney et al., 2017). Имеются исследования, подтверждающие ассоциативные связи SNP в районе гена миостатина со скоростью роста молодняка и массой тушки (Dushyanth et al., 2016). Изучение результатов селекции кур по маркерному аллелю гена миостатина у непромышленных пород ранее не проводили. В начале нашей работы такими маркерами выступали несколько SNP в гене миостатина (Дементьева и др., 2018; Митрофанова, Дементьева, 2018), но лишь замена в положении rs313744840 этого гена оказалась эффективной для отбора на увеличение живой массы у птицы. Селекция по маркеру у группы птиц пушкинской породы осуществлялась в течение пяти лет. В результате были достигнуты увеличение живой массы опытного поголовья и сдвиг частоты встречаемости аллеля A с 0.11 до 0.50.

Нарушение равновесия Харди–Вайнберга отмечено в поколении P3 из-за непропорционального отбора гомозигот AA по сравнению с поколением P2, а в поколении P5 – из-за избытка гетерозигот, взятых в племенное ядро по продуктивности. В процессе селекции главными критериями были не только генотип особи, но и воспроизводительные качества матерей (яйценоскость, масса яйца). Возможно, увеличение ширины таза (см. табл. 3) стало следствием отбора на увеличение массы яйца.

Ассоциация SNP-маркера с ростом и развитием в опытной группе привела к изменениям экстерьерного профиля взрослой птицы в возрасте 330 дней. Наибольшей живой массой и длинным корпусом отличались особи с гетерозиготным генотипом и генотипом AA. Для них отмечено достоверное увеличение размеров корпуса по трем показателям: длине корпуса, длине корпуса лентой, косой длине туловища. В результате селекции птица в среднем стала крупнее за счет увеличения количества гетерозиготных особей в популяции с длинным корпусом и большим обхватом груди. Глубина груди и ширина таза возросли благодаря росту частоты аллеля A в опытной популяции птиц по сравнению с курами из контрольной группы. Обнаружена тенденция к увеличению этих показателей при замене G на A в генотипе. Поскольку отбор осуществлялся в течение нескольких лет, это привело к повышению средних показателей живой массы кур. Анализ экстерьерных параметров взрослой птицы показал, что это достигается вследствие увеличения глубины и объема корпуса птицы, а не за счет роста длины конечностей.

Заключение

Таким образом, в опытной популяции пушкинской породы кур на протяжении пяти поколений проведена маркерная селекция на увеличение живой массы, которая достоверно ($p < 0.001$) изменила экстерьерный профиль взрослой птицы (330 дней). Увеличились длина корпуса, обхват и глубина груди, длина киля, ширина груди и ширина таза.

Список литературы / References

Дементьева Н.В., Митрофанова О.В., Кудинов А.А. Эффект различных рационов кормления на результат ассоциативного анализа полиморфизма в гене MSTN и роста живой массы у молодняка кур. Междунар. журн. прикл. и фундамент. исследований. 2018;6:145-148. DOI 10.23670/IRJ.2018.72.6.017.
[Dementeva N.V., Mitrofanova O.V., Kudinov A.A. The effect of different diets of feeding on the result of associative analysis of polymorphism in the MSTN gene and growth of live weight in chickens. Mezhdunarodnyy Zhurnal Prikladnykh i Fundamentalnykh Issledovaniy = International Journal of Applied and Fundamental Research. 2018;6:145-148. DOI 10.23670/IRJ.2018.72.6.017. (in Russian)]

Митрофанова О.В., Дементьева Н.В. Влияние различных рационов кормления на ассоциацию полиморфизма в гене MSTN и роста живой массы у молодняка петушков пушкинской породы. Молочнохоз. вестник. 2018;3(31):62-70. DOI 10.24411/2225-4269-2018-00021.
[Mitrofanova O.V., Dementeva N.V. Influence of various diets on association of polymorphism in MSTN gene and live weight increase in young cockerels of pushkin breed. Molochnokhozyaystvennyy Vestnik = Dairy Farming Journal. 2018;3(31):62-70. DOI 10.24411/2225-4269-2018-00021. (in Russian)]

Baron E.E., Wenceslau A.A., Alvares L.E., Nones K., Ruy D.C., Schmidt G., Zanella E., Coutinho L., Ledur M.C. High level of polymorphism in the *Myostatin* chicken gene. In: Proc. of the 7th World Congress on Genetics Applied to Livestock Production, August 19–23, 2002. Montpellier, France. 2002;19-23.

Boschiero C., Moreira G., Gheyas A., Godoy T., Gasparin G., Mariani P., Paduan M., Cesar A., Ledur M., Coutinho L. Genome-wide characterization of genetic variants and putative regions under selection in meat and egg-type chicken lines. BMC Genomics. 2018;19(83):1-18. DOI 10.1186/s12864-018-4444-0.

Collard B., Jahufer M., Brouwer J., Pang E. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica. 2005; 142:169-196. DOI 10.1007/s10681-005-1681-5.

Cruz V., Schenkel F., Savegnago R., Grupioni N., Stafuzza N., Sargolzaei M., Ibelli A., Peixoto J., Ledur M., Munari D. Association of apolipoprotein B and adiponectin receptor 1 genes with carcass, bone integrity and performance traits in a paternal broiler line. PLoS One. 2015;10(8):e0136824. DOI 10.1371/journal.pone.0136824.

Dementeva N., Mitrofanova O., Tyshchenko V., Terletskiy V., Yakovlev A. The rate of weight gain and productivity of a chicken broiler cross with various polymorphic types of the myostatin gene. Russ. J. Genet.: Appl. Res. 2017;7(1):1-5. DOI 10.1134/S207905971701004X.

Dushyanth K., Bhattacharya T.K., Shukla R., Chatterjee R.N., Sitaramamma T., Paswan C., Guru Vishnu P. Gene expression and polymorphism of *Myostatin* gene and its association with growth traits in chicken. Anim. Biotechnol. 2016;27(4):269-277. DOI 10.1080/10495398.2016.1182541.

Fornari M., Zanella R., Ibelli A., Fernandes L., Cantao M., Thomaz-Soccol V., Ledur M., Peixoto J. Unraveling the associations of osteoprotegerin gene with production traits in a paternal broiler line. SpringerPlus. 2014;3(682). DOI 10.1186/2193-1801-3-682.

Grupioni N., Stafuzza N., Carvajal A., Ibelli A., Peixoto J., Ledur M., Munari D. Association of *RUNX2* and *TNFSF11* genes with production traits in a paternal broiler line. Genet. Mol. Res. 2017;16(1). DOI 10.4238/gmr16019443.

Guo X., Fang Q., Ma C., Zhou B., Wan Y., Jiang R. Whole-genome resequencing of Xishuangbanna fighting chicken to identify signatures of selection. Genet. Sel. Evol. 2016;48(62). DOI 10.1186/s12711-016-0239-4.

He Y.L., Wu Y.H., Quan F.S., Liu Y.G., Zhang Y. Comparative analysis of myostatin gene and promoter sequences of Qinchuan and Red Angus cattle. Genet. Mol. Res. 2013;12(3):3398-3406. DOI 10.4238/2013.September.4.6.

Hope M., Haynes F., Oddy H., Koohmaraie M., Al-Owaimer A., Geesink G. The effects of the myostatin g+6723G>A mutation on carcass and meat quality of lamb. Meat Sci. 2013;95(1):118-122. DOI 10.1016/j.meatsci.2013.03.029.

Hu W., Chen S., Zhang R., Liu Y. Single nucleotide polymorphisms in the upstream regulatory region alter the expression of myostatin. In Vitro Cell. Dev. Biol. Anim. 2013;49(6):417-423. DOI 10.1007/s11626-013-9621-5.

Khlestkina E.K. Molecular markers in genetic studies and breeding. Russ. J. Genet.: Appl. Res. 2014;4(3):236-244. DOI 10.1134/S2079059714030022.

Li H., Zhu W., Chen K., Song W., Shu J., Han W. Associations between GHR and IGF-1 gene polymorphisms, and reproductive traits in Wenchang chickens. Turk. J. Vet. Anim. Sci. 2008;32(4):281-285.

Mitrofanova O., Dementeva N., Krutikova A., Yurchenko O., Vakhrameev A., Terletskiy V. Association of polymorphic variants in *MSTN*, *PRL*, and *DRD2* genes with intensity of young animal growth in Pushkin breed chickens. Cytol. Genet. 2017;51(3);179-184. DOI 10.3103/S0095452717030082.

Rooney M.F., Porter R.K., Katz L.M., Hill E.W. Skeletal muscle mitochondrial bioenergetics and associations with myostatin genotypes in the Thoroughbred horse. PLOS One. 2017;12(11):e0186247. DOI 10.1371/journal.pone.0186247.

Smaragdov M.G. Genomic selection as a possible accelerator of traditional selection. Russ. J. Genet. 2009;45(6):633-636. DOI 10.1134/S1022795409060015.

Sun Y., Liu R., Zhao G., Zheng M., Sun Y., Yu X., Li P., Wen J. Genome-wide linkage analysis identifies loci for physical appearance traits in chickens. G3 (Bethesda). 2015;5(10):2037-2041. DOI 10.1534/g3.115.020883.

Tu P.-A., Lo L.-L., Chen Y.-C., Hsu C.-C., Shiau J.-W., Lin E.-C., Wang P.-H. Polymorphisms in the promoter region of *myostatin* gene are associated with carcass traits in pigs. J. Anim. Breed. Genet. 2013;131(2):116-122. DOI 10.1111/jbg.12053.

Ye X., Brown S., Nones K., Coutinho L., Dekkers J., Lamont S. Associations of myostatin gene polymorphisms with performance and mortality traits in broiler chickens. Genet. Sel. Evol. 2007;39(1):73-89. DOI 10.1051/gse:2006029.

Zhang G., Fan Q., Zhang T., Wang J., Wang W., Xue Q., Wang Y. Genome-wide association study of growth traits in the Jinghai Yellow chicken. Genet. Mol. Res. 2015;14(4):15331-15338.

Zhang G., Zhang L., Wei Y., Wang J., Ding F., Dai G., Xie K. Polymorphisms of the myostatin gene and its relationship with reproduction traits in the Bian chicken. Anim. Biotechnol. 2012;23(3):184-193. DOI 10.1080/10495398.2012.681411.

ORCID ID
N.V. Dementeva orcid.org/0000-0003-0210-9344
A.B. Vakhrameev orcid.org/0000-0001-5166-979X
T.A. Larkina orcid.org/0000-0002-7764-1338
O.V. Mitrofanova orcid.org/0000-0003-4702-2736

Благодарности. Исследование выполнено в рамках государственного задания АААА-А18-118021590138-1 с использованием популяций кур из биоресурсной коллекции ЦКП «Генетическая коллекция редких и исчезающих пород кур» (ВНИИГРЖ, Санкт-Петербург, Пушкин). Авторы выражают благодарность О.П. Юрченко, создателю пушкинской породы кур, за предоставленную возможность использовать эту породу в исследованиях.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 01.11.2018. После доработки 05.06.2019. Принята к публикации 10.06.2019.

ГЕНЕТИКА И СЕЛЕКЦИЯ ЖИВОТНЫХ
Оригинальное исследование / Original article
УДК 599.735.51:578.828.1

Вавиловский журнал генетики и селекции. 2019;23(8):999-1005
DOI 10.18699/VJ19.576

Ассоциация однонуклеотидного полиморфизма rs110861313 в межгенном районе хромосомы 23 с развитием лейкоза у крупного рогатого скота черно-пестрой породы

Р.Б. Айтназаров[1, 2], Е.В. Игнатьева[1, 2], Т.А. Агаркова[3], Н.Г. Двоеглазов[3], Н.А. Осипова[3], В.В. Храмцов[3], Н.С. Юдин[1, 2] ✉

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Сибирский федеральный научный центр агробиотехнологий Российской академии наук, р. п. Краснообск, Новосибирская область, Россия
✉ e-mail: yudin@bionet.nsc.ru

Недавно с помощью полногеномного анализа ассоциаций было идентифицировано несколько однонуклеотидных полиморфизмов (ОНП), связанных с чувствительностью крупного рогатого скота (КРС) к лейкозу. Однако все эти исследования были выполнены на животных голштинской породы и их гибридах. В связи с этим актуальна проверка информативности выявленных ранее генетических маркеров для российских пород КРС. Целью исследования была проверка ассоциации ОНП rs110861313 в межгенном участке хромосомы 23 с лейкозом у черно-пестрой породы КРС. При гематологическом исследовании крови животных, в сыворотке крови которых серологическими методами обнаружены специфические антитела к вирусу лейкоза КРС, были сформированы три группы: здоровые животные (n = 115), бессимптомные вирусоносители (n = 145) и больные лейкозом животные с персистентным лимфоцитозом (n = 107). Генотипирование ОНП rs110861313 проводили с использованием полимеразной цепной реакции с последующим анализом полиморфизма длин рестрикционных фрагментов. Выявлено достоверное снижение частоты генотипа A/A (11.2 %) у животных с персистентным лимфоцитозом по сравнению с вирусоносителями (27.6 %) ($p < 0.002$). Частота животных с генотипом C/C у животных с персистентным лимфоцитозом (41.1 %) была достоверно выше, чем у вирусоносителей (21.4 %) ($p < 0.001$). При этом бессимптомных вирусоносителей можно считать более адекватным контролем, чем здоровых, но не контактировавших с вирусом животных. По данным биоинформатического анализа, устойчивость к инфекции вирусом лейкоза КРС может быть связана с наличием в районе ОНП rs110861313 сайта связывания транскрипционного фактора FOXM1, который экспрессируется в клетках иммунной системы и потенциально может влиять на экспрессию ближайших генов (*LY6G5B*, *GPANK1*, *ABHD16A*, *LY6G6F*, *LY6G6E*, *CSNK2B*, *ApoM*). Таким образом, нами установлено, что ОНП rs110861313 в межгенном участке хромосомы 23 ассоциирован с развитием лейкоза при инфицировании вирусом лейкоза у животных черно-пестрой породы КРС.
Ключевые слова: крупный рогатый скот; черно-пестрая порода; лейкоз; вирусоноситель; персистентный лимфоцитоз; однонуклеотидный полиморфизм; rs110861313; ассоциация.

Для цитирования: Айтназаров Р.Б., Игнатьева Е.В., Агаркова Т.А., Двоеглазов Н.Г., Осипова Н.А., Храмцов В.В., Юдин Н.С. Ассоциация однонуклеотидного полиморфизма rs110861313 в межгенном районе хромосомы 23 с развитием лейкоза у крупного рогатого скота черно-пестрой породы. Вавиловский журнал генетики и селекции. 2019;23(8):999-1005. DOI 10.18699/VJ19.576

Single nucleotide polymorphism rs110861313 in the intergenic region of chromosome 23 is associated with the development of leukosis in the Russian Black Pied cattle

R.B. Aitnazarov[1, 2], E.V. Ignatieva[1, 2], T.A. Agarkova[3], N.G. Dvoeglazov[3], N.A. Osipova[3], V.V. Khramtsov[3], N.S. Yudin[1, 2] ✉

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Siberian Federal Research Center for AgroBioTechnologies, RAS, Krasnoobsk, Novosibirsk region, Russia
✉ e-mail: yudin@bionet.nsc.ru

In recent years, using a genome-wide association study (GWAS), a number of single nucleotide polymorphisms (SNPs) have been suggested to be associated with susceptibility to leukemia in cattle. However, all studies have been done with purebred Holstein cows and their hybrids. In this regard, it is important to confirm the functional role of polymorphisms previously identified in a GWAS study in Russian cattle breeds. The aim of this study was to verify the association between rs110861313 in the intergenic region of bovine chromosome 23 and leukemia in the Russian Black Pied cattle. Based on the levels of bovine leukemia virus (BLV)-specific antibodies detected in serum using serodiagnostic techniques, animals were divided into three groups: healthy animals (n = 115), asymptomatic virus carriers (n = 145) and animals with leukemia (n = 107). Genotyping of rs110861313 was carried out using polymerase chain reaction followed by analysis of restriction fragment length polymorphisms. A significant decrease in the frequency of the A/A geno-

© Айтназаров Р.Б., Игнатьева Е.В., Агаркова Т.А., Двоеглазов Н.Г., Осипова Н.А., Храмцов В.В., Юдин Н.С., 2019

type (11.2 %) was revealed in animals with persistent lymphocytosis compared to virus carriers (27.6 %) ($p < 0.002$). At the same time, the frequency of animals with the C/C genotype in animals with persistent lymphocytosis (41.1 %) was significantly higher than that of virus carriers (21.4 %) ($p < 0.001$). In this case, asymptomatic virus carriers can be considered a more suitable control than healthy animals that have not been in contact with the virus. According to bioinformatics analysis, resistance to BLV can be due to the presence of the transcription factor FOXM1 binding site in the region of rs110861313. FOXM1 is expressed in immune cells and can potentially affect the expression of the neighboring genes (*LY6G5B*, *GPANK1*, *ABHD16A*, *LY6G6F*, *LY6G6E*, *CSNK2B*, *ApoM*). Thus, we found that SNP rs110861313 in the intergenic region of bovine chromosome 23 is associated with the development of leukemia following BLV infection in the Russian Black Pied cattle.
Key words: cattle; Russian Black Pied cattle; bovine leucosis; virus carrier; persistent lymphocytosis; single nucleotide polymorphism; rs110861313; association.

For citation: Aitnazarov R.B., Ignatieva E.V., Agarkova T.A., Dvoeglazov N.G., Osipova N.A., Khramtsov V.V., Yudin N.S. Single nucleotide polymorphism rs110861313 in the intergenic region of chromosome 23 is associated with the development of leukosis in the Russian Black Pied cattle. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):999-1005. DOI 10.18699/VJ19.576 (in Russian)

Введение

Лейкоз крупного рогатого скота (КРС) – хроническая инфекционная болезнь, связанная со злокачественным разрастанием клеток кроветворных органов, в результате чего образуются отдельные опухолевые массы или диффузная инфильтрация различных тканей и органов. Возбудитель инфекции – онкогенный В-лимфоцитотропный вирус лейкоза крупного рогатого скота (ВЛКРС), относящийся к семейству ретровирусов (Gillet et al., 2007). Близким родственником ВЛКРС является Т-лимфотропный вирус человека типа 1 (T lymphotropic virus type 1, HTLV-I), вызывающий у человека Т-клеточный лейкоз и Т-клеточную лимфому. Большинство инфицированных животных – бессимптомные носители вируса. Приблизительно у одной трети инфицированных животных развивается доброкачественная поликлональная пролиферация В-клеток (персистентный лимфоцитоз), а у 5–10 % – летальные лимфоидные опухоли (Ferrer, 1980; Burny et al., 1988).

В России ежегодно лейкоз КРС наносит многомиллиардный ущерб молочному животноводству (Stepanova, 2016). Поскольку коэффициент наследуемости чувствительности к ВЛКРС составляет не менее 0.08, эффективной мерой профилактики этого заболевания может стать геномная селекция (Abdalla et al., 2013). Устойчивость к лейкозу связана с большим количеством генетических маркеров, многие из которых специфичны для отдельных пород КРС (Юдин и др., 2018б). В частности, недавно с помощью полногеномного анализа ассоциаций (GWAS) было идентифицировано несколько однонуклеотидных полиморфных маркеров (ОНП), связанных с чувствительностью КРС к инфекции ВЛКРС (Abdalla et al., 2016; Brym et al., 2016; Carignano et al., 2018). Но все эти исследования были выполнены на КРС голштинской породы и их гибридах. Ввиду этого для российских пород КРС актуальна проверка информативности генетических маркеров, ранее идентифицированных с помощью полногеномного анализа ассоциаций.

Целью исследования была проверка ассоциации ОНП rs110861313 в межгенном участке хромосомы 23 с лейкозом у крупного рогатого скота черно-пестрой породы. Ранее по результатам полногеномного анализа этот ОНП был ассоциирован с провирусной нагрузкой ВЛКРС и количеством лейкоцитов в крови у голштинов и гибридов голштинов с джерси в Аргентине (Carignano et al., 2018).

Материалы и методы

Животные. Исследовано 367 коров черно-пестрой породы в возрасте 1.5–8 лет из хозяйств Новосибирской области и Алтайского края. Для забора венозной крови брали вакуумные пробирки с антикоагулянтом (ЭДТА). Диагностические процедуры проводили по общепринятой методике. Инфицированных ВЛКРС животных выявляли с применением реакции иммунодиффузии в геле агара (РИД) с помощью набора для серологической диагностики лейкоза крупного рогатого скота (производства Курской биофабрики) и методом иммуноферментного анализа (ИФА), при использовании теста-набора IDEXX Leucose Serum Screening Ab Test. Идентификацию больных животных, а также дифференциальную диагностику форм и стадий заболевания осуществляли путем гематологического анализа крови животных, в сыворотке крови которых серологическими методами (РИД, ИФА) были обнаружены специфические антитела к ВЛКРС. Количество лейкоцитов подсчитывали посредством гематологического анализатора Mindray BC-288-Vet в соответствии с инструкцией по его эксплуатации. При обнаружении в пробе крови повышенного количества лейкоцитов и лимфоцитов животное считали больным лейкозом (Методические указания по диагностике лейкоза…, 2000). В результате животных разделили на три группы: здоровые ($n = 115$), бессимптомные вирусоносители ($n = 145$) и больные лейкозом ($n = 107$). Протокол экспериментов был одобрен комиссией по биоэтике ИЦиГ СО РАН.

Генотипирование однонуклеотидного полиморфизма rs110861313. Геномную ДНК выделяли из периферической крови стандартным методом протеолитической обработки с последующей экстракцией фенолом (Sambrook, Russell, 2006). Генотипирование аллельных вариантов ОНП rs110861313 в межгенном участке хромосомы 23 проводили с помощью полимеразной цепной реакции с последующим анализом полиморфизма длин рестрикционных фрагментов (ПЦР-ПДРФ). Дизайн праймеров и подбор условий ПЦР осуществляли с использованием программы Vector NT версии 11.5.1. Нуклеотидные последовательности генома коровы (сборка Bos_taurus_UMD_3.1.1/bosTau8) получали из геномного браузера UCSC (https://genome.ucsc.edu/). Эндонуклеазу рестрикции для генотипирования ОНП подбирали с помощью программы dCAPS (Neff et al., 2002).

Для проведения ПЦР применяли прямой 5′-TACAA AAACCCATCTAAAAGGGTGC-3′ и обратный 5′-AAG AAGCACCAGGACAAGCA-3′ (подчеркнута замена относительно референсной последовательности коровы) праймеры. Фрагмент 145 п. н. амплифицировали в 25 мкл реакционной смеси, содержащей 1x SE-буфер и 5 ед. активности Taq ДНК-полимеразы («СибЭнзим», Россия), 3 мМ $MgCl_2$, по 100 мкМ каждого dNTP, по 0.2 мкМ каждого праймера и 0.1–0.5 мкг ДНК. Режим амплификации: 35 циклов (95 °C – 30 с, 62 °C – 30 с, 72 °C – 30 с) с предварительной денатурацией при 95 °C – 3 мин. Продукт амплификации подвергали рестрикции ферментом AccB1 I («СибЭнзим», Россия) при условиях, рекомендуемых производителем. Ампликон размером 351 п. н. либо оставался неразрезанным (генотип AA), либо его разрезали на фрагменты 326 и 25 п. н. (генотип CC) или 351, 326 и 25 п. н. (генотип AC). Электрофоретический анализ продуктов рестрикции проводили в 4 % полиакриламидном геле с окрашиванием бромистым этидием.

Статистический и биоинформатический анализы. Проверку отклонения распределения генотипов от равновесия Харди–Вайнберга выполняли с использованием программы Hardy–Weinberg equilibrium calculator (http://www.oege.org/software/hwe-mr-calc.shtml). Сравнение частот генотипов и аллелей проводили с помощью критерия χ^2 для таблицы сопряженности 2 × 2. Критическое значение уровня статистической значимости при проверке нулевых гипотез принимали равным 0.05. Результаты обрабатывали с помощью пакета Statistica версии 8.0.

Для оценки возможного функционального эффекта нуклеотидной замены в районе межгенного ОНП rs110861313 на потенциальные сайты связывания транскрипционных факторов использовали ресурс PERFECTOS-APE (http://opera.autosome.ru/perfectosape). Анализ проводили с применением данных о позиционных весовых матрицах сайтов связывания транскрипционных факторов человека из базы HOCOMOCO (Kulakovskiy et al., 2018). Районы ДНК размером 25 п. н., фланкирующие ОНП rs110861313 с 5′- и 3′-концов, получали из сборки генома коровы Bos_taurus_UMD_3.1.1/bosTau8 с помощью геномного браузера UCSC. Пороговое значение *p*-value, оценивающее достоверность сходства участка ДНК с весовой матрицей сайта, принимали равным $5 \cdot 10^{-4}$. Считали, что ОНП достоверно нарушает сайт связывания транскрипционного фактора, если отношение значений *p*-value для референсного и альтернативного аллелей было не менее 4. Официальные символы генов, кодирующих транскрипционные факторы (ТФ), соответствующие обнаруженным таким образом сайтам, идентифицировали по базе EntrezGene (http://www.ncbi.nlm.nih.gov/gene). Поскольку информация по виду *Bos taurus* весьма ограниченная, в запросах к базам был выбран вид *Homo sapiens*. Данные о тканеспецифическом характере экспрессии генов получали из базы данных GeneCards (http://www.genecards.org/).

Результаты

Частоты генотипов и аллелей ОНП rs110861313 у больных лейкозом, здоровых животных и вирусоносителей приведены в табл. 1. Распределение частот генотипов во всех изученных выборках соответствовало равновесию Харди–Вайнберга. Обнаружено статистически значимое повышение частоты животных с генотипом A/A по ОНП rs110861313 у вирусоносителей (27.6 %) по сравнению со здоровыми (17.4 %) ($p = 0.053$), а также достоверное снижение частоты животных с генотипом C/C у вирусоносителей (21.4 %) по сравнению со здоровыми (37.4 %) ($p = 0.005$). Соответственно, у вирусоносителей была существенно повышена частота аллеля A (53.1 %) и снижена частота аллеля C (46.9 %) по сравнению со здоровыми животными (40.0 и 60.0 % соответственно) ($p < 0.01$).

Таблица 1. Частоты генотипов и аллелей по ОНП rs110861313 у больных лейкозом животных черно-пестрой породы крупного рогатого скота, а также в группах здоровых животных и бессимптомных вирусоносителей

Генотип или аллель	Частота генотипа (аллеля): % (число)[а]			p[в]
	Здоровые (1), (n = 115)[б]	Вирусоносители (2), (n = 145)[б]	Больные (3), (n = 107)[б]	
A/A	17.4 (20)	27.6 (40)[г]	11.2 (12)[д]	$\mathbf{p_{1,2} = 0.053}$ $p_{1,3} = 0.191$ $\mathbf{p_{2,3} = 0.002}$
A/C	45.2 (52)	51.0 (74)	47.7 (51)	$p_{1,2} = 0.351$ $p_{1,3} = 0.715$ $p_{2,3} = 0.597$
C/C	37.4 (43)	21.4 (31)[г]	41.1 (44)[д]	$\mathbf{p_{1,2} = 0.005}$ $p_{1,3} = 0.569$ $\mathbf{p_{2,3} = 0.001}$
A	40.0 (92)	53.1 (154)[г]	35.0 (75)[д]	$\mathbf{p_{1,2} = 0.003}$ $p_{1,3} = 0.282$ $\mathbf{p_{2,3} = 0.0001}$
C	60.0 (138)	46.9 (136)[г]	65.0 (139)[д]	$\mathbf{p_{1,2} = 0.003}$ $p_{1,3} = 0.282$ $\mathbf{p_{2,3} = 0.0001}$
$p_{ХВ}$[г]	$\chi^2 = 0.39$, $p = 0.5$	$\chi^2 = 0.09$, $p = 0.8$	$\chi^2 = 0.24$, $p = 0.6$	

[а] – число животных с данным генотипом или хромосом с данным аллелем; [б] – число животных; [в] – значения *p*, вычисленные для сравнения между группами здоровых (1), вирусоносителей (2) и больных (3) животных; [г] – результаты проверки соответствия распределения частот генотипов и аллелей равновесию Харди–Вайнберга с помощью критерия χ^2.

Выявлено достоверное снижение частоты генотипа A/A по ОНП rs110861313 у животных с персистентным лимфоцитозом (11.2 %) по сравнению с вирусоносителями (27.6 %) ($p < 0.01$). При этом частота особей с генотипом C/C у животных с персистентным лимфоцитозом (41.1 %) была достоверно выше, чем у вирусоносителей (21.4 %) ($p < 0.001$). В результате у животных с персистентным лимфоцитозом была существенно снижена частота аллеля A (35.0 %) и увеличена частота аллеля C (65.0 %) по сравнению с вирусоносителями (53.1 и 46.9 % соответственно) ($p = 0.003$). Не найдено статистически значимых различий в частотах генотипов и аллелей ОНП rs110861313 между группами больных и здоровых животных.

С применением программы PERFECTOS-APE проведен поиск потенциальных сайтов связывания транскрип-

Таблица 2. Характеристика транскрипционных факторов, экспрессирующихся в клетках иммунной системы, потенциальные сайты связывания которых имели большее сходство* с весовой матрицей для аллеля A (rs110861313), ассоциированного с лейкозом

Название предсказанного ССТФ	Официальный символ гена, кодирующего ТФ	Официальное название гена, кодирующего ТФ	Последовательность сайта для аллеля A/цепь ДНК	Отношение значений *p*-value для референсного и альтернативного аллелей	Экспрессия в иммунной системе**
TBR1	*TBR1*	T-box, brain 1	aaaagtgtgcAaa/+	5.03	Костный мозг, лейкоциты, лимфоузел, тимус
FOXM1	*FOXM1*	Forkhead box M1	tgttTgcacact/–	11.5	Костный мозг, лейкоциты, лимфоузел, тимус
FOXG1	*FOXG1*	Forkhead box G1	gttgttTgcacactttt/–	5.22	Костный мозг, лимфоузел, тимус
FOXD3	*FOXD3*	Forkhead box D3	gttgttTgcacactt/–	9.11	Костный мозг, лимфоузел, тимус
ZSCA4	*ZSCAN4*	Zinc finger and SCAN domain containing 4	aaagtgtgcAaa/+	6.39	Костный мозг, лимфоузел, тимус
MTF1	*MTF1*	Metal regulatory transcription factor 1	ctaaaagtgtgcAaaca/+	12.8	Костный мозг, лейкоциты, лимфоузел, тимус

* Сходство оценивалось программой PERFECTOS-APE на основе значения *p*-value.
** По данным базы GeneCards.

ционных факторов (ССТФ), сходство которых с весовой матрицей могло измениться в результате нуклеотидной замены в районе rs110861313. Благодаря программе было выявлено шесть ССТФ, сходство которых с весовой матрицей оценивалось значением $p\text{-value} < 5 \cdot 10^{-4}$ и было выше для альтернативного аллеля A, связанного с устойчивостью к лейкозу. Оказалось, что все шесть сайтов соответствовали транскрипционным факторам, экспрессирующимся в клетках иммунной системы (TBR1, FOXM1, FOXG1, ZSCA4, FOXD3, MTF1) (табл. 2).

Обсуждение

С помощью метода ПЦР-ПДРФ была исследована частота генотипов и аллелей ОНП rs110861313 в межгенном участке хромосомы 23 в выборках здоровых и больных лейкозом животных, а также бессимптомных вирусоносителей черно-пестрой породы крупного рогатого скота. Наблюдаемое распределение генотипов ОНП rs110861313 существенно не отличалось от ожидаемого по распределению Харди–Вайнберга (см. табл. 1). Хотя популяции сельскохозяйственных животных и человека не отвечают условиям действия закона Харди–Вайнберга (популяция бесконечно большого размера, в которой не действует естественный отбор, не идет мутационный процесс, отсутствует обмен особями с другими популяциями, не происходит дрейф генов, все скрещивания случайны), этот факт косвенно может свидетельствовать об отсутствии ошибок генотипирования в нашем исследовании (Hosking et al., 2004).

В литературе неоднократно обсуждался вопрос об адекватном контроле при поиске генетических вариантов, ассоциированных с развитием инфекционных заболеваний (Loeb, 2013; Юдин и др., 2018a). Популяционную выборку здоровых животных нельзя считать оптимальным контролем, поскольку она состоит из животных как с устойчивым, так и чувствительным генотипами, которые просто не успели контактировать с инфекционным агентом. Поэтому считается, что контрольную группу целесообразно формировать из животных с клинически доказанным событием инфицирования, у которых заболевание протекало в бессимптомной форме. В нашем случае это животные-вирусоносители.

Нами обнаружено достоверное снижение частоты генотипа A/A по ОНП rs110861313 у больных лейкозом животных по сравнению с вирусоносителями (см. табл. 1). Частота животных с генотипом C/C, наоборот, у больных животных была достоверно выше, чем у вирусоносителей. При этом у больных животных была существенно снижена частота аллеля A и увеличена частота аллеля C по сравнению с вирусоносителями. Эти результаты хорошо соответствуют данным H.A. Carignano с коллегами (2018), которые с помощью полногеномного анализа ассоциаций показали ассоциацию аллеля C ОНП rs110861313 с величиной провирусной нагрузки и числом лейкоцитов у животных, инфицированных ВЛКРС.

ОНП rs110861313 локализован в межгенном участке хромосомы 23 крупного рогатого скота. На расстоянии 30 тыс. п. н. от него находятся девять генов (табл. 3). Из них пять генов – *LY6G5B*, *GPANK1*, *ABHD16A*, *LY6G6F*, *LY6G6E* – кодируют белки иммунной системы. Ген *CSNK2B* кодирует бета-субъединицу казеинкиназы 2 (CK2) – высококонсервативной полифункциональной серин/треониновой протеинкиназы, экспрессирующейся во всех тканях. Повышение уровня и активности CK2 характерно для опухолевых клеток. Для онкогенеза особенно важны антиапоптотические функции CK2: эта протеинкиназа способствует репарации ДНК, влияет на сигнальные каскады NF-kB, Wnt, PI3K/Akt и JAK-STAT, взаимодействует с шаперонами, активирует антиапоптотические и инактивирует проапоптотические белки, в том

Таблица 3. Функциональная характеристика генов, локализованных на расстоянии не более 30 тыс. п. н. от ОНП rs110861313 в межгенном участке хромосомы 23 крупного рогатого скота

Официальный символ гена	Официальное название гена	Расстояние от ОНП rs110861313 в п. н./область гена	Функция
LY6G5B	Lymphocyte antigen 6 family member G5B	899/3′-область	Ортолог этого гена у человека кодирует белок, участвующий в иммунном ответе и распознавании клеток
CSNK2B	Casein kinase 2 beta	3447/3′-область	Ген «домашнего хозяйства», кодирует бета-субъединицу казеин-киназы II, которая регулирует метаболические пути, сигнальную трансдукцию, транскрипцию, трансляцию и репликацию
GPANK1	G-patch domain and ankyrin repeats 1	7753/5′-область	Ортолог этого гена у человека расположен в кластере генов III главного комплекса гистосовместимости, кодирует белок, участвующий в иммунных реакциях
ABHD16A	Abhydrolase domain containing 16A	8346/3′-область	»
C23H6orf47	Chromosome 23 C6orf47 homolog	13442/5′-область	Функция неизвестна
ApoM	Apolipoprotein M	15211/3′-область	Белок, кодируемый этим геном, является аполипопротеином и членом семейства белков липокалина. Он связан с липопротеинами высокой плотности и, в меньшей степени, с липопротеинами низкой плотности и богатыми триглицеридами липопротеинами
BAG6	BAG cochaperone 6	19374/5′-область	Ортолог этого гена у человека расположен в кластере генов III главного комплекса гистосовместимости, кодирует ядерный белок, который расщепляется каспазой 3 и участвует в регуляции апоптоза
LY6G6F	Lymphocyte antigen 6 family member G6F	23454/5′-область	Оролог этого белка у человека участвует в иммунном ответе и распознавании клеток
LY6G6E	Lymphocyte antigen 6 family member G6E	27718/3′-область	»

числе каспазы (Володина, Штиль, 2012). Аполипопротеин M (ApoM) – переносчик липидного медиатора сфингозин-1-фосфата (S1P). Сигнальные пути, включающие сфингозинкиназы (SphKs) и рецепторы S1P (S1PR), играют важную роль в онкогенезе множественных опухолей, включая немелкоклеточный рак легкого (NSCLC). Показано, что сверхэкспрессия ApoM может стимулировать пролиферацию и инвазию клеток NSCLC *in vitro* и рост опухоли *in vivo*, вероятно, через активацию S1PR1 и сигнальных путей ERK1/2 и PI3K/AKT (Zhu et al., 2018).

В литературе нами не встречено упоминания о каком-либо из вышеперечисленных генов в связи с инфекцией ВЛКРС. Тем не менее у людей, больных Т-клеточным лейкозом взрослых (Adult T cell leukemia/lymphoma) и инфицированных вирусом HTLV-I, была обнаружена повышенная частота соматических мутаций в гене *CSNK2B* (Kataoka et al., 2015). Интересно, что мутации в этом гене у человека ассоциированы с развитием шизофрении (Yang et al., 2018; Niu et al., 2019), судорожных припадков (Nakashima et al., 2019) и эпилепсии (Poirier et al., 2017; Sakaguchi et al., 2017). В связи с этим представляется перспективным исследовать изменение поведения животных после инфицирования ВЛКРС.

С помощью биоинформатического анализа нами установлено, что замена нуклеотида C→A rs110861313 может быть связана с образованием сайтов связывания транскрипционных факторов TBR1, FOXM 1, FOXG1, ZSCAN4, FOXD3, и MTF1 (см. табл. 2). Все эти транскрипционные факторы экспрессируются в клетках иммунной системы. Наибольший интерес представляет фактор FOXM1, служащий активатором транскрипции (Wang et al., 2008). Известно, что этот фактор регулирует экспрессию генов клеточного цикла, важных для репликации ДНК и митоза, контролирует пролиферацию клеток и репарацию ДНК (Nandi et al., 2018). Так, FOXM1 опосредует пролиферацию, выживание, миграцию/инвазию, прогрессию и онкогенез клеток «трижды негативного» рака молочной железы (Hamurcu et al., 2016). Этот же белок посредством регуляции экспрессии экзонуклеазы I может модулировать чувствительность клеток рака яичников к химиотерапии цисплатином (Zhou et al., 2014). Уровень экспрессии FOXM1 коррелирует с прогнозом при раке желудка (Li et al., 2018) и почки (Liang et al., 2018). Наконец, мета-анализ паттернов экспрессии 18000 образцов опухолей человека, относившихся к 39 типам злокачественных новообразований, идентифицировал регуляторную сеть FOXM1 как основной предиктор неблагоприятного прогноза заболевания (Gentles et al., 2015).

Мы предполагаем, что у животных с генотипом A/A ОНП rs110861313 образуется сайт на ДНК, с которым может связываться транскрипционный фактор FOXM1 и, таким образом, активировать транскрипцию близлежащих генов. В конечном счете это может привести к подавлению инфекции ВЛКРС и отсутствию персистентного лимфоцитоза у вирусоносителей. Однако для проверки гипотезы о наличии взаимодействия ДНК с факторами

транскрипции необходимо провести эксперименты *in vivo* и *in vitro* (Merkulova et al., 2007). Экспериментальное подтверждение связывания транскрипционных факторов с полиморфным участком ДНК, а также влияния нуклеотидной замены на интенсивность связывания будет свидетельствовать о возможном влиянии ОНП rs110861313 на экспрессию близлежащих генов и, вследствие этого, на устойчивость к лейкозу.

Заключение

Нами установлено, что однонуклеотидный полиморфизм rs110861313 в межгенном участке хромосомы 23 ассоциирован с развитием лейкоза при инфицировании ВЛКРС у животных черно-пестрой породы крупного рогатого скота. По данным биоинформатического анализа, устойчивость к инфекции ВЛКРС может быть ассоциирована с наличием в районе ОНП rs110861313 сайта связывания транскрипционного фактора FOXM1, который потенциально может влиять на экспрессию ближайших генов. Наши результаты указывают на возможность использования этого генетического маркера для маркер-ориентированной и геномной селекции (Юдин, Воевода, 2015).

Список литературы / References

Володина Ю.Л., Штиль А.А. Казеинкиназа 2 – универсальный регулятор выживаемости клеток. Молекуляр. биология. 2012; 46(3):423-433.
[Volodina Y.L., Shtil A.A. Casein kinase 2, a versatile regulator of cell survival. Mol. Biol. 2012;46(3):381-390.]

Методические указания по диагностике лейкоза крупного рогатого скота (утв. Минсельхозом РФ 23.08.2000 № 13-7-2/2130). Доступно по адресу: http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=371364#08727407013115032. Проверено 18 сентября 2019 г.
[Guidelines on the Diagnosis of Bovine Leukemia (approved by the RF Ministry of Agriculture on Aug. 23, 2000, No. 13-7-2/2130). Available from http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=371364#0872740 7013115032 (18.09.2019). (in Russian)]

Юдин Н.С., Бархаш А.В., Максимов В.Н., Игнатьева Е.В., Ромащенко А.Г. Генетическая предрасположенность человека к заболеваниям, вызываемым флавивирусами. Молекуляр. биология. 2018а;52(2):190-209. DOI 10.7868/S0026898418020039.
[Yudin N.S., Barkhash A.V., Maksimov V.N., Ignatieva E.V., Romaschenko A.G. Human genetic predisposition to diseases caused by viruses from flaviviridae family. Mol. Biol. 2018a;52(2):165-181. DOI 10.1134/S0026893317050223.]

Юдин Н.С., Воевода М.И. Молекулярно-генетические маркеры экономически важных признаков у молочного скота. Генетика. 2015;51(5):600-612. DOI 10.7868/S0016675815050082.
[Yudin N.S., Voevoda M.I. Molecular genetic markers of economically important traits in dairy cattle. Russ. J. Genet. 2015;51(5):506-517. DOI 10.1134/S1022795415050087.]

Юдин Н.С., Подколодный Н.Л., Агаркова Т.А., Игнатьева Е.В. Приоритизация генов, ассоциированных с патогенезом лейкоза у крупного рогатого скота. Вавиловский журнал генетики и селекции. 2018б;22(8):1063-1069. DOI 10.18699/VJ18.451.
[Yudin N.S., Podkolodnyy N.L., Agarkova T.A., Ignatieva E.V. Prioritization of genes associated with the pathogenesis of leukosis in cattle. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018b;22(8):1063-1069. DOI 10.18699/VJ18.451. (in Russian)]

Abdalla E.A., Rosa G.J., Weigel K.A., Byrem T. Genetic analysis of leukosis incidence in United States Holstein and Jersey populations. J. Dairy Sci. 2013;96(9):6022-6029. DOI 10.3168/jds.2013-6732.

Brym P., Bojarojć-Nosowicz B., Oleński K., Hering D.M., Ruść A., Kaczmarczyk E., Kamiński S. Genome-wide association study for host response to bovine leukemia virus in Holstein cows. Vet. Immunol. Immunopathol. 2016;175:24-35. DOI 10.1016/j.vetimm.2016.04.012.

Burny A., Cleuter Y., Kettmann R., Mammerickx M., Marbaix G., Portetelle D., Broeke A., Willems L., Thomas R. Bovine leukemia: facts and hypotheses derived from the study of an infectious cancer. Vet. Microbiol. 1988;17:197-218.

Carignano H.A., Roldan D.L., Beribe M.J., Raschia M.A., Amadio A., Nani J.P., Gutierrez G., Alvarez I., Trono K., Poli M.A., Miretti M.M. Genome-wide scan for commons SNPs affecting bovine leukemia virus infection level in dairy cattle. BMC Genomics. 2018; 19(1):142. DOI 10.1186/s12864-018-4523-2.

Ferrer J.F. Bovine lymphosarcoma. Adv. Vet. Sci. Comp. Med. 1980; 24:1-68.

Gentles A.J., Newman A.M., Liu C.L., Bratman S.V., Feng W., Kim D., Nair V.S., Xu Y., Khuong A., Hoang C.D., Diehn M., West R.B., Plevritis S.K., Alizadeh A.A. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat. Med. 2015; 21(8):938-945. DOI 10.1038/nm.3909.

Gillet N., Florins A., Boxus M., Burteau C., Nigro A., Vandermeers F., Balon H., Bouzar A.B., Defoiche J., Burny A., Reichert M., Kettmann R., Willems L. Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in human. Retrovirology. 2007;4:18.

Hamurcu Z., Ashour A., Kahraman N., Ozpolat B. FOXM1 regulates expression of eukaryotic elongation factor 2 kinase and promotes proliferation, invasion and tumorgenesis of human triple negative breast cancer cells. Oncotarget. 2016;7(13):16619-16635. DOI 10.18632/oncotarget.7672.

Hosking L., Lumsden S., Lewis K., Yeo A., McCarthy L., Bansal A., Riley J., Purvis I., Xu C.F. Detection of genotyping errors by Hardy–Weinberg equilibrium testing. Eur. J. Hum. Genet. 2004;12(5): 395-399.

Kataoka K., Nagata Y., Kitanaka A., Shiraishi Y., Shimamura T., Yasunaga J., Totoki Y., Chiba K., Sato-Otsubo A., Nagae G., Ishii R., Muto S., Kotani S., Watatani Y., Takeda J., Sanada M., Tanaka H., Suzuki H., Sato Y., Shiozawa Y., Yoshizato T., Yoshida K., Makishima H., Iwanaga M., Ma G., Nosaka K., Hishizawa M., Itonaga H., Imaizumi Y., Munakata W., Ogasawara H., Sato T., Sasai K., Muramoto K., Penova M., Kawaguchi T., Nakamura H., Hama N., Shide K., Kubuki Y., Hidaka T., Kameda T., Nakamaki T., Ishiyama K., Miyawaki S., Yoon S.S., Tobinai K., Miyazaki Y., Takaori-Kondo A., Matsuda F., Takeuchi K., Nureki O., Aburatani H., Watanabe T., Shibata T., Matsuoka M., Miyano S., Shimoda K., Ogawa S. Integrated molecular analysis of adult T cell leukemia/lymphoma. Nat. Genet. 2015;47(11):1304-1315. DOI 10.1038/ng.3415.

Kulakovskiy I.V., Vorontsov I.E., Yevshin I.S., Sharipov R.N., Fedorova A.D., Rumynskiy E.I., Medvedeva Y.A., Magana-Mora A., Bajic V.B., Papatsenko D.A., Kolpakov F.A., Makeev V.J. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Res. 2018;46(D1):D252-D259. DOI 10.1093/nar/gkx1106.

Li X., Ma K., Song S., Shen F., Kuang T., Zhu Y., Liu Z. Tight correlation between FoxM1 and FoxP3+ Tregs in gastric cancer and their clinical significance. Clin. Exp. Med. 2018;18(3):413-420. DOI 10.1007/s10238-018-0505-6.

Liang J., Liu Z., Zou Z., Tang Y., Zhou C., Yang J., Wei X., Lu Y. The correlation between the immune and epithelial-mesenchymal transition signatures suggests potential therapeutic targets and prognosis prediction approaches in kidney cancer. Sci. Rep. 2018;8(1): 6570. DOI 10.1038/s41598-018-25002-w.

Loeb M. Host genomics in infectious diseases. Infect. Chemother. 2013;45(3):253-259. DOI 10.3947/ic.2013.45.3.253.

Merkulova T.I., Oshchepkov D.Y., Ignatieva E.V., Ananko E.A., Levitsky V.G., Vasiliev G.V., Klimova N.V., Merkulov V.M., Kol-

chanov N.A. Bioinformatical and experimental approaches to investigation of transcription factor binding sites in vertebrate genes. Biochemistry (Mosc). 2007;72(11):1187-1193.

Nakashima M., Tohyama J., Nakagawa E., Watanabe Y., Siew C.G., Kwong C.S., Yamoto K., Hiraide T., Fukuda T., Kaname T., Nakabayashi K., Hata K., Ogata T., Saitsu H., Matsumoto N. Identification of de novo CSNK2A1 and CSNK2B variants in cases of global developmental delay with seizures. J. Hum. Genet. 2019;64(4):313-322. DOI 10.1038/s10038-018-0559-z.

Nandi D., Cheema P.S., Jaiswal N., Nag A. FoxM1: Repurposing an oncogene as a biomarker. Semin. Cancer Biol. 2018;52(Pt. 1):74-84. DOI 10.1016/j.semcancer.2017.08.009.

Neff M.M., Turk E., Kalishman M. Web-based primer design for single nucleotide polymorphism analysis. Trends Genet. 2002;18(12):613-615.

Niu H.M., Yang P., Chen H.H., Hao R.H., Dong S.S., Yao S., Chen X.F., Yan H., Zhang Y.J., Chen Y.X., Jiang F., Yang T.L., Guo Y. Comprehensive functional annotation of susceptibility SNPs prioritized 10 genes for schizophrenia. Transl. Psychiatry. 2019;9(1):56. DOI 10.1038/s41398-019-0398-5.

Poirier K., Hubert L., Viot G., Rio M., Billuart P., Besmond C., Bienvenu T. CSNK2B splice site mutations in patients cause intellectual disability with or without myoclonic epilepsy. Hum. Mutat. 2017; 38(8):932-941. DOI 10.1002/humu.23270.

Sakaguchi Y., Uehara T., Suzuki H., Kosaki K., Takenouchi T. Truncating mutation in CSNK2B and myoclonic epilepsy. Hum. Mutat. 2017;38(11):1611-1612. DOI 10.1002/humu.23307.

Sambrook J., Russell D.W. The Condensed Protocols from Molecular Cloning: a Laboratory Manual, Cold Spring Harbor; New York: Cold Spring Harbor Laboratory Press, 2006.

Stepanova T.V. Analysis of the economic damage caused by bovine leukemia from 2010 to 2014 in the Russian Federation. Russ. J. Agric. Socio-Economic Sci. 2016;8(56):49-56. DOI 10.18551/rjoas.2016-08.08.

Wang I.C., Chen Y.J., Hughes D.E., Ackerson T., Major M.L., Kalinichenko V.V., Costa R.H., Raychaudhuri P., Tyner A.L., Lau L.F. FoxM1 regulates transcription of JNK1 to promote the G1/S transition and tumor cell invasiveness. J. Biol. Chem. 2008;283(30):20770-20778. DOI 10.1074/jbc.M709892200.

Yang C.P., Li X., Wu Y., Shen Q., Zeng Y., Xiong Q., Wei M., Chen C., Liu J., Huo Y., Li K., Xue G., Yao Y.G., Zhang C., Li M., Chen Y., Luo X.J. Comprehensive integrative analyses identify GLT8D1 and CSNK2B as schizophrenia risk genes. Nat. Commun. 2018;9(1):838. DOI 10.1038/s41467-018-03247-3.

Zhou J., Wang Y., Wang Y., Yin X., He Y., Chen L., Wang W., Liu T., Di W. FOXM1 modulates cisplatin sensitivity by regulating EXO1 in ovarian cancer. PLoS One. 2014;9(5):e96989. DOI 10.1371/journal.pone.0096989.

Zhu Y., Luo G., Jiang B., Yu M., Feng Y., Wang M., Xu N., Zhang X. Apolipoprotein M promotes proliferation and invasion in non-small cell lung cancers via upregulating S1PR1 and activating the ERK1/2 and PI3K/AKT signaling pathways. Biochem. Biophys. Res. Commun. 2018;501(2):520-526. DOI 10.1016/j.bbrc.2018.05.029.

ORCID ID
E.V. Ignatieva orcid.org/0000-0002-8588-6511
N.S. Yudin orcid.org/0000-0002-1947-5554

Благодарности. Работа выполнена при поддержке Комплексной программы фундаментальных исследований СО РАН.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 03.10.2019. После доработки 14.11.2019. Принята к публикации 15.11.2019.

ГЕНЕТИКА И СЕЛЕКЦИЯ ЖИВОТНЫХ
Оригинальное исследование / Original article
УДК 636.028:615.357

Вавиловский журнал генетики и селекции. 2019;23(8):1006-1010
DOI 10.18699/VJ19.577

Влияние экзогенного хорионического гонадотропина человека на овуляцию у мышей

С.Я. Амстиславский[1, 2] ✉, С.В. Раннева[1, 2], Д.С. Рагаева[1], Э.А. Чуйко[2], А.М. Попкова[2], Е.Ю. Брусенцев[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ e-mail: amstis@bionet.nsc.ru

Применение некоторых вспомогательных репродуктивных технологий (ВРТ), в частности гормональной стимуляции овуляции, может привести к изменению числа и качества получаемых ооцитов. Целью исследования было изучение характера овуляции у мышей линии CD1 после воздействия на самок хорионическим гонадотропином человека (ХГЧ) и исследование последствий этого воздействия на длительность беременности, эмбриональные потери и вес рожденных потомков. При воздействии препаратом ХГЧ не было обнаружено достоверных различий по общему числу овулировавших ооцитов по сравнению с контрольными самками, а также различий по числу незрелых (без полярного тела) и зрелых (с формирующимся полярным телом) форм овулировавших ооцитов. Число же ооцитов с полярным телом у мышей экспериментальной группы после гормональной стимуляции овуляции препаратом ХГЧ было достоверно больше ($p < 0.05$) по сравнению с контрольной группой (6.2 ± 0.86 и 2.2 ± 0.9 соответственно). Не было получено достоверных различий между группами по продолжительности беременности и числу рожденных потомков, включая долю живо- и мертворожденных особей. Однако вес детенышей на пятый день после рождения в экспериментальной группе был достоверно меньше ($p < 0.001$), чем в контроле (3.16 ± 0.09 и 3.76 ± 0.07 соответственно). Таким образом, введение экзогенного ХГЧ самкам мышей стимулирует развитие ооцитов *in vivo*, что приводит к формированию большего числа их зрелых форм, но потомки, рожденные самками, стимулированными препаратом ХГЧ, имеют меньший вес тела в первые дни после рождения.
Ключевые слова: мыши; ооциты; хорионический гонадотропин человека.

Для цитирования: Амстиславский С.Я., Раннева С.В., Рагаева Д.С., Чуйко Э.А., Попкова А.М., Брусенцев Е.Ю. Влияние экзогенного хорионического гонадотропина человека на овуляцию у мышей. Вавиловский журнал генетики и селекции. 2019;23(8):1006-1010. DOI 10.18699/VJ19.577

Effect of exogenous human chorionic gonadotropin on ovulation in mice

S.Ya. Amstislavsky[1, 2] ✉, S.V. Ranneva[1, 2], D.S. Ragaeva[1], E.A. Chuyko[2], A.M. Popkova[2], E.Yu. Brusentsev[1]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
✉ e-mail: amstis@bionet.nsc.ru

The implementation of assisted reproductive technologies (ART), hormonal stimulation in particular, may change the quality of ovulated oocytes. The purpose of our work was to study ovulation in CD1 mice after their stimulation with human chorionic gonadotropin (hCG) and to investigate the effects of such hormonal stimulation on the pregnancy duration, fetal losses and the weight of the offspring. No significant differences were found in the total number of ovulated oocytes or in the number of immature (without a polar body) ovulated oocytes; nor were there differences between the groups in the number of oocytes with a developing polar body. However, the number of matured oocytes with a distinct polar body was significantly higher ($p < 0.05$) in mice stimulated with hCG (experimental group) as compared with the controls (6.2 ± 0.86 and 2.2 ± 0.97, respectively). No significant differences were observed between the experimental and control mice in the duration of pregnancy or in the numbers of term offspring, including the percentage of live and stillborn pups. However, the body weight of the offspring in the experimental group was significantly lower ($p < 0.001$) as compared with the controls on the fifth day after birth (3.16 ± 0.09 and 3.76 ± 0.07, respectively). Thus, exogenous hCG facilitates the development of mouse oocytes *in vivo*, which leads to the larger number of their mature forms at ovulation, however, the offspring born after hCG-stimulated pregnancy was characterized by a lower body weight on the fifth day after birth.
Key words: mice; oocytes; human chorionic gonadotropin.

For citation: Amstislavsky S.Ya., Ranneva S.V., Ragaeva D.S., Chuyko E.A., Popkova A.M., Brusentsev E.Yu. Effect of exogenous human chorionic gonadotropin on ovulation in mice. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1006-1010. DOI 10.18699/VJ19.577 (in Russian)

Введение

Хорионический гонадотропин человека (ХГЧ) играет важную роль во время беременности. У человека ХГЧ начинает вырабатываться уже на 6–8-й день беременности тканью хориона, отсюда и происходит название гормона; его концентрация возрастает в несколько тысяч раз к 7–11-й неделе, а затем начинает постепенно снижаться (Fournier et al., 2015). Этот гормон обладает лютеинизи-

© Амстиславский С.Я., Раннева С.В., Рагаева Д.С., Чуйко Э.А., Попкова А.М., Брусенцев Е.Ю., 2019

рующей активностью, будучи агонистом рецепторов лютеинизирующего гормона (ЛГ), и превосходит его по этим свойствам (Keay et al., 2004). Благодаря ХГЧ в первые месяцы беременности сохраняется активность желтого тела, которое обеспечивает продукцию прогестерона – «гормона беременности».

ХГЧ характеризуется структурным сходством с ЛГ, оба гормона воздействуют на один и тот же рецептор – ЛГ/ХГЧ (Drakakis et al., 2009). Фактор, отличающий ХГЧ от ЛГ, – более длительный период полувыведения, равный 36 ч, тогда как период полувыведения рекомбинантного ЛГ оценивается примерно в 10–12 ч (Cole, 2010). ХГЧ демонстрирует более сильную аффинность связывания с ЛГ/ХГЧ-рецептором (Drakakis et al., 2009). Независимо от наличия в схеме гормональной стимуляции фолликулостимулирующего гормона, низкие дозы ХГЧ могут поддерживать развитие и созревание наиболее крупных фолликулов яичника, клетки которых приобрели рецепторы ЛГ/ХГЧ в процессе своего роста (Cole, 2010).

Хорионический гонадотропин человека применяется для контроля овуляции у человека и других млекопитающих, а также при лечении бесплодия у обоих полов (Delvigne, Rozenberg, 2002; Homburg, 2004; Амстиславский, 2006). Благодаря ранней продукции гормона трофобластом эмбриона, ХГЧ служит гормональным маркером хромосомных нарушений зародыша у человека (Борисова и др., 2017). Безопасными и эффективными протоколами стимуляции суперовуляции у женщин признаются протоколы с применением гормональных препаратов (Homburg, 2004). В частности, для этих целей широко используют группу гонадотропинов, важнейшим из которых является ХГЧ (Delvigne, Rozenberg, 2002; Homburg, 2004).

В ранних работах было показано, что введение обладающего фолликулостимулирующей активностью гонадотропина сыворотки жеребых кобыл (ГСЖК) на разных стадиях эстрального цикла мышей влияет на качество получаемых после этой процедуры ооцитов (Tarin et al., 2002). В работе (Wang et al., 2006) было установлено, что преимплантационные эмбрионы, полученные из незрелых ооцитов после последовательной инъекции ГСЖК и ХГЧ, с последующим дозреванием *in vitro* (*in vitro* maturation, IVM) и после искусственного оплодотворения (*in vitro* fertilization, IVF), хуже развиваются при культивировании *in vitro* (*in vitro* culture IVC), чем полученные из дозревших *in vivo* ооцитов. Суперовуляция мышей с использованием инъекций ГСЖК и ХГЧ дает меньшее число зрелых ооцитов, но при этом повышает число ооцитов с деформациями митохондрий, сниженной митохондриальной активностью и продукцией АТФ, по сравнению с ооцитами, полученными при естественной овуляции (Lee et al., 2017). Лишь в единичных работах изучали эффекты гормональной стимуляции препаратом ХГЧ (без комбинации с ГСЖК) на развитие эмбрионов мышей (Ertzeid, Storeng, 1992; Dinopoulou et al., 2016). Тем не менее результаты этих работ противоречивы, требуют дальнейшей проверки.

В настоящее время исследователи разрабатывают новые протоколы суперовуляции. Так, синхронизация эстрального цикла у мышей с последующей инъекцией антиингибиновой сыворотки и ХГЧ позволяет получить в два раза больше зрелых ооцитов, чем стандартная схема ГСЖК и ХГЧ (Hasegawa et al., 2016). Было показано (Takeo, Nakagata, 2015), что совместное введение антиингибиновой сыворотки с препаратом ГСЖК и последующей инъекцией препарата ХГЧ повышало выход зрелых ооцитов по сравнению с их введением по отдельности. Однако во всех используемых протоколах гормональной стимуляции мышей с целью вызвать суперовуляцию обязательно применяется ХГЧ.

В нашей ранней работе на мышах линии DD выявлено, что доля зрелых форм овулировавших ооцитов была наибольшей после введения гонадотропинов (ГСЖК и ХГЧ) самкам на стадии эструса (Redina et al., 1994). В этом исследовании также было установлено, что после такого сочетанного воздействия двух гонадотропных препаратов у мышей овулируют ооциты разной степени зрелости (Redina et al., 1994). Тем не менее пока неизвестно, как отразится стимуляция самок мышей только препаратом ХГЧ на стадии эструса непосредственно перед спариванием с самцом на качестве ооцитов и дальнейшем протекании беременности. Задачи нашего исследования – изучить характер овуляции у мышей после воздействия экзогенным ХГЧ; установить, как влияет этот гормон на такие параметры, как длительность беременности, эмбриональные потери и вес рожденных потомков.

Материалы и методы

Экспериментальные животные. В эксперименте было использовано 10 самок мышей линии CD1 в возрасте 2.5 мес. для получения ооцитов и 14 самок мышей тех же линии и возраста для получения потомков. В контрольной и экспериментальной группах было по пять особей для сбора ооцитов и, соответственно, шесть и восемь особей для получения потомков. Для стерильного спаривания использовали пять вазэктомированных самцов для выделения ооцитов и пять фертильных самцов (для беременности). Животных содержали в клетках с подстилкой из опилок в стандартных условиях конвенционального вивария Института цитологии и генетики СО РАН (Новосибирск, Россия): при комфортной температуре 22–24 °C, свободном доступе к стандартному корму («Чара», Россия) и очищенной воде, при естественном режиме освещения. Все эксперименты на животных одобрены Комиссией по биоэтике Института цитологии и генетики СО РАН (протокол № 5 от 13.05.2011) и соответствуют Европейской конвенции о защите позвоночных животных, используемых для экспериментальных и других научных целей.

Протокол гормональной стимуляции овуляции. В нашей работе был использован протокол, описанный ранее (Redina et al., 1994), с модификациями. Для гормональной стимуляции овуляции использовали препарат ХГЧ Chorulon (Intervet, MSD Animal Health, Новая Зеландия). Самкам делали по одной инъекции препарата, растворенного в физиологическом растворе в дозировке 5 МЕ (100 мкл) внутрибрюшинно. Введение препарата осуществляли на стадии эструса, с 17:00 до 18:00 ч. Стадию цикла самки определяли по влагалищным мазкам. После инъекций самок подсаживали по одной к каждому вазэктомированному самцу для последующего сбора ооцитов либо к фертильному самцу для получения по-

томства. На утро следующего дня у самок проверяли наличие вагинальных пробок, по которым судили о том, что спаривание состоялось. Контрольным самкам вводили 100 мкл физиологического раствора и спаривали с самцами по аналогичной схеме.

Вазэктомия самцов. Процедуру вазэктомии проводили не менее чем за две недели до начала эксперимента, как было описано ранее, с небольшими модификациями (Hogan et al., 1994). Половозрелых самцов мышей линии CD1 наркотизировали путем внутрибрюшинного введения 0.25 мг/кг препарата медетомедина гидрохлорида (Domitor 1 мг/мл, Orion-Corporation, Финляндия) и через 10 мин 50 мг/кг препарата золетила (Zoletil, SA, Virbac Sante Animale, Франция). После наркотизации мышам подкожно вводили антибиотик: 0.01 мл амоксициллина тригидрата 150 мг/мл (ОАО «Синтез», Россия). Затем животных помещали на подогреваемый столик, шерсть в зоне операционного поля сбривали, а кожу обрабатывали 70 % этиловым спиртом. При помощи хирургических ножниц делали горизонтальный надрез длиной около 5 мм кожных покровов мошонки. Подтягивали эпидидимисы к краю хирургической раны и разворачивали их так, чтобы было видно семенные канатики. Семявыносящие каналы отделяли от сопряженных тканей и пережигали раскаленным пинцетом в двух местах, удаляя участок между ними. Эпидидимисы перемещали в первоначальное положение. Затем в рану присыпали 2 мг амоксициллина тригидрата (ОАО «Синтез», Россия). Закрывали надрезы наложением двух швов и обрабатывали их Раносаном (ООО «АПИ-САН», Россия).

Оценка ооцитов. Самок после продуктивного спаривания подвергали эвтаназии при помощи дислокации шейных позвонков. Ооциты извлекали через 20–22 ч после спаривания путем рассечения ампулярной части яйцеводов в питательной среде M2 (Merck, Германия). Для удаления кумулюсных клеток использовали гиалуронидазу (Merck) в концентрации 80 МЕ/мл (Brinster, 1971). Оценивали как по общему числу ооцитов, так и по отдельным категориям: без полярного тела, с формирующимся полярным телом и со сформированным полярным телом (стадия MII мейоза), как было описано ранее (Redina et al., 1994). Оценку ооцитов производили с помощью инвертированного светового микроскопа (DM IL LED, Leica Microsystems, Германия) при увеличении 10×, для фотодокументирования использовали камеру (DFC 295, Leica Microsystems, Германия).

Оценка продолжительности беременности и взвешивание рожденных потомков. Самок, имеющих вагинальные пробки после покрытия фертильными самцами, отсаживали в отдельные клетки. Учитывали долю покрытых самок, а также продолжительность беременности (день родов). Оценивали число рожденных потомков, а также долю живо- и мертворожденных детенышей. Кроме того, на пятый день после рождения производили взвешивание потомков.

Статистический анализ. Данные анализировали с помощью программы Statistica 6.0. Соответствие полученных значений нормальному распределению определяли при помощи критерия Шапиро–Уилка. Результаты по ооцитам, а также продолжительности беременности, числу рожденных потомков и весу тела представлены как среднее ± ошибка среднего. Для сравнения групп использовали *t*-критерий Стьюдента. Данные по покрытию самок, а также по числу живо- и мертворожденных потомков представлены как процент от общего. При сравнении групп использовали критерий χ^2. Значения при $p < 0.05$ считали статистически достоверными.

Результаты

Данные по оценке числа и качества ооцитов приведены в табл. 1. Статистический анализ не выявил достоверных различий по общему числу овулировавших ооцитов, а также по числу ооцитов без полярного тела и числу ооцитов с формирующимся полярным телом между экспериментальной и контрольной группами. При этом число ооцитов со сформированным полярным телом было достоверно больше после гормональной стимуляции самок препаратом ХГЧ ($p < 0.05$) по сравнению с контрольной группой (6.2 ± 0.86 и 2.2 ± 0.97 соответственно). Ооциты трех оцениваемых категорий показаны на рисунке.

Данные по оценке продолжительности беременности и рожденных потомков представлены в табл. 2. Статисти-

Таблица 1. Овуляция у мышей линии CD1 после воздействия ХГЧ

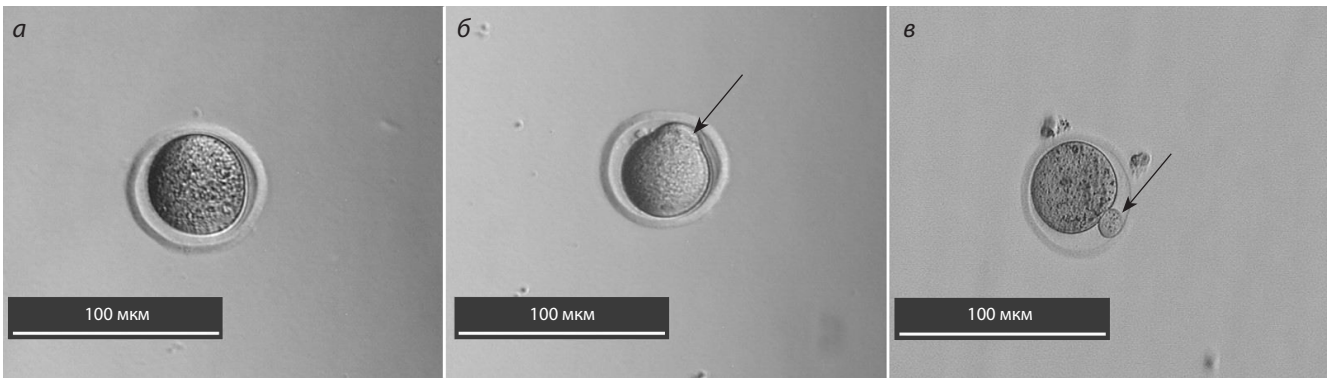
Ооциты	Группа (число животных)	
	Контроль ($n = 5$)	ХГЧ ($n = 5$)
Среднее число овулировавших ооцитов	12.80 ± 0.58	13.40 ± 1.81
Без полярного тела	3.40 ± 1.12	1.60 ± 0.81
С формирующимся полярным телом	7.20 ± 1.93	5.60 ± 1.60
Со сформированным полярным телом	2.20 ± 0.97	6.20 ± 0.86*

* $p < 0.05$ по сравнению с контролем.

Таблица 2. Оценка продолжительности беременности и рожденного потомства мышей линии CD1 после введения экзогенного ХГЧ

Параметр	Группа (число животных)	
	Контроль ($n = 6$)	ХГЧ ($n = 8$)
Число самок с влагалищными пробками, %	6 (100)	6 (75.00)
Продолжительность беременности, дни	19.67 ± 0.33	19.00 ± 0.32
Число потомков на самку	11.83 ± 3.31	10.83 ± 5.78
живорожденных, %	71 (100)	65 (90.28)
мертворожденных, %	0 (0)	7 (9.72)
Вес детенышей, г	3.76 ± 0.07	3.16 ± 0.09***

*** $p < 0.001$ по сравнению с контролем.



Ооциты трех оцениваемых категорий: *а* – без полярного тела; *б* – с формирующимся полярным телом; *в* – с полярным телом. Полярное тело указано стрелкой.

ческий анализ не выявил достоверных различий между экспериментальной и контрольной группами по доле покрытых самок, а также по продолжительности беременности и числу рожденных потомков, включая живо- и мертворожденных особей. Тем не менее вес детенышей на пятый день после рождения в экспериментальной группе был достоверно ниже ($p < 0.001$), чем в контроле (3.16 ± 0.09 и 3.76 ± 0.07 соответственно).

Обсуждение

Оценка качества ооцитов, полученных после гормонального воздействия на самок мышей препаратом ХГЧ, показала возрастание числа зрелых форм со сформированным полярным телом по сравнению с контрольной группой. Наши данные согласуются с недавними результатами по IVM ооцитов мышей со стадии герминального везикула, где было установлено, что скорость дозревания в среде с добавлением ХГЧ была значительно выше по сравнению с контрольной группой, а также улучшалось эмбриональное развитие при воздействии препаратом ХГЧ в ходе IVC (Dinopoulou et al., 2016). В работе по влиянию гонадотропинов на доимплантационное развитие мышей (Ertzeid, Storeng, 1992) было выявлено, что после стимуляции овуляции препаратом ХГЧ увеличилось среднее число эмбрионов на мышь, но уменьшилась доля морфологически нормальных эмбрионов по сравнению с группой со спонтанной овуляцией. Улучшение качественного состава овулировавших ооцитов, обнаруженное в нашей работе, по сравнению с исследованием G. Ertzeid и R. Storeng (1992), может быть обусловлено методическими различиями при введении препарата ХГЧ. В наших экспериментах препарат ХГЧ вводили мышам лишь на стадии эструса, в то время как G. Ertzeid и R. Storeng (1992) стадию цикла мышей на момент введения препарата не учитывали. Как было установлено нами ранее (Redina et al., 1994), качество овулировавших ооцитов при введении гормональных препаратов существенно зависит от стадии эстрального цикла на начало гормональной стимуляции.

Результаты изучения воздействия препаратом ХГЧ во время беременности у людей на развитие эмбрионов достаточно противоречивы. Так, в настоящее время накоплено множество данных по позитивному влиянию ХГЧ на пре- и постимплантационный периоды развития зародыша человека (Kane et al., 2009; Strug et al., 2016; Li et al., 2017; Makrigiannakis et al., 2017). В одном из исследований на людях было обнаружено, что ХГЧ является регулятором пролиферации маточных NK-клеток (natural killer), которые играют важную роль в успешном протекании беременности у человека, воздействуя через рецептор маннозы CD206 (Kane et al., 2009). В работе на приматах было установлено, что после стимуляции яичников внутриматочная инфузия ХГЧ повышает число рецепторов ESR1 и PGR эндометрия, задерживает развитие его стромы, а также способствует синхронизации донора и реципиента, что в конечном счете способствует успешной трансплантации эмбрионов (Strug et al., 2016).

ХГЧ активно используют в репродуктивной медицине для иммуномодуляции эндометрия перед трансплантацией эмбрионов, поскольку он выступает в качестве активатора введенных в полость матки аутологичных мононуклеарных клеток периферической крови, что, в свою очередь, повышает эффективность имплантации (Li et al., 2017). Все эти данные свидетельствуют лишь об опосредованном влиянии ХГЧ на развивающийся плод и не учитывают отдаленных эффектов на рожденное потомство.

В более раннем исследовании на мышах было отмечено негативное влияние как комбинации двух гормональных препаратов, ГСЖК и ХГЧ, так и, собственно, экзогенного ХГЧ на пре- и постимплантационное развитие зародышей и возрастание процента резорбированных плодов (Ertzeid, Storeng, 1992). Наши результаты, напротив, подтверждают позитивное влияние ХГЧ на качество овуляции, но при этом продемонстрирована тенденция к повышению доли мертворожденных особей. Обнаруженное нами достоверное снижение веса тела у потомков после введения ХГЧ на пятый день после их рождения согласуется с обнаруженным ранее снижением веса плодов в конце беременности после воздействия экзогенного ХГЧ (Ertzeid, Storeng, 1992). В исследовании на людях было установлено, что зачатые при помощи вспомогательных репродуктивных технологий дети часто рождаются недоношенными и имеют меньший вес тела по сравнению с естественно зачатыми (Sazonova et al., 2011; Henningsen et al., 2015), что также может быть связано со стимуляцией их матерей в период овуляции препаратом ХГЧ.

Заключение

Таким образом, наше исследование показало, что введение экзогенного ХГЧ самкам мышей перед спариванием их с самцами влияет на овуляцию, улучшая качество ооцитов. Между тем такого рода воздействие оказывает негативное влияние на потомство, отличающееся сниженным весом в первые дни после рождения. Механизмы этих отдаленных эффектов пока остаются непонятными и требуют дальнейшего изучения.

Список литературы / References

Борисова М.А., Моисеенко Д.Ю., Смирнова О.В. Хорионический гонадотропин человека: неизвестное об известном. Физиол. человека. 2017;43(1):97-110. DOI 10.7868/S0131164616060059. [Borisova M.A., Moiseenko D.Yu., Smirnova O.V. Human chorionic gonadotropin: Unknown about known. Human Physiology. 2017; 43(1):93-104. DOI 10.1134/S0362119716060050].

Brinster R.L. Measuring embryonic enzyme activity. In: Daniel J.C. (Ed.) Methods in Mammalian Embryology. San Francisco: Freeman, 1971;215-227.

Cole L.A. Biological functions of hCG and hCG-related molecules. Reprod. Biol. Endocrinol. 2010;8(1):102. DOI 10.1186/1477-7827-8-102.

Delvigne A., Rozenberg S. Epidemiology and prevention of ovarian hyperstimulation syndrome (OHSS): a review. Hum. Reprod. Update. 2002;8:559-577. DOI 10.1093/humupd/8.6.559.

Dinopoulou V., Drakakis P., Kefala S., Kiapekou E., Bletsa R., Anagnostou E., Kallianidis K., Loutradis D. Effect of recombinant-LH and hCG in the absence of FSH on *in vitro* maturation (IVM) fertilization and early embryonic development of mouse germinal vesicle (GV)-stage oocytes. Reprod. Biol. 2016;16(2):138-146. DOI 10.1016/j.repbio.2016.01.004.

Drakakis P., Loutradis D., Beloukas A., Sypsa V., Anastasiadou V., Kalofolias G., Arabatzi H., Kiapekou E., Stefanidis K., Paraskevis D., Makrigiannakis A., Hatzakis A., Antsaklis A. Early hCG addition to rFSH for ovarian stimulation in IVF provides better results and the cDNA copies of the hCG receptor may be an indicator of successful stimulation. Reprod. Biol. Endocrinol. 2009;7:110. DOI 10.1186/1477-7827-7-110.

Ertzeid G., Storeng R. Adverse effects of gonadotrophin treatment on pre- and postimplantation development in mice. J. Reprod. Fertil. 1992;96:649-655. DOI 10.1530/jrf.0.0960649.

Fournier T., Guibourdenche J., Evain-Brion D. Review: hCGs: different sources of production, different glycoforms and functions. Placenta. 2015;36(1):60. DOI 10.1016/j.placenta.2015.02.002.

Hasegawa A., Mochida K., Inoue H., Noda Y., Endo T., Watanabe G., Ogura A. High-yield superovulation in adult mice by anti-inhibin serum treatment combined with estrous cycle synchronization. Biol. Reprod. 2016;94(1):21. DOI 10.1095/biolreprod.115.134023.

Henningsen A.A., Gissler M., Skjaerven R., Bergh C., Tiitinen A., Romundstad L.B., Wennerholm U.B., Lidegaard O., Nyboe Andersen A., Forman J.L., Pinborg A. Trends in perinatal health after assisted reproduction: a Nordic study from the CoNARTaS group. Hum. Reprod. 2015;30:710-716. DOI 10.1093/humrep/deu345.

Hogan B., Beddington R., Costantini F., Lacy E. Manipulating the Mouse Embryo. A Laboratory Manual. 2nd ed., New York; Cold Spring Harbor: Cold Spring Harbor Laboratory, 1994.

Homburg R. Management of infertility and prevention of ovarian hyperstimulation in women with polycystic ovary syndrome. Best Pract. Res. Clin. Obstet. Gynaecol. 2004;18:773-788. DOI 10.1016/j.bpobgyn.2004.05.006.

Kane N., Kelly R., Saunders P.T., Critchley H.O. Proliferation of uterine natural killer cells is induced by human chorionic gonadotropin and mediated via the mannose receptor. Endocrinology. 2009;150: 2882-2888. DOI 10.1210/en.2008-1309.

Keay S.D., Vatish M., Karteris E., Hillhouse E.W., Randeva H.S. The role of hCG in reproductive medicine. Br. J. Obstset. Gynecol. 2004; 111(11):1218. DOI 10.1111/j.1471-0528.2004.00412.x.

Lee M., Ahn J.I., Lee A.R., Ko D.W., Yang W.S., Lee G., Ahn J.Y., Lim J.M. Adverse effect of superovulation treatment on maturation, function and ultrastructural integrity of murine oocytes. Mol. Cells. 2017;40(8):558-566. DOI 10.14348/molcells.2017.0058.

Li S., Wang J., Cheng Y., Zhou D., Yin T., Xu W., Yu N., Yang J. Intrauterine administration of hCG-activated autologous human peripheral blood mononuclear cells (PBMC) promotes live birth rates in frozen/thawed embryo transfer cycles of patients with repeated implantation failure. J. Reprod. Immunol. 2017;119:15-22. DOI 10.1016/j.jri.2016.11.006.

Makrigiannakis A., Vrekoussis T., Zoumakis E., Kalantaridou S.N., Jeschke U. The role of hCG in implantation: A mini-review of molecular and clinical evidence. Int. J. Mol. Sci. 2017;18(6):1305. DOI 10.3390/ijms18061305.

Redina O.E., Amstislavsky S.Ya., Maksimovsky L.F. Induction of superovulation in DD mice at different stages of the oestrous cycle. J. Reprod. Fertil. 1994;(102):263-267. DOI 10.1530/jrf.0.1020263.

Sazonova A., Kallen K., Thurin-Kjellberg A., Wennerholm U.B., Bergh C. Obstetric outcome after *in vitro* fertilization with single or double embryo transfer. Hum. Reprod. 2011;26(2):442-450. DOI 10.1093/humrep/deq325.

Strug M.R., Su R., Young J.E., Dodds W.G., Shavell V.I., Diaz-Gimeno P., Ruiz-Alonso M., Simon C., Lessey B.A., Leach R.E., Fazleabas A.T. Intrauterine human chorionic gonadotropin infusion in oocyte donors promotes endometrial synchrony and induction of early decidual markers for stromal survival: A randomized clinical trial. Hum. Reprod. 2016;31:1552-1561. DOI 10.1093/humrep/dew080.

Takeo T., Nakagata N. Superovulation using the combined administration of inhibin antiserum and equine chorionic gonadotropin increases the number of ovulated oocytes in C57BL/6 female mice. PLoS One. 2015;10(5):e0128330. DOI 10.1371/journal.pone.0128330.

Tarin J.J., Perez-Albala S., Cano A. Stage of the estrous cycle at the time of pregnant mare's serum gonadotropin injection affects the quality of ovulated oocytes in the mouse. Mol. Reprod. Dev. 2002; 61(3):398-405. DOI 10.1002/mrd.10042.

Wang Y., Ock S.A., Chian R.C. Effect of gonadotrophin stimulation on mouse oocyte quality and subsequent embryonic development *in vitro*. Reprod. Biomed. Online. 2006;12(3):304-314. DOI 10.1016/S1472-6483(10)61002-4.

Благодарности. Работа выполнена при поддержке РФФИ, № 19-016-00025 и бюджетного проекта ИЦиГ СО РАН, № 0259-2019-0003.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 20.05.2019. После доработки 13.06.2019. Принята к публикации 13.06.2019.

МЕДИЦИНСКАЯ ГЕНЕТИКА
Оригинальное исследование / Original article
УДК 612.821:575.174.015.3(571.5)

Вавиловский журнал генетики и селекции. 2019;23(8):1011-1019
DOI 10.18699/VJ19.578

Полиморфизм генов липидного обмена в некоторых популяциях Южной и Восточной Сибири

Л.Э. Табиханова[1, 3] ✉, Л.П. Осипова[1, 3], Е.Н. Воронина[2, 3], А.О. Брагин[3], М.Л. Филипенко[2, 3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ e-mail: tabikhan@bionet.nsc.ru

Нарушения липидного обмена лежат в основе патогенеза ряда заболеваний. Коренные народы Сибири отличаются особым генетически обусловленным типом метаболизма, который поддерживает благоприятные липидные показатели крови при повышенном, по сравнению с европеоидами, вкладе в рацион продуктов животного происхождения. При этом коренные сибирские этносы меньше подвержены заболеваниям метаболического спектра. Цель настоящего исследования – изучение частот аллелей генов липидного обмена в коренных популяциях Сибири. Была поставлена задача на примере бурят, телеутов и русских Восточной Сибири выявить этнические особенности в распределении частот полиморфных вариантов генов *CETP* (*G1264A*, rs5882), *LPL* (*C1791G*, rs328) и *FTO* (*C83401A*, rs8050136) и сравнить с данными по мировым популяциям. Выборки восточных (N = 132) и западных (N = 278) бурят, телеутов (N = 120), русских (N = 122) и потомков смешанных браков бурят с русскими (N = 56) генотипированы с помощью ПЦР в режиме реального времени с использованием конкурирующих TaqMan-зондов. В настоящей работе впервые показано, что по частотам полиморфных вариантов генов *CETP* и *FTO* изученные выборки бурят находятся в промежуточном положении между европеоидными группами и популяциями Восточной Азии. Показана статистически значимо меньшая встречаемость у бурят, по сравнению с русскими, аллеля *83401A* гена *FTO*, ассоциированного с ожирением, что согласуется с меньшей подверженностью этого коренного сибирского этноса метаболическим нарушениям. Не выявлено популяционных различий в распределении полиморфного варианта *LPL 1791G*, связанного с дислипидемией, который, вероятно, не вносит заметного вклада в этнические особенности липидного профиля. Промежуточное значение частот аллелей *CETP 1264G* и *FTO 83401A* в группе потомков смешанных браков бурят с русскими позволяет предположить больший риск ассоциированных с этими вариантами метаболических нарушений у лиц смешанного происхождения, чем среди бурят. У телеутов продемонстрирована пониженная частота *FTO 83401A* по сравнению с некоторыми европеоидными группами и выявлена наименьшая частота *CETP 1264G*, ассоциированного с благоприятными липидными показателями крови.
Ключевые слова: буряты; телеуты; русские Восточной Сибири; метисы; ПЦР в режиме реального времени; липидный обмен; генетический полиморфизм; *CETP* (rs5882); *LPL* (rs328); *FTO* (rs8050136).

Для цитирования: Табиханова Л.Э., Осипова Л.П., Воронина Е.Н., Брагин А.О., Филипенко М.Л. Полиморфизм генов липидного обмена в некоторых популяциях Южной и Восточной Сибири. Вавиловский журнал генетики и селекции. 2019;23(8):1011-1019. DOI 10.18699/VJ19.578

Polymorphism of lipid exchange genes in some populations of South and East Siberia

L.E. Tabikhanova[1, 3] ✉, L.P. Osipova[1, 3], E.N. Voronina[2, 3], A.O. Bragin[3], M.L. Filipenko[2, 3]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
✉ e-mail: tabikhan@bionet.nsc.ru

Lipid metabolism disorders underlie the pathogenesis of a number of diseases. Indigenous peoples of Siberia have a specific genetically determined type of metabolism supporting such lipid blood parameters that favor increased consumption (in comparison with Caucasians) of animal products. At the same time, indigenous Siberian ethnic groups are less susceptible to metabolic diseases. The objective of the presented study was to investigate the allele frequencies of lipid metabolism genes in indigenous populations of Siberia to identify the ethnic features of allele frequency distribution for polymorphic variants in genes *CETP* (*G1264A*, rs5882), *LPL* (*C1791G*, rs328) and *FTO* (*C83401A*, rs8050136) in the samples taken from Buryats, Teleuts and Russians of Eastern Siberia, and to compare them with data on world populations. Samples of the Eastern (N = 132) and Western (N = 278) Buryats, Teleuts (N = 120), Russians (N = 122) and persons of mixed Buryat-Russian origin (N = 56) were genotyped by real-time PCR using competitive TaqMan-probes. The obtained results have for the first time demonstrated that the *CEPT* and *FTO* allele frequencies in the Buryat samples are

© Табиханова Л.Э., Осипова Л.П., Воронина Е.Н., Брагин А.О., Филипенко М.Л., 2019

intermediate between European and East Asian populations. Significantly lower incidence of the obesity-assossiated *83401A* allele of the *FTO* gene has been shown in Buryats, compared with Russians, which is consistent with lower susceptibility of the indigenous ethnic groups to metabolic disorders. There have been no population differences in the distribution of *LPL* gene polymorphic variants associated with dyslipidemia, which means they probably do not contribute to the ethnic characteristics of the lipid profile. The intermediate frequencies of the *CETP 1264G* and *FTO 83401A* alleles found in the metis group demonstrate that the metabolic disorders associated with these variants can be rather expected in the descendants of mixed marriages than among Buryats. It has also been demonstrated that Teleuts differ by *FTO 83401A* allele frequency from some of the European groups and have the lowest detected frequency of the allele *CETP 1264G* associated with the favorable lipid blood parameters.
Key words: Buryats; Teleuts; Russians of Eastern Siberia; mixed origin; real-time PCR; lipid metabolism; genetic polymorphism; *CETP* (rs5882); *LPL* (rs328); *FTO* (rs8050136).

For citation: Tabikhanova L.E., Osipova L.P., Voronina E.N., Bragin A.O., Filipenko M.L. Polymorphism of lipid exchange genes in some populations of South and East Siberia. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1011-1019. DOI 10.18699/VJ19.578 (in Russian)

Введение

Коренным народам Сибири, в течение многих веков адаптировавшимся к суровым климато-географическим условиям и преимущественно белково-липидной диете, присущ особый тип метаболизма, с усилением белково-липидного и минимизацией углеводного обмена (Панин, 1978). Ранее показано, что представители сибирских этносов, ведущие традиционный образ жизни, по сравнению с пришлым населением характеризуются благоприятным липидным профилем крови, предотвращающим развитие сердечно-сосудистых и других метаболически обусловленных заболеваний: пониженным уровнем общего холестерина, триглицеридов, липопротеинов низкой и очень низкой плотности (ЛПНП и ЛПОНП), повышенным уровнем липопротеинов высокой плотности (ЛПВП) (Отева и др., 1993; Севостьянова, 2013; Даренская, 2014; Поляков и др., 2015; Цыганкова и др., 2017). В настоящее время урбанизация затрагивает коренное население, меняя условия проживания и хозяйственный уклад, и приводит к прогрессированию у сибирских народов «болезней цивилизации», связанных с нарушениями метаболизма (Овсянникова и др., 2007; Людинина и др., 2014).

Изучение полиморфизма генов метаболического профиля в коренных популяциях Сибири важно как для понимания молекулярно-генетических основ адаптивного потенциала, сложившегося в процессе их становления в конкретных климато-географических условиях проживания и приспособления к особенностям питания, так и для выявления генетических резервов этносов в меняющемся мире (Hsieh et al., 2017; Hallmark et al., 2018). Цель настоящей работы – изучить в сибирских популяциях частоты полиморфных вариантов генов липидного обмена, функциональная значимость которых установлена по результатам исследований мирового населения.

Ген *CETP* кодирует транспортер (белок-переносчик) эфиров холестерина, работа которого заключается в переводе холестерина из антиатерогенных ЛПВП в атерогенные ЛПНП (Кох и др., 2014). При изучении ряда мировых популяций показано, что у носителей аллеля *G* полиморфного локуса *CETP* (*G1264A*, rs5882) снижена активность белка-переносчика и, следовательно, повышен уровень ЛПВП (Thompson et al., 2008), а также понижен уровень триглицеридов по сравнению с гомозиготами *AA* (Hosseini-Esfahani et al., 2019). Обнаружены обратные ассоциации аллеля *1264G* с риском развития атеросклероза (Thompson et al., 2008; Cyrus et al., 2016). В эксперименте, проведенном на близнецах в европейской популяции, было показано, что у гомозигот *GG* меньше индекс массы тела, масса жира и толщина подкожной жировой клетчатки, а при избыточном питании гомозиготы *AA* быстрее набирали вес (Teran-Garcıa et al., 2008). В исследованиях европеоидов, американцев африканского происхождения и жителей Китая выявлено, что генотип *GG* ассоциирован также с долгожительством, снижением риска болезни Альцгеймера и сосудистой деменции, аномалий серого вещества в задних отделах головного мозга у здоровых пожилых людей (Sanders et al., 2010; Yu et al., 2012; Chen et al., 2014; Salminen et al., 2015).

В усвоении жиров тканями важную роль играет фермент липопротеинлипаза (LPL), функция которого заключается в гидролизе триглицеридов плазмы крови до свободных жирных кислот и глицерина и преобразовании, таким образом, липопротеинов очень низкой плотности до липопротеинов низкой плотности и обеспечении энергетических потребностей ткани (Кох и др., 2014). Представляет интерес полиморфный локус *C1791G* (rs328) гена *LPL*: в исследованиях ряда мировых популяций показано, что аллель *G* ассоциирован с благоприятными изменениями липидного состава – снижением уровня триглицеридов и повышением уровня липопротеинов высокой плотности у здоровых испытуемых (Sagoo et al., 2008; Webster et al., 2009; Tang et al., 2010; Shatwan et al., 2016). Показана ассоциация генотипа *CC* с метаболическим синдромом как у жительниц Мексики европеоидного и африканского происхождения, так и у потомков коренного населения страны (Cahua-Pablo et al., 2015). Аллель *1791G* также является протективным по отношению к болезни Альцгеймера (Ren L., Ren X., 2016).

Ген *FTO* (fat mass and obesity associated gene) кодирует 2-оксоглутаратзависимую деметилазу нуклеиновых кислот, которая участвует в центральном контроле гомеостаза энергии (Кудрявцева и др., 2010). С помощью полногеномного анализа ассоциаций (GWAS) продемонстрировано, что полиморфизм гена *FTO* связан с риском ожирения (Babenko et al., 2019). Исследования полиморфного локуса *FTO* (*C83401A*, rs8050136) в различных по происхождению популяциях мира показали, что у носителей аллеля *A* снижена экспрессия гена *FTO* и повышен риск развития ожирения (Park et al., 2013; Chen et al., 2018). Продемонстрирована ассоциация варианта *FTO 83401A* с

сахарным диабетом второго типа в русской популяции и среди населения Восточной Азии (Суплотова и др., 2014; Yang et al., 2017).

Нами не найдено литературных данных о частотах аллелей и генотипов *CETP* (*G1264A*, rs5882), *LPL* (*C1791G*, rs328) и *FTO* (*C83401A*, rs8050136) у коренных сибирских народов. Задачи настоящей работы: изучить встречаемость этих полиморфных вариантов у бурят, одного из самых многочисленных народов Восточной Сибири (более 460 тыс. человек по данным переписи 2010 г.), телеутов, коренной малочисленной народности Южной Сибири (около 2.5 тыс. человек), и русских Восточной Сибири; сопоставить с литературными данными по мировым популяциям; соотнести полученные результаты с распространенностью дислипидемий среди этих этносов. Не меньший интерес представляет распределение вариантов генов липидного обмена у потомков смешанных браков бурят с русскими, поэтому они также были включены в исследование.

Материалы и методы

Генетический материал для данной работы был собран сотрудниками лаборатории популяционной этногенетики ИЦиГ СО РАН (руководитель – к.б.н. Л.П. Осипова) во время экспедиций 2003–2006 гг. Забор крови производился у добровольцев, практически здоровых на момент исследования, с использованием «Информированного согласия» и с одобрения местных органов здравоохранения и этического комитета ИЦиГ СО РАН. Перед сдачей крови каждый испытуемый заполнял специально разработанную демографическую анкету, в которой уточнял национальную принадлежность предков до третьего-четвертого поколения. На основании собранной информации были сформированы пять выборок населения Южной и Восточной Сибири.

Лица бурятской национальности, не имеющие иноэтнических предков, проживающие в селах Алханай и Орловский Агинского Бурятского округа (АБО) Забайкальского края, вошли в группу восточных бурят (N = 132). Этнические буряты, проживающие в селах Эхирит-Булагатского района Усть-Ордынского Бурятского округа (УОБО) Иркутской области (N = 278), вошли в западную выборку. Потомки первого и второго поколения от смешанных браков бурят с русскими составили выборку метисов (N = 56). Выборка телеутов была сформирована из коренного населения Беловского района Кемеровской области (N = 120). Русские старожилы Восточной Сибири, несколько поколений предков которых проживали в селах Забайкальского края и Иркутской области, вошли в пятую выборку (N = 122).

Образцы ДНК были выделены из лейкоцитарных фракций венозной крови с помощью наборов «Биосилика» (Россия). Генотипирование однонуклеотидных замен в генах *CETP* (*G1264A*, rs5882), *LPL* (*C1791G*, rs328), *FTO* (*C83401A*, rs8050136) проводили методом ПЦР в режиме реального времени с использованием конкурирующих TaqMan-зондов, комплементарных полиморфным участкам ДНК. Структуру праймеров и зондов подбирали с помощью последовательностей, доступных в базе данных NCBI (http://www.ncbi.nlm.nih.gov/), с использованием программ UGENE (version 1.14, http://ugene.unipro.ru/) и Oligo Analyzer (version 1.0.3, https://eu.idtdna.com/pages/tools/oligoanalyzer) (табл. 1).

Амплификацию выполняли в объеме 25 мкл, ПЦР-смесь включала в себя праймеры (300 нМ), TaqMan-зонды (100 нМ), TrisHCl (65 мМ, pH 8.9), $(NH_4)_2SO_4$ (16 мМ), $MgCl_2$ (2.5 мМ), Tween-20 (0.05 %), dNTP (0.2 мМ), ДНК (0.5–10 нг) и Taq-ДНК полимеразу (0.5 ед., hot-start, Biosan, IHBFM). Реакцию проводили в следующих условиях: начальная денатурация 3 мин при 96 °C; затем 46 циклов, включающих денатурацию при 96 °C – 5 с, отжиг праймеров и последующую элонгацию при 61 °C – 30 с (каждый шаг сопровождался регистрацией флуоресцентного сигнала на длине волны эмиссии флуорофоров FAM и R6G).

Популяционные частоты аллельных вариантов определяли на основе наблюдаемых частот генотипов. Соответствие эмпирически наблюдаемого распределения частот генотипов теоретически ожидаемому распределению, равновесному по закону Харди–Вайнберга, проверяли с помощью критерия χ^2 (Пирсона) (при $p > 0.05$ равновесие выполняется). Оценку достоверности различий в частотах аллелей между исследованными выборками проводили по критерию χ^2 с применением поправки Йейтса на непрерывность; при $p < 0.017$ (с поправкой на множественность сравнения, 0.017 = 0.05/3) результаты считались статистически значимыми.

Результаты

Распределение генотипов полиморфных локусов генов *CETP* (*G1264A*, rs5882), *LPL* (*C1791G*, rs328) и *FTO* (*C83401A*, rs8050136) в выборках бурят, их метисов, телеутов и русских Восточной Сибири приведено в табл. 2.

Распределение генотипов соответствовало равновесию Харди–Вайнберга для всех полиморфных локусов

Таблица 1. Структуры праймеров и зондов, использованных для генотипирования однонуклеотидных замен в генах *CETP*, *LPL* и *FTO*

Локус	Праймеры	Зонды
CETP (*G1264A*, rs5882)	5′-CCTTGTGGGTCACTTCTGTACT-3′	5′-R6G-CCGAGTCCGTCCAGAGC-BHQ-3′
	5′-CACACTTACGAGACATGACCT-3′	5′-FAM-CCGAGTCCATCCAGAGCT-BHQ-3′
LPL (*C1791G*, rs328)	5′-CCATTTTCTTCCACAGGG-3′	5′-FAM-CACCAGCCTGACTTCTTATTC-BHQ-3′
	5′-AAGCTCAGGATGCCCAGTC-3′	5′-R6G-CACCAGCCTCACTTCTTATTC-BHQ-3′
FTO (*C83401A*, rs8050136)	5′-TCAGTTATGCATTTAGAATGTCTGA-3′	5′-FAM-CTGTGAATTTTGTGATGCACTTG-BHQ-3′
	5′-CACTCCATTTCTGACTGTTACCT-3′	5′-HEX-CTGTGAATTTAGTGATGCACTTG-BHQ-3′

Таблица 2. Распределение генотипов *CETP*, *LPL* и *FTO* в выборках бурят, их метисов, телеутов и русских Восточной Сибири

Популяция			Буряты		Метисы	Телеуты	Русские Восточной Сибири
			восточные	западные			
CETP (*G1264A*, rs5882)	Распределение генотипов, *n* (%)	*AA*	51 (38.6)	125 (44.9)	27 (48.2)	68 (56.7)	58 (47.5)
		AG	63 (47.7)	120 (43.2)	22 (39.3)	39 (32.5)	52 (42.6)
		GG	18 (13.7)	33 (11.9)	7 (12.5)	13 (10.8)	12 (9.8)
	N, человек		132	278	56	120	122
	p (H–W)		0.835	0.611	0.457	0.052	0.945
LPL (*C1791G*, rs328)	Распределение генотипов, *n* (%)	*CC*	114 (86.4)	227 (81.9)	47 (87.0)	96 (82.8)	102 (83.6)
		CG	17 (12.9)	48 (17.4)	7 (13.0)	16 (13.8)	19 (15.6)
		GG	1 (0.7)	2 (0.7)	0	4 (3.4)	1 (0.8)
	N, человек		132	277	54	116	122
	p (H–W)		0.936	0.945	0.926	0.519	0.981
FTO (*C83401A*, rs8050136)	Распределение генотипов, *n* (%)	*CC*	69 (52.3)	158 (56.9)	21 (38.9)	57 (49.1)	47 (38.5)
		CA	56 (42.4)	106 (38.1)	25 (46.3)	46 (39.7)	56 (45.9)
		AA	7 (5.3)	14 (5.0)	8 (14.8)	13 (11.2)	19 (15.6)
	N, человек		132	278	54	116	122
	p (H–W)		0.308	0.481	0.9	0.428	0.733

Примечание. *N* – объем выборки; *n* – численность; *p* (H–W) – значение вероятности отклонения от равновесного распределения Харди–Вайнберга. Здесь и далее: «метисы» – потомки смешанных браков бурят с русскими.

Таблица 3. Частота аллеля *CETP 1264G* в некоторых популяциях (этнических группах) и сравнение популяций (*p*-value)

Популяция/этническая группа	*N*, человек	Частота *CETP 1264G*, %	Сравнение популяций (*p*-value)				
			Восточные буряты	Западные буряты	Метисы	Телеуты	Русские Восточной Сибири
Восточные буряты*	132	37.5		0.296	0.378	**0.017**	0.155
Западные буряты*	278	33.5	0.296		0.859	0.089	0.559
Метисы*	56	32.1	0.378	0.860		0.401	0.947
Телеуты*	120	27.1	**0.017**	0.089	0.401		0.385
Русские Восточной Сибири*	122	31.1	0.155	0.559	0.947	0.385	
Китайцы хань, Пекин, Китай**	103	47.6	**0.035**	***p* < 0.001**	**0.011**	***p* < 0.001**	***p* < 0.001**
Южные китайцы хань, Китай**	105	43.3	0.236	**0.015**	0.066	***p* < 0.001**	**0.009**
Японцы, Токио, Япония**	104	52.9	**0.001**	***p* < 0.001**	***p* < 0.001**	***p* < 0.001**	***p* < 0.001**
Кинь (вьеты) Хошимин, Вьетнам**	99	38.4	0.920	0.247	0.324	**0.016**	0.132
Население штата Юта, потомки выходцев из Северной и Западной Европы**	99	35.9	0.798	0.600	0.581	0.061	0.335
Финны, Финляндия**	99	35.4	0.714	0.691	0.643	0.077	0.393
Англичане и шотландцы**	91	29.7	0.109	0.391	0.762	0.632	0.838
Иберы, Испания**	107	30.8	0.151	0.529	0.909	0.444	0.974
Тосканцы, Италия**	107	33.6	0.430	0.953	0.881	0.160	0.638

Примечание. Здесь и в табл. 4, 5: * собственные данные; ** литературные данные (The 1000 Genomes…, 2012). Жирным шрифтом выделены значения $p < 0.017$, при которых различия считались статистически значимыми.

и выборок. Частоты аллелей *CETP 1264G*, *LPL 1791G* и *FTO 83401A* в изученных выборках, а также в некоторых этнических группах, описанных в литературе (The 1000 Genomes…, 2012), и сравнение популяций (*p*-value) представлены в табл. 3–5.

Аллель *CETP 1264G*, ассоциированный с благоприятными липидными показателями крови, широко распространен в Африке (в среднем 63.8 %), в нигерийской популяции его частота достигает 69 % (The 1000 Genomes…, 2012). Несколько реже он встречается в популяциях

Таблица 4. Частота аллеля *LPL 1791G* в некоторых популяциях (этнических группах) и сравнение популяций (*p*-value)

Популяция/этническая группа	*N*, человек	Частота *LPL 1791G*, %	Сравнение популяций (*p*-value)				
			Восточные буряты	Западные буряты	Метисы	Телеуты	Русские Восточной Сибири
Восточные буряты*	132	7.2		0.362	0.987	0.287	0.674
Западные буряты*	277	9.4	0.362		0.434	0.797	0.820
Метисы*	54	6.5	0.987	0.434		0.351	0.647
Телеуты*	116	10.3	0.287	0.797	0.351		0.633
Русские Восточной Сибири*	122	8.6	0.674	0.820	0.647	0.633	
Китайцы хань, Пекин, Китай**	103	10.2	0.322	0.846	0.377	0.902	0.675
Южные китайцы хань, Китай**	103	12.4	0.080	0.280	0.151	0.588	0.245
Японцы, Токио, Япония**	104	11.1	0.189	0.572	0.263	0.907	0.463
Кинь (вьеты) Хошимин, Вьетнам**	99	13.1	0.050	0.184	0.113	0.451	0.170
Население штата Юта, потомки выходцев из Северной и Западной Европы**	99	12.6	0.072	0.255	0.141	0.551	0.225
Финны, Финляндия**	99	11.6	0.143	0.454	0.218	0.783	0.374
Англичане и шотландцы**	91	11.5	0.163	0.497	0.234	0.817	0.407
Иберы, Испания**	107	16.4	**0.003**	**0.009**	0.021	0.078	**0.016**
Тосканцы, Италия**	107	12.6	0.066	0.239	0.136	0.539	0.214

Таблица 5. Частота аллеля *FTO 83401A* в некоторых популяциях (этнических группах) и сравнение популяций (*p*-value)

Популяция/этническая группа	*N*, человек	Частота *FTO 83401A*, %	Сравнение популяций (*p*-value)				
			Восточные буряты	Западные буряты	Метисы	Телеуты	Русские Восточной Сибири
Восточные буряты*	132	26.5		0.511	**0.038**	0.314	**0.005**
Западные буряты*	278	24.1	0.511		**0.004**	0.055	***p* < 0.001**
Метисы*	54	38.0	0.038	**0.004**		0.249	0.976
Телеуты*	116	31.0	0.314	0.055	0.249		0.105
Русские Восточной Сибири*	122	38.5	**0.005**	***p* < 0.001**	0.976	0.105	
Китайцы хань, Пекин, Китай**	103	15.0	**0.004**	**0.009**	***p* < 0.001**	***p* < 0.001**	***p* < 0.001**
Южные китайцы хань, Китай**	105	13.8	**0.001**	**0.003**	***p* < 0.001**	***p* < 0.001**	***p* < 0.001**
Японцы, Токио, Япония**	104	17.3	0.023	0.056	***p* < 0.001**	**0.001**	***p* < 0.001**
Кинь (вьеты) Хошимин, Вьетнам**	99	21.7	0.281	0.557	**0.003**	0.039	***p* < 0.001**
Население штата Юта, потомки выходцев из Северной и Западной Европы**	99	44.4	***p* < 0.001**	***p* < 0.001**	0.336	**0.006**	0.247
Финны, Финляндия**	99	38.9	**0.006**	***p* < 0.001**	0.975	0.106	0.990
Англичане и шотландцы**	91	39.6	**0.005**	***p* < 0.001**	0.884	0.085	0.897
Иберы, Испания**	107	37.4	**0.014**	***p* < 0.001**	0.987	0.185	0.884
Тосканцы, Италия**	107	46.3	***p* < 0.001**	***p* < 0.001**	0.194	**0.001**	0.111

Южной (в среднем 45 %) и Восточной Азии (43.7 %), а также среди населения Центральной и Южной Америки (40.1 %). Европейские популяции характеризуются пониженной частотой этого протективного к атеросклерозу аллеля (33.1 %).

Группа русских Восточной Сибири по частоте аллеля *CETP 1264G* соответствует другим описанным в литературе европеоидным популяциям (см. табл. 3). Наши исследования выявили повышенную частоту этого аллеля среди бурят, по сравнению с русскими, но статистически

значимо меньшую, чем у ряда популяций Восточной Азии. Группа потомков смешанных браков бурят с русскими находится в промежуточном положении между родительскими популяциями. У телеутов показана наименьшая частота *CETP 1264G* из изученных нами, а также из всех описанных ранее популяций. Различия телеутов с восточными бурятами и рядом популяций Восточной Азии являются статистически значимыми.

Встречаемость аллеля *LPL 1791G*, протективного к ряду метаболических нарушений, относительно невысока в популяциях человека (The 1000 Genomes…, 2012). Так, в африканских популяциях она варьирует от 2.5 до 10 %, среди коренного населения Америки составляет 2.4–8.7 %, в популяциях Южной Азии – 5.2–10.8 %. В популяциях Восточной Азии встречаемость этого аллеля 10.2–14 %, у европейцев – 11.5–16.4 %.

Между исследованными нами выборками бурят, телеутов и русских значимых отличий по встречаемости аллеля *LPL 1791G* выявлено не было (см. табл. 4). Во всех группах она несколько снижена в сравнении с данными литературы по популяциям Восточной Азии и европейцам. Однако статистически значимые отличия найдены только у бурят по сравнению с европеоидной популяцией иберийцев, в которой встречаемость *1791G* максимальна.

Аллель *FTO 83401A*, ассоциированный с высоким индексом массы тела, широко распространен в африканских (38.1–46.8 %) и европеоидных (37.4–46.3 %) популяциях (The 1000 Genomes…, 2012), реже – в коренных популяциях Центральной и Южной Америки (21.9–34.6 %) и Южной Азии (24.3–33.8 %). В популяциях Восточной Азии – китайских, японских, вьетнамских – встречаемость *83401A* еще ниже (13.8–21.7 %).

Частота аллеля *FTO 83401A* в выборках восточных и западных бурят статистически значимо ниже, чем в выборках русских Восточной Сибири и в европеоидных группах, описанных в литературе (см. табл. 5). Среди телеутов аллель *83401A* также встречается реже, чем у русских, статистически значимые различия выявлены между выборками телеутов и европеоидной выборкой населения штата Юта, а также группой тосканцев Италии. С другой стороны, частоты варианта *FTO 83401A* в сибирских выборках значительно повышены по сравнению с рядом групп Восточной Азии. Ранее, по частотам других генов, уже было показано промежуточное положение этих коренных сибирских популяций между европеоидами и восточными азиатами (Табиханова и др., 2018а, б).

Обсуждение

Настоящее исследование выполнено в рамках актуального направления, задачей которого является изучение особенностей популяционно-генетической структуры коренных народностей Сибири в медико-биологическом и генно-географическом контексте. Впервые проведено изучение частот функционально значимых аллелей *CETP 1264G*, *LPL 1791G* и *FTO 83401A* в популяциях бурят, телеутов и русских Восточной Сибири, а также потомков от смешанных браков бурят с русскими.

Показано, что выявленные частоты согласуются с общим географическим градиентом распределения полиморфных вариантов *CETP* (*G1264A*, rs5882) и *FTO* (*C83401A*, rs8050136). Группа русских Восточной Сибири соответствует другим европеоидным популяциям, а сибирские народы находятся в промежуточном положении между европеоидами и популяциями Восточной Азии. Исключение составляет частота аллеля *CETP 1264G* у телеутов, наименьшая из известных популяционных частот. Возможно, на становление данной популяции оказали влияние факторы микроэволюции (дрейф генов, прохождение через «бутылочное горлышко»). Для окончательных выводов необходимо дальнейшее изучение этой малочисленной народности, направленное на расширение выборки и охват исследованием других территориальных групп телеутов.

Что касается полиморфизма гена *LPL* (*C1791G*, rs328), то он не демонстрирует выраженного градиента частот на карте мира. Не найдено закономерностей и в распределении аллеля *LPL 1791G* в наших выборках.

Отмечено, что в рационе бурят, по сравнению с русскими, выше удельный вес богатых холестерином животных продуктов, и соотношение ингредиентов с превышением доли животной пищи является характерной особенностью их питания (Баирова и др., 2013). При этом негативные изменения липидного профиля (повышение уровня триглицеридов, общего холестерина ЛПНП и ЛПОНП) в климактерическом периоде у буряток менее выражены, чем у представительниц русской этнической группы (Семенова и др., 2018). Показано, что уровень холестерина у больных с инфарктом миокарда и сахарным диабетом второго типа в бурятской популяции был ниже аналогичного показателя у пациентов европеоидной популяции (Бардымова и др., 2015). По-видимому, адаптация к условиям сибирского региона с его резко континентальным климатом, к животноводческому типу хозяйства и потреблению калорийной пищи с высоким содержанием белка и жира оказала влияние на частоты полиморфных вариантов генов липидного обмена этого коренного сибирского этноса.

Меньшая распространенность нарушений липидного обмена среди бурят согласуется с повышенной популяционной частотой *CETP 1264G*, ассоциированного с благоприятными липидными показателями крови, и пониженной встречаемостью *FTO 83401A*, ассоциированного с ожирением. Не выявлено популяционных различий по частоте аллеля *LPL 1791G*, что может свидетельствовать об отсутствии его вклада в этнические особенности липидного профиля. Дальнейшее изучение генов липидного обмена в этих популяциях поможет более точно установить природу этнических различий.

Промежуточные значения частот *CETP 1264G* и *FTO 83401A* в группе метисов позволяют предположить больший риск ассоциированных с этими вариантами метаболических нарушений у потомков смешанных браков. Однако среди метисов наблюдается наименьшая встречаемость *LPL 1791G*; возможно, это связано с малочисленностью выборки.

Литературных данных об особенностях липидного спектра у телеутов нами не найдено, но в исследованиях шорцев, другого коренного малочисленного народа Кемеровской области, обнаружена характерная для сибирских этносов тенденция к меньшей распространенности дислипидемий, чем у пришлого европеоидного населения

(Цыганкова и др., 2017). В то же время представители телеутского этноса, проживающие на урбанизированной территории Кузбасса в экологически неблагоприятной зоне (межгорной котловине), демонстрируют в последние десятилетия повышение частоты сердечно-сосудистых заболеваний (Овсянникова и др., 2007). В настоящей работе у телеутов, как и у бурят, продемонстрирована пониженная частота *FTO 83401A*, ассоциированного с ожирением, однако генетические отличия этого коренного сибирского этноса от европеоидов менее выражены, как было показано нами ранее и для других генов (Табиханова и др., 2018а, б). Низкая популяционная частота *CETP 1264G*, ассоциированного с благоприятными липидными показателями крови, может свидетельствовать о повышенном риске дислипидемий в рассматриваемой территориальной группе телеутов.

Заключение

В настоящей работе впервые изучены этнические особенности в распределении частот полиморфных вариантов генов *CETP* (*G1264A*, rs5882), *LPL* (*C1791G*, rs328) и *FTO* (*C83401A*, rs8050136) в сибирских популяциях бурят, телеутов, русских и у потомков смешанных браков бурят с русскими.

Показано, что по частотам *CETP 1264G* и *FTO 83401A* выборки бурят находятся в промежуточном положении между европеоидными группами и популяциями Восточной Азии. По сравнению с русскими у бурят статистически значимо меньшая встречаемость аллеля *83401A* гена *FTO*, ассоциированного с ожирением. Это согласуется с описанной в литературе меньшей, по сравнению с европеоидным населением, подверженностью бурят метаболическим нарушениям. Промежуточное значение частот аллелей *CETP 1264G* и *FTO 83401A* в группе метисов позволяет предположить больший риск ассоциированных с этими вариантами метаболических нарушений у потомков смешанных браков, чем среди бурят.

У телеутов также продемонстрирована пониженная частота *FTO 83401A*, однако генетические отличия этого коренного сибирского этноса от европеоидов менее выражены; показана наименьшая частота *CETP 1264G* из всех изученных популяций. Результаты нашего исследования могут свидетельствовать о повышенном риске дислипидемий в данной территориальной группе телеутов.

Не выявлено популяционных различий в распределении варианта *LPL 1791G*, связанного с дислипидемией, который, вероятно, не вносит вклад в особенности липидного профиля этих сибирских народов.

Список литературы / References

Баирова Т.А., Долгих В.В., Колесникова Л.И., Первушина О.А. Нутрициогенетика и факторы риска сердечно-сосудистой патологии: ассоциативные исследования в популяциях Восточной Сибири. Бюл. ВСНЦ СО РАМН. 2013;4(92):87-92.
[Bairova T.A., Dolgikh V.V., Kolesnikova L.I., Pervushina O.A. Nutriciogenetics and risk factors of cardiovascular disease: associated research in Eastern Siberia populations. Byulleten' VSNTs SO RAMN = Bulletin ESSC SB RAMS. 2013;4(92):87-92. (in Russian)]

Бардымова Т.П., Протасов К.В., Донирова О.С., Цыреторова С.С., Березина М.В. Этнические особенности инфаркта миокарда и сахарного диабета у больных бурятской популяции. Сборник тезисов VII Всерос. диабетологического конгр. «Сахарный диабет в XXI веке – время объединения усилий». Москва, 24–27 февр., 2015. М., 2015;121.
[Bardymova T.P., Protasov K.V., Donirova O.S., Tsyretorova S.S., Berezina M.V. Ethnic features of myocardial infarction and diabetes mellitus in patients of the Buryat population. Abstracts from the VII Russian Diabetology Congress "Diabetes mellitus in the 21st century: time to unite efforts", 2015 Feb. 24–27. Moscow, 2015;121. (in Russian)]

Даренская М.А. Особенности метаболических реакций у коренного и пришлого населения Севера и Сибири. Бюл. ВСНЦ СО РАМН. 2014;2(96):97-103.
[Darenskaya M.A. Features of metabolic reactions in indigenous and migrant populations of Northern Russia and Siberia. Byulleten' VSNTs SO RAMN = Bulletin ESSC SB RAMS. 2014;2(96):97-103. (in Russian)]

Кох Н.В., Лифшиц Г.И., Воронина Е.Н. Возможности анализа полиморфизма генов липидного обмена для выявления факторов риска атеросклероза. Рос. кардиол. журн. 2014;10(114):53-57. DOI 10.15829/1560-4071-2014-10-53-57.
[Koch N.V., Lifschitz G.I., Voronina E.N. Approaches to the lipid metabolism genes polymorphism analysis in screening for atherosclerosis risk factors. Rossiyskiy Kardiologicheskiy Zhurnal = Russian Journal of Cardiology. 2014;10(114):53-57. DOI 10.15829/1560-4071-2014-10-53-57. (in Russian)]

Кудрявцева Е.А., Воронина Е.Н., Лифшиц Г.И., Крапивина Н.А., Цветовская Г.А., Филипенко М.Л. Отсутствие влияния полиморфных локусов генов *INSIG2*, *FTO*, *GNB3* на степень выраженности ожирения у больных метаболическим синдромом. Вестн. Новосиб. гос. ун-та. Сер.: Биология, клиническая медицина. 2010;8(3):32-39.
[Kudryavtseva E.A., Voronina E.N., Lifshits G.I., Krapivina N.A., Tsvetovskaya G.A., Filipenko M.L. No influence of polymorphic variants of genes *INSIG2*, *FTO*, *GNB3* on the severity of obesity in patients with metabolic syndrome. Vestnik Novosibirskogo Gosudarstvennogo Universiteta. Seriya: Biologiya, klinicheskaya meditsina = Bulleten of Novosibirsk State University. Ser.: Biology, clinical medicine. 2010;8(3):32-39. (in Russian)]

Людинина А.Ю., Потолицына Н.Н., Солонин Ю.Г., Осадчук Л.В., Гуторова Н.В., Петрова П.Г., Троев И.П., Остобунаев В.В., Бойко Е.Р. Липидный профиль у мужчин коми и якутской этнической принадлежности с избыточной массой тела и ожирением. Экология человека. 2014;1:13-19.
[Lyudinina A.Yu., Potolitsyna N.N., Solonin Yu.G., Osadchuk L.V., Gutorova N.V., Petrova P.G., Troev I.P., Ostobunaev V.V., Boyko E.R. Lipid profile in men of Komi and Yakut ethnic groups with overweight and obesity. Ekologiya Cheloveka = Human Ecology. 2014;1:13-19. (in Russian)]

Овсянникова О.В., Подхомутников В.М., Колбаско А.В., Лузина Ф.А., Гуськова Е.В. Сердечно-сосудистая патология у коренных жителей Кузбасса – телеутов. Рос. кардиол. журн. 2007; 12(6):59-62.
[Ovsyannikova O.V., Podkhomutnikov V.M., Kolbasko A.V., Luzina F.A., Gus'kova E.V. Cardiovascular disease in rural Kuzbass aborigines – Teleut. Rossiyskiy Kardiologicheskiy Zhurnal = Russian Journal of Cardiology. 2007;12(6):59-62. (in Russian)]

Отева Е.А., Масленников А.В., Николаева А.А., Осипова Л.П., Евсеева О.Л., Шеренталь И.С., Филимонова Т.А., Пиковская Н.В. Особенности липидного состава сыворотки крови у северных селькупов. Терапевт. архив. 1993;65(1):21-24.
[Oteva E.A., Maslennikov A.V., Nikolaeva A.A., Osipova L.P., Evseeva O.L., Sherental' I.S., Filimonova T.A., Pikovskaja N.V. Features of serum lipid composition in northern Selkups. Terapevticheskiy Arkhiv = Therapeutic Archive. 1993;65(1):21-24. (in Russian)]

Панин Л.Е. Энергетические аспекты адаптации. Л., 1978.
[Panin L.E. Energy Aspects of Adaptation. Leningrad, 1978. (in Russian)]

Поляков Л.М., Розуменко А.А., Осипова Л.П., Куницын В.Г., Гольцова Т.В. Липидный спектр сыворотки крови у представителей коренного и пришлого населения Ямало-Ненецкого автономного округа. Сиб. науч. мед. журн. 2015;35(6):66-69.
[Polyakov L.M., Rozumenko A.A., Osipova L.P., Kunitsyn V.G., Goltsova T.V. Serum lipid spectrum of indigenous and alien population of Yamalo-Nenets Autonomous Okrug. Sibirskiy Nauchnyy Meditsinskiy Zhurnal = The Siberian Scientific Medical Journal. 2015;35(6):66-69. (in Russian)]

Севостьянова Е.В. Особенности липидного и углеводного метаболизма человека на Севере (литературный обзор). Бюл. сиб. медицины. 2013;12(1):93-100.
[Sevostyanova Ye.V. Some features of human lipid and carbohydrate metabolism in the North. Byulleten' Sibirskoy Meditsiny = Bulletin of Siberian Medicine. 2013;12(1):93-100. (in Russian)]

Семёнова Н.В., Мадаева И.М., Даренская М.А., Гаврилова О.А., Жамбалова Р.М., Колесникова Л.И. Липидный профиль у женщин двух этнических групп в климактерическом периоде. Acta Biomedica Scientifica. 2018;3(3):93-98. DOI 10.29413/ABS.2018-3.3.14.
[Semenova N.V., Madaeva I.M., Darenskaya M.A., Gavrilova O.A., Zhambalova R.M., Kolesnikova L.I. Lipid profile in menopausal women of two ethnic groups. Acta Biomedica Scientifica. 2018; 3(3):93-98. DOI 10.29413/ABS.2018-3.3.14. (in Russian)]

Суплотова Л.А., Вахромеева К.А., Бельчикова Л.Н., Носиков В.В. Поиск ассоциаций полиморфных генетических маркеров с сахарным диабетом 2 типа в русской популяции. Мед. наука и образование Урала. 2014;4:51-57.
[Suplotova L.A., Vakhromeeva K.A., Belchikova L.N., Nosikov V.V. Evalaution of association of genetic markers with type 2 diabetes in Russian population. Meditsinskaya Nauka i Obrazovanie Urala = Medical Science and Education of Ural. 2014;4:51-57. (in Russian)]

Табиханова Л.Э., Осипова Л.П., Чуркина Т.В., Воронина Е.Н., Филипенко М.Л. Полиморфизм генов *CYP1A1* и *CYP2D6* в популяциях бурят, телеутов и у русских Восточной Сибири. Вавиловский журнал генетики и селекции. 2018а;22(2):205-211. DOI 10.18699/VJ18.348.
[Tabikhanova L.E., Osipova L.P., Churkina T.V., Voronina E.N., Filipenko M.L. Genetic polymorphism of *CYP1A1* and *CYP2D6* in populations of Buryats, Teleuts and Russians of Eastern Siberia. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(2):205-211. DOI 10.18699/VJ18.348. (in Russian)]

Табиханова Л.Э., Осипова Л.П., Чуркина Т.В., Воронина Е.Н., Филипенко М.Л. Полиморфизм генов *F2*, *F5* и *VKORC1* в популяциях бурят, телеутов и русских жителей Восточной Сибири. Молекуляр. медицина. 2018б;16(3):31-36. DOI 10.29296/24999490-2018-03-06.
[Tabikhanova L.E., Osipova L.P., Churkina T.V., Voronina E.N., Filipenko M.L. Genetic polymorphism of *F2*, *F5* and *VKORC1* in populations of Buryats, Teleuts and Russians of Eastern Siberia. Molekulyarnaya Meditsina = Molecular Medicine. 2018;16(3):31-36. DOI 10.29296/24999490-2018-03-06. (in Russian)]

Цыганкова Д.П., Мулерова Т.А., Огарков М.Ю., Михалина Е.В., Саарела Е.Ю., Казачек Я.В., Кузьмина А.А., Барбараш О.Л. Показатели липидного обмена у жителей Горной Шории: этнические особенности и влияние условий проживания. Атеросклероз и дислипидемии. 2017;1(26):68-76.
[Tsygankova D.P., Mulerova T.A., Ogarkov M.Yu., Mikhalina E.V., Saarela E.Yu., Kazachek Ya.V., Kuzmina A.A., Barbarash O.L. Indicators of lipid metabolism in the inhabitants of Mountain Shoria: ethnic peculiarities and the impact of living conditions. Ateroskleroz i Dislipidemii = The Journal of Atherosclerosis and Dyslipidemias. 2017;1(26):68-76. (in Russian)]

Babenko V.N., Babenko R.O., Gamieldien G., Markel A.L. FTO haplotyping underlines high obesity risk for European populations. BMC Med. Genomics. 2019;12(2):46. DOI 10.1186/s12920-019-0491-x.

Cahua-Pablo J.A., Cruz M., Méndez-Palacios A., Antúnez-Ortiz D.L., Vences-Velázquez A., Alarcón-Romero L.C., Parra E.J., Tello-Flores V.A., Leyva-Vázquez M.A., Valladares-Salgado A., Pérez-Macedonio C.P., Flores-Alfaro E. Polymorphisms in the *LPL* and *CETP* genes and haplotype in the *ESR1* gene are associated with metabolic syndrome in women from Southwestern Mexico. Int. J. Mol. Sci. 2015;16:21539-21554. DOI 10.3390/ijms160921539.

Chen B., Li Z., Chen J., Ji J., Shen J., Xu Y., Zhao Y., Liu D., Shen Y., Zhang W., Shen J., Wang Y., Shi Y. Association of fat mass and obesity-associated and retinitis pigmentosa guanosine triphosphatase (GTPase) regulator-interacting protein-1 like polymorphisms with body mass index in Chinese women. Endocr. J. 2018;65(7):783-791. DOI 10.1507/endocrj.EJ17-0554.

Chen J.J., Li Y.M., Zou W.Y., Fu J.L. Relationships between *CETP* genetic polymorphisms and Alzheimer's disease risk: a meta-analysis. DNA Cell Biol. 2014;33(11):807-815. DOI 10.1089/dna.2013.2265.

Cyrus C., Vatte C., Al-Nafie A., Chathoth S., Al-Ali R., Al-Shehri A., Akhtar M.S., Almansori M., Al-Muhanna F., Keating B., Al-Ali A. The impact of common polymorphisms in *CETP* and *ABCA1* genes with the risk of coronary artery disease in Saudi Arabians. Hum. Genomics. 2016;10(8). DOI 10.1186/s40246-016-0065-3.

Hallmark B., Karafet T.M., Hsieh P.H., Osipova L.P., Watkins J.C., Hammer M.F. Genomic evidence of local adaptation to climate and diet in indigenous Siberians. Mol. Biol. Evol. 2018;36(2):315-327. DOI 10.1093/molbev/msy211.

Hosseini-Esfahani F., Esfandiar Z., Mirmiran P., Daneshpour M.S., Ghanbarian A., Azizi F. The interaction of cholesteryl ester transfer protein gene variations and diet on changes in serum lipid profiles. Eur. J. Clin. Nutr. 2019; 73:1291-1298. DOI 10.1038/s41430-019-0397-x.

Hsieh P.H., Hallmark B., Watkins J.C., Karafet T.M., Osipova L.P., Gutenkunst R.N., Hammer M.F. Exome sequencing provides evidence of polygenic adaptation to a fat-rich animal diet in indigenous Siberian populations. Mol. Biol. Evol. 2017;34(11):2913-2926. DOI 10.1093/molbev/msx226.

Park S.L., Cheng I., Pendergrass S.A., Kucharska-Newton A.M., Lim U., Ambite J.L., Caberto C.P., Monroe K.R., Schumacher F., Hindorff L.A., Oetjens M.T., Wilson S., Goodloe R.J., Love S.A., Henderson B.E., Kolonel L.N., Haiman C.A., Crawford D.C., North K.E., Heiss G., Ritchie M.D., Wilkens L.R., Le Marchand L. Association of the *FTO* obesity risk variant rs8050136 with percentage of energy intake from fat in multiple racial/ethnic populations. Am. J. Epidemiol. 2013;178(5):780-790. DOI 10.1093/aje/kwt028.

Ren L., Ren X. Meta-analyses of four polymorphisms of lipoprotein lipase associated with the risk of Alzheimer's disease. Neurosci. Lett. 2016;619:73-78. DOI 10.1016/j.neulet.2016.03.021.

Sagoo G.S., Tatt I., Salanti G., Butterworth A.S., Sarwar N., Maarle M., Jukema J.W., Wiman B., Kastelein J.J., Bennet A.M., Faire U., Danesh J., Higgins J.P. Seven lipoprotein lipase gene polymorphisms, lipid fractions, and coronary disease: a huge association review and meta-analysis. Am. J. Epidemiol. 2008;168(11):1233-1246. DOI 10.1093/aje/kwn235.

Salminen L.E., Schofield P.R., Pierce K.D., Luo X., Zhao Y., Laidlaw D.H., Cabeen R.P., Conturo T.E., Lane E.M., Heaps J.M., Bolzenius J.D., Baker L.M., Cooley S.A., Scott S., Cagle L.M., Paul R.H. Genetic markers of cholesterol transport and gray matter diffusion: a preliminary study of the *CETP* I405V polymorphism. J. Neural. Transm. 2015;122:1581-1592. DOI 10.1007/s00702-015-1434-0.

Sanders A.E., Wang C., Katz M., Derby C.A., Barzilai N., Ozelius L., Lipton R.B. Association of a functional polymorphism in the cholesteryl ester transfer protein (*CETP*) gene with memory decline and incidence of dementia. JAMA. 2010;303(2):150-158. DOI 10.1001/jama.2009.1988.

Shatwan I.M., Minihane A.M., Williams C.M., Lovegrove J.A., Jackson K.G., Vimaleswaran K.S. Impact of lipoprotein lipase gene polymorphism, *S447X*, on postprandial triacylglycerol and glucose response to sequential meal ingestion. Int. J. Mol. Sci. 2016;17(3):397. DOI 10.3390/ijms17030397.

Tang W., Apostol G., Schreiner P.J., Jacobs D.R., Boerwinkle E., Fornage M. Associations of lipoprotein lipase gene polymorphisms with longitudinal plasma lipid trends in young adults: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. Circ. Cardiovasc. Genet. 2010;3(2):179-186. DOI 10.1161/CIRCGENETICS.109.913426.

Teran-Garcıa M., Despres J.P., Tremblay A., Bouchard C. Effects of cholesterol ester transfer protein (CETP) gene on adiposity in response to long-term overfeeding. Atherosclerosis. 2008;196:455-460.

The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65. DOI 10.1038/nature11632.

Thompson A., Angelantonio E., Sarwar N., Erqou S., Saleheen D., Dullaart R., Keavney B., Ye Z., Danesh J. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. JAMA. 2008;299(23):2777-2788. DOI 10.1001/jama.299.23.2777.

Webster R.J., Warrington N.M., Weedon M.N., Hattersley A.T., McCaskie P.A., Beilby J.P., Palmer L.J., Frayling T.M. The association of common genetic variants in the *APOA5*, *LPL* and *GCK* genes with longitudinal changes in metabolic and cardiovascular traits. Diabetologia. 2009;52:106-114. DOI 10.1007/s00125-008-1175-9.

Yang Y., Liu B., Xia W., Yan J., Liu H.Y., Hu L., Liu S.M. *FTO* genotype and type 2 diabetes mellitus: spatial analysis and meta-analysis of 62 case-control studies from different regions. Genes. 2017;8(2):70. DOI 10.3390/genes8020070.

Yu L., Shulman J.M., Chibnik L., Leurgans S., Schneider J.A., Jager P.L., Bennett D.A. The *CETP I405V* polymorphism is associated with an increased risk of Alzheimer's disease. Aging Cell. 2012;11(2):228-233. DOI 10.1111/j.1474-9726.2011.00777.x.

ORCID ID
L.E. Tabikhanova orcid.org/0000-0002-5547-8189
L.P. Osipova orcid.org/0000-0001-7602-1156

Благодарности. Исследование выполнено за счет гранта Российского научного фонда, проект № 19-15-00219 «Исследование генетических детерминант мультифакторных заболеваний у коренных народов Сибири». Авторы выражают благодарность к.б.н. Т.М. Карафет, к.б.н. Д.В. Личман, Н.А. Вавиловой, Н.А. Молетотовой, к.б.н. С.С. Сангаеву и А.О. Лихачёвой за участие в экспедициях и техническую помощь.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 29.05.2019. После доработки 02.10.2019. Принята к публикации 07.10.2019.

МЕДИЦИНСКАЯ ГЕНЕТИКА
Оригинальное исследование / Original article
УДК 577.125.8+577.323.36

Вавиловский журнал генетики и селекции. 2019;23(8):1020-1025
DOI 10.18699/VJ19.579

Изменения липидного обмена мыши при одновременном воздействии антисмысловыми олигонуклеотидными производными к мРНК генов *anoB*, *PCSK9* и *anoCIII*

С.И. Ошевский[1] ✉, Ю.И. Рагино[1], Е.В. Каштанова[1], Я.В. Полонская[1], Е.М. Стахнева[1], В.П. Николин[1], Н.А. Попова[1, 2], Н.А. Колчанов[1], М.И. Воевода[1, 2, 3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Федеральный исследовательский центр фундаментальной и трансляционной медицины, Новосибирск, Россия
✉ e-mail: oshevski@bionet.nsc.ru

Для пациентов с наследственной гиперхолестеринемией и/или не переносящих статины остается актуальной разработка новых средств, снижающих уровень «плохого» холестерина, например на основе антисмысловых олигонуклеотидных производных (АСО). Целью исследования была оценка изменения липидного обмена в вариантах совместного воздействия АСО, адресованными на мРНК его важнейших генов: *anoB*, *PCSK9* и *anoCIII*. Использованы: самки мыши линии C57BL/6J; защищенные от действия нуклеаз олигонуклеотидные производные длиной 13 и 20 нуклеотидов; стандартные методы количественного определения липопротеидов – ХС ЛПВП, ХС не-ЛПВП, общего холестерина (ХС) и аланинаминотрансферазы (АЛТ) в сыворотке крови. АСО в комбинациях: 1) АСО к апоB, 2) АСО к апоCIII, 3) АСО к апоB и АСО к PCSK9, 4) АСО к апоB, АСО к PCSK9 и АСО к апоCIII, 5) 3АСО к апоB, АСО к PCSK9 и 2АСО к апоCIII (цифры – это кратность количества вводимого препарата), 6) АСО к PCSK9 и (АСО к апоCIII – только в четвертое введение) вводили в хвостовую вену мышам по четыре раза. Вследствие трехкратного еженедельного введения в этих комбинациях уровень ХС не-ЛПВП относительно контроля был ниже на 25, 16, 35, 47, 60 и 7 % соответственно, а соотношение ХС ЛПВП/ХС не-ЛПВП выше в 1.8, 1.5, 1.9, 2.4, 3.1 и 1.24 раза. При последующем введении АСО с одновременным переходом на жировую диету через одну неделю в случае комбинаций 1–6 АСО уровень ХС не-ЛПВП был на 28, 2, 28, 70, 33 и 49 % ниже, чем в контроле, а соотношение ХС ЛПВП/ХС не-ЛПВП лучше в 1.5, 1.1, 2, 3.7, 1.9 и 2 раза соответственно. Концентрация АЛТ при всех комбинациях АСО оставалась в пределах уровня у контрольных животных, что указывает на отсутствие гепатотоксического эффекта. Для наибольшей эффективности антисмысловых олигонуклеотидных производных в комбинациях важно учитывать необходимость подбора концентраций, в том числе и в сочетаниях, а также порядка и периодов введения. Впервые предлагается временной комбинационно-концентрационный антисмысловой подход.
Ключевые слова: антисмысловые олигонуклеотидные производные; регуляция липидного обмена.

Для цитирования: Ошевский С.И., Рагино Ю.И., Каштанова Е.В., Полонская Я.В., Стахнева Е.М., Николин В.П., Попова Н.А., Колчанов Н.А., Воевода М.И. Изменения липидного обмена мыши при одновременном воздействии антисмысловыми олигонуклеотидными производными к мРНК генов *anoB*, *PCSK9* и *anoCIII*. Вавиловский журнал генетики и селекции. 2019;23(8):1020-1025. DOI 10.18699/VJ19.579

Changes induced in mouse lipid metabolism by simultaneous impact of antisense oligonucleotide derivatives to *apoB*, *PCSK9*, and *apoCIII* mRNAs

S.I. Oshevski[1] ✉, Y.I. Ragino[1], E.V. Kashtanova[1], Y.V. Polonskaya[1], E.M. Stakhneva[1], V.P. Nikolin[1], N.A. Popova[1, 2], N.A. Kolchanov[1], M.I. Voevoda[1, 2, 3]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Federal Research Center of Fundamental and Translational Medicine, Novosibirsk, Russia
✉ e-mail: oshevski@bionet.nsc.ru

Development of new drugs able to decrease the level of “bad” cholesterol, in particular, based on antisense oligonucleotide derivatives (ASOs), remains relevant for the patients with familial hypercholesterolemia and/or intolerant to statins. The goal of the work was to assess the changes in the lipid metabolism caused by variants of joint impact of the ASOs targeted to the mRNAs of its key genes: *apoB*, *PCSK9*, and *apoCIII*. Female C57BL/6J mice; nuclease-protected 13- and 20-nucleotide ASOs, and standard protocols for quantification of lipoproteins (HDL CHL, non-HDL CHL, and total CHL) and ALT in the blood serum were used in the work. The following combi-

© Ошевский С.И., Рагино Ю.И., Каштанова Е.В., Полонская Я.В., Стахнева Е.М., Николин В.П., Попова Н.А., Колчанов Н.А., Воевода М.И., 2019

nations of ASOs were four times injected to the mouse caudal vein: 1) ASO to apoB, 2) ASO to apoCIII, 3) ASO to apoB and ASO to PCSK9, 4) ASO to apoB, ASO to PCSK9, and ASO to apoCIII, 5) ASO to apoB (three doses), ASO to PCSK9, and ASO to apoCIII (two doses), 6) ASO to PCSK9 and (ASO to apoCIII – only in the fourth administration). Triple weekly administration of these ASO combinations resulted in a decrease in non-HDL CHL by 25, 16, 35, 47, 60, and 7 %, respectively, as compared with the control and 1.8-, 1.5-, 1.9-, 2.4-, 3.1, and 1.24-fold higher HDL CHL/non-HDL CHL ratio. The subsequent ASO injection with concurrent switching to a high-fat diet after 1 week resulted in a decrease in the non-HDL CHL by 28, 2, 28, 70, 33, and 49 % for ASOs (1–6), respectively, as compared with the control; the HDL CHL/non-HDL CHL ratio was 1.5-, 1.1-, 2-, 3.7-, 1.9-, and 2-fold better. The ALT concentration for all ASO combinations remained within the norm for the control animals, demonstrating the absence of any hepatotoxic effect. The best efficiency of ASOs requires selection of concentrations for single ASOs and their combinations as well as of the order and timing of administration. Thus, a new antisense approach is proposed.
Key words: antisense oligonucleotide derivatives; lipid metabolism regulation.

For citation: Oshevski S.I., Ragino Y.I., Kashtanova E.V., Polonskaya Y.V., Stakhneva E.M., Nikolin V.P., Popova N.A., Kolchanov N.A., Voevoda M.I. Changes induced in mouse lipid metabolism by simultaneous impact of antisense oligonucleotide derivatives to *apoB*, *PCSK9*, and *apoCIII* mRNAs. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1020-1025. DOI 10.18699/VJ19.579 (in Russian)

Введение

Дислипидемия – значимый фактор развития раннего атеросклероза и сердечно-сосудистых заболеваний, поэтому снижение уровня холестерина липопротеидов невысокой плотности (ХС не-ЛПВП) во многих случаях представляется важным для предотвращения этих болезней. В качестве профилактического средства хорошо зарекомендовали себя статины. Однако для пациентов с наследственной гиперхолестеринемией и/или не переносящих статины (Visser et al., 2012) существует потребность в дополнительных или альтернативных стратегиях эффективного снижения ХС не-ЛПВП. Такой стратегией в последние годы становится терапия антисмысловыми олигонуклеотидными производными (АСО), базирующаяся на антисмысловом подходе (Antisense Drug Technology…, 2008), когда вводимое олигонуклеотидное производное комплементарно связывается с мРНК или пре-мРНК значимого для липидного обмена белка и, таким образом, блокирует ее трансляцию, т. е. в итоге снижает концентрацию данного белка в клетке.

К 2018 г. Food and Drug Administration (FDA, США) допустило на рынок только один препарат на основе антисмысловых олигонуклеотидных производных, предназначенный для лечения дислипидемии. KYNAMRO (мипомерсен) предназначен для лечения гомозиготной наследственной гиперхолестеринемии. Он снижает уровень апоB в плазме и уровень ХС ЛПНП в дозозависимой и экспозиционной манере (Geary et al., 2015), но имеет существенные побочные эффекты (Panta et al., 2015). Другие препараты: IONIS-APOCIIIRx (volanesorsen) – для лечения наследственной хиломикронемии и семейной частичной липодистрофии – прошел третью стадию клинических испытаний, но не был допущен FDA на рынок в 2018 г.; GalNAc-коньюгированные IONIS-APO(a)-LRx – для снижения уровня липопротеина (a) и IONIS-ANGPTL3-LRx – для преодоления смешанной дислипидемии – на второй стадии клинических испытаний (Hegele et al., 2019). Все эти подходы относятся к монотерапии. По-видимому, монотерапия вообще не может быть достаточно эффективной без побочных эффектов. Очевидно, что любое резкое изменение в концентрации (активности) одного из компонентов системы может приводить к неоднозначным последствиям. Недавно нами было сформулировано представление, что комплексное воздействие может быть эффективнее, так как система регуляции липидного обмена сложна и ее компоненты – это взаимосвязанные участники. Множественное, менее агрессивное воздействие АСО на ряд участников системы, возможное в разной степени, а, может быть, и разнонаправленное, представляется перспективным (Ошевский и др., 2015). Была показана принципиальная возможность влияния на липидный обмен сразу несколькими АСО (Ошевский и др., 2015). При воздействии АСО к трем мРНК важнейших генов липидного обмена, непосредственно связанных согласно генным сетям, – *anoB*, *PCSK9* и *anoCIII* – наибольший эффект снижения уровня ХС ЛПНП и наилучшее соотношение ХС ЛПВП/ХС ЛПНП достигалось при одновременном воздействии тремя АСО. Максимальную эффективность действия АСО в комбинациях наблюдали при индукции жировой диетой.

В настоящей работе задействованы ранее использованные варианты и добавлены: АСО к апоCIII как самостоятельный реагент, набор трех АСО в большей концентрации двух из них и комбинированное сочетание АСО к PCSK9 с АСО к апоCIII (комбинация 6). Такое расширенное воздействие развивает стратегию антисмыслового подхода. Фактически впервые предлагается временной комбинационно-концентрационный подход.

Антисмысловые олигонуклеотидные производные, использованные в работе, относятся к Гэпмерам (Gapmers). Это олигонуклеотидные производные, имеющие на 5′- и 3′-концах короткие фрагменты олигомера – крылья (wings), отличные по химической природе от центральной части Гэпмера. Крылья придают Гэпмеру дополнительные свойства, например большую устойчивость к нуклеазам и т. д. Олигонуклеотидные производные длиной 13 и 20 нуклеотидов, защищенные от действия нуклеаз наличием межнуклеотидных фосфоротиоатных связей и блоками LNA-нуклеотидов (locked nucleic acids) на 5′- и 3′-концах (Antisense Drug Technology…, 2008), предложены нами ранее (Ошевский и др., 2015). Они представлены в таблице, где приведены также данные о специфичности действия их аналогов, имеющих такие же нуклеотидные адреса. Химическое отличие АСО в нашей работе от используемых другими авторами – это межнуклеотидные фосфатные остатки в LNA-блоках вместо фосфоротио-

Свойства использованных АСО

Структура АСО (LNA-нуклеотиды подчеркнуты)	Действие на мышь аналогов – полных фосфоротиоатов (лит. источник)	Вводимое кол-во, мг (в работе)
5′-$\underline{GC}p_sAp_sTp_sTp_sGp_sGp_sTp_sAp_sTp_s\underline{TCA}$ мишень: мРНК апоB	Снижает уровень ХС, ХС ЛПНП (на 80 %) при снижении ХС ЛПВП (на 50 %) (Straarup et al., 2010)	0.04
5′-$\underline{GT}p_sCp_sTp_sGp_sTp_sGp_sGp_sAp_sAp_s\underline{GCG}$ мишень: мРНК PCSK9	Снижает уровень мРНК PCSK9, что приводит к увеличению содержания LDLR в 2–3 раза (Gupta et al., 2010)	0.08
5′-$\underline{GA}p_sGp_sAp_sAp_sCp_sAp_sTp_sAp_sCp_sTp_s\underline{TT}p_sCp_sCp_sCp_sCp_s\underline{TTA}$ мишень: мРНК апоCIII, курсивом обозначена сделанная замена: T (у крысы) заменено на *C*	Должен снижать уровень апоCIII и ТГ, его аналог снижал в два раза уровень апоCIII и на 40–60 % уровень ТГ в крови крыс на 40-й день воздействия (Holmberg et al., 2011)	0.1

атных. Эта особенность может быть весьма полезна, так как для АСО с межнуклеотидными фосфоротиоатными остатками в LNA-блоках, особенно с последовательностями TCC и TGC, показана гепатотоксичность (Burdick et al., 2014). Кроме того, олигонуклеотидные производные имеют структуры, исключающие значимые комплементационные взаимодействия их друг с другом при совместном использовании, которое, очевидно, приводило бы к снижению их эффективности. Мы исследовали последовательные изменения ХС ЛПВП, ХС не-ЛПВП, ХС и АЛТ в течение длительного периода у каждого животного индивидуально, в том числе и в контроле. Ситуация соответствует понятию «персонализированная медицина». Исследование воздействия АСО на одну пациентку было недавно опубликовано (Van Poelgeest et al., 2013).

Цель нашей работы – на примере липидного обмена мыши показать возможность нового подхода в антисмысловой технологии – воздействовать одновременно несколькими антисмысловыми олигонуклеотидными производными, направленными на мРНК взаимосвязанных генов, используя производные в различных сочетаниях, изменяя сочетания производных в разные временные периоды и варьируя концентрации производных в сочетаниях.

Материалы и методы

Антисмысловые олигонуклеотидные производные синтезированы фирмой «Биосинтез» (Новосибирск, Россия), выделены методом обращенно-фазовой хроматографии и охарактеризованы с помощью электрофореза в денатурирующем ПААГ. Использовали самок мыши линии C57BL/6J в возрасте 20 дней, имеющих вес 20–21 г, выращенных в виварии конвенциональных животных ИЦиГ СО РАН. Эксперименты выполняли в условиях вивария, как в работе (Ошевский и др., 2015), при температуре 22–24 °C и влажности около 65 %. Рацион составлял корм «Чара» SPF-категории (ЗАО «Ассортимент-Агро», Россия). Для индукции жировой диетой в корм добавляли жир до общего содержания 17 %.

АСО вводили в физиологическом растворе в объеме 200 мкл в хвостовую вену, кровь забирали из ретроорбитального синуса в количестве 250–300 мкл. Перед забором крови мышей лишали пищи на 16 ч. Сыворотку крови получали стандартным методом: через 1.5 ч при комнатной температуре после свертывания крови сгусток аккуратно отделяли носиком автоматической пипетки от стенок пробирки и центрифугировали содержимое при 1500 g (4000 об./мин) при комнатной температуре на микроцентрифуге Eppendorf в течение 4 мин. Сыворотку аккуратно собирали автоматической пипеткой и хранили замороженной при –70 °C. Во всех образцах сыворотки определяли ХС, холестерин липопротеидов высокой плотности (ХС ЛПВП), ХС не-ЛПВП и уровень АЛТ энзиматическими методами с использованием стандартных реактивов Biocon Fluitest (Германия) на биохимическом анализаторе «Фотометр 5010+». В экспериментах использовали по одному животному, в контрольном эксперименте – четыре животных, которым вводили физиологический раствор.

Результаты и обсуждение

Общая схема эксперимента представлена на рис. 1. АСО вводили растворенными в физиологическом растворе в комбинациях: 1) АСО к апоB, 2) АСО к апоCIII, 3) АСО к апоB и АСО к PCSK9, 4) АСО к апоB, АСО к PCSK9 и АСО к апоCIII, 5) 3АСО к апоB, АСО к PCSK9 и 2АСО к апоCIII, 6) АСО к PCSK9 и (АСО к апоCIII)[1] в количестве, указанном в таблице, или кратно.

Результаты анализа сыворотки крови приведены на рис. 2. В контрольном эксперименте показаны средние значения. Стандартные отклонения значений концентрации ХС не-ЛПВП от средних на 15-, 23- и 38-й дни составляли 5.7, 1 и 4 соответственно. У опытных мышей на 15-й день при воздействии каждым из АСО (комбинации 1, 2 и 6) уровни ХС не-ЛПВП были существенно выше уровня контрольных мышей. По-видимому, воздействие после двух введений АСО оказалось недостаточным для снижения ХС не-ЛПВП, что могло привести к ответной реакции с повышением уровня ХС не-ЛПВП, т. е. было преодолено с ответным превышением. При более комплексном воздействии: двумя (комбинация 3), тремя (комбинация 4) и тремя с увеличенной концентрацией компонентов (комбинация 5) АСО наблюдается значимое результативное снижение уровней ХС не-ЛПВП, что коррелирует с полученными соотношениями ХС ЛПВП/ХС не-ЛПВП для комбинаций АСО на 15-й день (рис. 3). При дальнейшем введении АСО к 23-му дню воздействие во всех комбинациях АСО оказалось эффективным: снижение ХС не-ЛПВП составляло относительно контроля 25, 16, 35, 47, 60 и 7 % соответственно.

[1] Три раза вводили АСО к PCSK9, а в четвертый раз на 31-й день – АСО к PCSK9 и АСО к апоCIII.

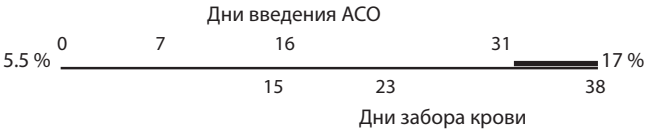


Рис. 1. Общая схема введения АСО, забора крови и диеты.

Линией обозначено время, ее толщина отражает содержание жира в диете животных (%).

Согласно рис. 2, снижение ХС не-ЛПВП больше в комбинации 3 – двух производных АСО к апоВ и АСО к PCSK9, чем при действии только АСО к апоВ, АСО к апоCIII или только АСО – к PCSK9. Снижение еще больше в полной комбинации АСО (47 %) и достигает величины 60 % в полной комбинации с увеличенными концентрациями АСО. При этом наблюдается и улучшение соотношения ХС ЛПВП/ХС не-ЛПВП относительно контроля: в 1.8, 1.5, 1.9, 2.4, 3.1 и 1.24 раза, коррелирующее с изменениями ХС не-ЛПВП при действии АСО в соответствующих сочетаниях (см. рис. 3). Таким образом, на 23-й день еженедельного воздействия АСО на мышей при обычной диете существенно снижается уровень ХС не-ЛПВП и улучшается соотношение ХС ЛПВП/ХС не-ЛПВП, особенно при действии АСО в максимальной комбинации и в максимальной комбинации с увеличенной концентрацией компонентов.

Далее систему липидного обмена стимулировали жировой диетой, увеличив содержание жира в три раза – от 5.5 до 17 % – непосредственно после введения следующей дозы АСО (31-й день). Результаты анализов сывороток на 7-й день после последнего введения представлены на рис. 2. Уровень ХС не-ЛПВП в контроле повысился в 1.34 раза и является самым высоким; на 2 % ниже уровень в случае АСО к апоCIII (повысился на 55 %). В случае АСО к апоВ и комбинации АСО к апоВ и PCSK9 уровни ХС не-ЛПВП равны и на 28 % ниже, чем в контроле (повысились на 29 и 47 % соответственно). В случае комбинации 6, когда к АСО к PCSK9 добавилось АСО к апоCIII, уровень ХС не-ЛПВП существенно понизился – в 1.36 раза, и он был на 49 % ниже, чем в контроле. В наиболее полной комбинации АСО наблюдается снижение уровня ХС не-ЛПВП в 1.3 раза, что на 70 % меньше, чем в контроле. При этом, как видно из рис. 3, соотношение ХС ЛПВП/ХС не-ЛПВП относительно контроля в комбинациях 1, 2, 3, 4, 6 лучше в 1.5, 1.1, 2, 3.7, 2 раза соответственно. По сравнению с 23-м днем соотношение стало существенно выше только при действии комбинации 6 и наиболее полной комбинации АСО в 1.6 и 1.5 раза соответственно, несмотря на жировую диету.

Важно, что при жировой диете максимально низкий уровень ХС не-ЛПВП по сравнению с контролем с разницей 70 % и наилучшее соотношение ХС ЛПВП/ ХС не-ЛПВП – 6.4 наблюдаются именно при наиболее полной комбинации АСО. Существенным является результат изменения уровня ХС не-ЛПВП при действии АСО в максимальной комбинации с увеличенной концентрацией компонентов (комбинация 5). На 15-й день он максимально низкий, к 23-му дню возрастает в 13 раз, оставаясь при этом самым низким (минус 60 % относительно контроля) и далее при

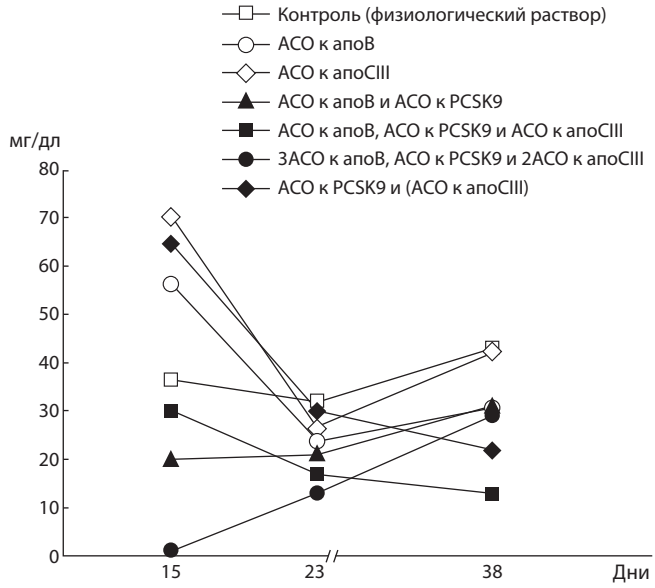


Рис. 2. Уровень концентрации ХС не-ЛПВП в сыворотке крови мышей линии C57BL/6J.

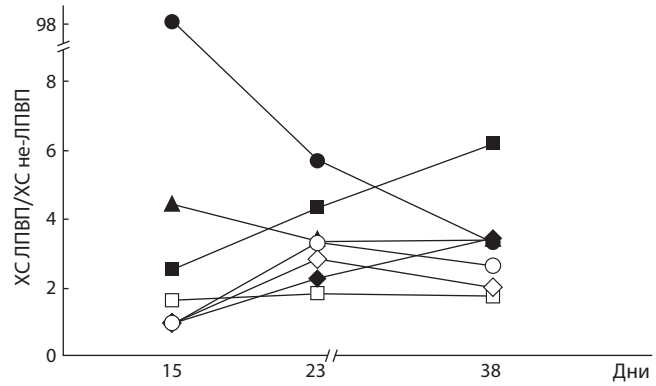


Рис. 3. Соотношение ХС ЛПВП/ХС не-ЛПВП в сыворотке крови мышей линии C57BL/6J.

Обозначения см. на рис. 2.

индукции жировой диетой увеличивается еще в 2.2 раза, достигая уровня на 5 % меньше, чем в случае действия АСО к апоВ и комбинации АСО к апоВ и PCSK9. Эти изменения коррелируют с уменьшением соотношения ХС ЛПВП/ХС не-ЛПВП от 98 до 5.8 и 3.3 к 38-му дню для комбинации 5. К 38-му дню соотношение становится равным соотношению для комбинации 3 АСО к апоВ и PCSK9 и близко соотношению для комбинации 1 АСО к апоВ, равному 2.6. Таким образом, на 15-й день уровень ХС не-ЛПВП чрезвычайно низкий и постепенно повышается, несмотря на действие трех АСО, два из которых в высокой концентрации. По-видимому, столь резкое снижение уровня ХС не-ЛПВП к 15-му дню приводит к активной ответной реакции организма мыши, направленной на преодоление эффекта препаратов АСО в высокой концентрации, что дополнительно стимулируется жировой диетой и завершается достижением приемлемого уровня. Следует отметить, что при максимальной нагрузке АСО концентрация АЛТ в 15-, 23- и 38-й дни составляла

35, 40, 40 ед./л соответственно, т. е. фактически оставалась без изменений и так же, как у контрольных мышей, была около 40 ед./л. Использованные олигонуклеотидные производные обладают высокой стоимостью[2]. Учитывая невозможность получения на первом этапе статистических данных, в работе оценивали только существенные количественные различия. Важно, что мы исследовали последовательные изменения ХС не-ЛПВП в течение длительного периода индивидуально у каждого животного. При сравнении эффектов воздействия трех АСО в двух концентрационных сочетаниях и двух АСО в комбинации 6, а также полной картины наблюдаемых изменений уровней ХС не-ЛПВП во всех вариантах действия АСО в 15-, 23- и 38-й дни можно заключить, что для достижения наибольшей эффективности антисмысловых олигонуклеотидных производных в комбинациях важно учитывать необходимость подбора концентраций, в том числе и в сочетаниях, и порядок и периоды введения АСО.

П. Замечник в 2007 г. впервые предложил одновременно использовать три фосфоротиоатные олигонуклеотидные производные для подавления бактериальной активности *Mycobacterium tuberculosis*, ингибируя три миколил-трансферазы (Harth et al., 2007). В 2012 г. группа исследователей воздействовала 10 АСО для коррекции сплайсинга РНК дистрофина (Aoki et al., 2012). Работа развивается и в настоящее время, показано действие коктейля конъюгатов морфолиновых олигонуклеотидных производных с пептидом на изменение сплайсинга экзонов 6 и 8 РНК дистрофина сердечной мышцы на модели дистрофии Дюшена собак (Echigoya et al., 2017). В 2015 г. появился очень интересный вариант воздействия на два гена мыши, *apoB* и *apoCIII*, полинуклеотидным производным, состоящим из двух LNA-содержащих Гэпмеров, соединенных между собой четырехнуклеотидным природным линкером, который гидролизуется внутри клетки, высвобождая два коротких функционирующих в качестве антисмысловых олигонуклеотидных производных агента. R.R. Subramanian с коллегами (2015) показали высокую эффективность конструкции, они считают, что ее использование способствует уменьшению количества побочных эффектов от АСО к РНК апоCIII.

Таким образом, наблюдается тенденция к развитию антисмысловой технологии в сторону воздействия не одним АСО. В приведенных выше вариантах не рассматривается значение влияния в комбинации концентрации каждого АСО, нацеленного на РНК каждого из генов, или их последовательности воздействия. Именно в случае воздействия на целую систему организма, а не на один процесс, и важны рассматриваемые в работе концентрационные зависимости в сочетаниях АСО и возможная последовательность их действия.

Заключение

Таким образом, впервые предлагается временной комбинационно-концентрационный подход антисмысловой технологии, открывающий новые перспективы создания терапевтических инструментов воздействия на системы организма. Подход согласуется с концепцией персонализированной медицины, так как допускает гибкость и вариабельность, учитывающие особенности организма пациента, в результате чего можно ожидать большей эффективности воздействия АСО, по сравнению с монотерапией, и снижения острых проявлений побочных эффектов.

[2] В настоящее время рано планировать применение комбинаций АСО к лечению дислипидемии у человека в силу большой стоимости производных использованного типа. Но по мере увеличения потребности и, как следствие, увеличения масштабов синтеза, стоимость АСО, как показывает практика, существенно снижается.

Список литературы / References

Ошевский С.И., Рагино Ю.И., Каштанова Е.В., Полонская Я.В., Стахнева Е.М., Николин В.П., Попова Н.А., Кораблев А.Н., Колчанов Н.А., Воевода М.И. Одновременное воздействие несколькими антисмысловыми олигонуклеотидными производными, его эффективность на примере липидного обмена мыши. Атеросклероз. 2015;11(3):72-78.
[Oshevski S.I., Ragino Y.I., Kashtanova E.V., Polonskaya Y.V., Stakhneva E.M., Nikolin V.P., Popova N.A., Korablev A.N., Kolchanov N.A., Voevoda M.I. The efficacy of simultaneous action of several antisense oligonucleotide derivatives by the example of lipid metabolism in mice. Ateroskleroz = Atherosclerosis. 2015;11(3):72-78. (in Russian)]

Antisense Drug Technology: Principles, Strategies, and Applications. Second Edition. Crooke S.T. (Ed.). CRC Press, 2008.

Aoki Y., Yokota T., Nagata T., Nakamura A., Tanihata J., Saito T., Duguez S.M., Nagaraju K., Hoffman E.P., Partridge T., Takeda S. Body-wide skipping of exons 45-55 in dystrophic mdx52 mice by systemic antisense delivery. Proc. Natl. Acad. Sci. USA. 2012;109(34): 13763-13768. DOI 10.1073/pnas.1204638109.

Burdick A.D., Sciabola S., Mantena S.R., Hollingshead B.D., Stanton R., Warneke J.A., Zeng M., Martsen E., Medvedev A., Makarov S.S., Reed L.A., Davis J.W. 2nd, Whiteley L.O. Sequence motifs associated with hepatotoxicity of locked nucleic acid–modified antisense oligonucleotides. Nucleic Acids Res. 2014;42(8):4882-4891. DOI 10.1093/nar/gku142.

Echigoya Y., Nakamura A., Nagata T., Urasawa N., Lim K.R.Q., Trieu N., Panesar D., Kuraoka M., Moulton H.M., Saito T., Aoki Y., Iversen P., Sazani P., Kole R., Maruyama R., Partridge T., Takeda S., Yokota T. Effects of systemic multiexon skipping with peptide-conjugated morpholinos in the heart of a dog model of Duchenne muscular dystrophy. Proc. Natl. Acad. Sci. USA. 2017;114(16):4213-4218. DOI 10.1073/pnas.1613203114.

Geary R.S., Baker B.F., Crooke S.T. Clinical and preclinical pharmacokinetics and pharmacodynamics of mipomersen (Kynamro®): a second-generation antisense oligonucleotide inhibitor of apolipoprotein B. J. Clin. Pharmacokinet. 2015;54(2):133-146. DOI 10.1007/s40262-014-0224-4.

Gupta N., Fisker N., Asselin M.C., Lindholm M., Rosenbohm C., Ørum H., Elmén J., Seidah N.G., Straarup E.M. A locked nucleic acid antisense oligonucleotide (LNA) silences PCSK9 and enhances LDLR expression *in vitro* and *in vivo*. PLoS One. 2010;5(5):e10682. DOI 10.1371/journal.pone.0010682.

Harth G., Zamecnik P.C., Tabatadze D., Pierson K., Horwitz M.A. Hairpin extensions enhance the efficacy of mycolyl transferase-specific antisense oligonucleotides targeting *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA. 2007;104(17):7199-7204. DOI 10.1073/pnas.0701725104.

Hegele R.A., Tsimikas S. Lipid-lowering agents. Circ. Res. 2019; 124(3):386-404. DOI 10.1161/CIRCRESAHA.118.313171.

Holmberg R., Refai E., Höög A., Crooke R.M., Graham M., Olivecrona G., Berggren P.O., Juntti-Berggren L. Lowering apolipoprotein CIII delays onset of type 1 diabetes. Proc. Natl. Acad. Sci. USA. 2011;108(26):10685-10689. DOI 10.1073/pnas.1019553108.

Panta R., Dahal K., Kunwar S. Efficacy and safety of mipomersen in treatment of dyslipidemia: A meta-analysis of randomized control-

led trials. J. Clin. Lipidol. 2015;(2):217-225. DOI 10.1016/j.jacl.2014.12.006.

Straarup E.M., Fisker N., Hedtjärn M., Lindholm M.W., Rosenbohm C., Aarup V., Hansen H.F., Ørum H., Hansen J.B., Koch T. Short locked nucleic acid antisense oligonucleotides potently reduce apolipoprotein B mRNA and serum cholesterol in mice and non-human primates. Nucleic Acids Res. 2010;38(20):7100-7111. DOI 10.1093/nar/gkq457.

Subramanian R.R., Wysk M.A., Ogilvie K.M., Bhat A., Kuang B., Rockel T.D., Weber M., Uhlmann E., Krieg A.M. Enhancing antisense efficacy with multimers and multi-targeting oligonucleotides (MTOs) using cleavable linkers. Nucleic Acids Res. 2015;43(19):9123-9132. DOI 10.1093/nar/gkv992.

van Poelgeest E.P., Swart R.M., Betjes M.G., Moerland M., Weening J.J., Tessier Y., Hodges M.R., Levin A.A., Burggraaf J. Acute kidney injury during therapy with an antisense oligonucleotide directed against PCSK9. Am. J. Kidney Dis. 2013;62(4):796-800. DOI 10.1053/j.ajkd.2013.02.359.

Visser M.E., Witztum J.L., Stroes E.S., Kastelein J.J. Antisense oligonucleotides for the treatment of dyslipidaemia. Eur. Heart J. 2012; 33(12):1451-1458. DOI 10.1093/eurheartj/ehs084.

ORCID ID

M.I. Voevoda orcid.org/0000-0001-9425-413X

Благодарности. Работа выполнена по бюджетному проекту ИЦиГ СО РАН № 0324-2019-0041 и поддержана грантом, полученным Новосибирским национальным исследовательским государственным университетом в рамках Проекта 5-100. Авторы выражают глубокую признательность В.И. Каледину и В.А. Напримерову за содействие в работе.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 09.11.2017. После доработки 04.10.2019. Принята к публикации 04.10.2019.

МЕДИЦИНСКАЯ ГЕНЕТИКА
Оригинальное исследование/ Original article
УДК 576.316.24:616-097.3

Вавиловский журнал генетики и селекции. 2019;23(8):1026-1031
DOI 10.18699/VJ19.580

Ассоциация длины теломер лейкоцитов и уровня специфических антител после вакцинации против клещевого энцефалита

Н.С. Юдин[1, 2, 3] ✉, В.А. Белявская[4], В.Н. Максимов[1, 2, 3], Д.Е. Иванощук[1, 2, 3], П.С. Орлов[1, 2, 3], М.И. Воевода[1, 2, 3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Научно-исследовательский институт терапии и профилактической медицины – филиал Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[4] Государственный научный центр вирусологии и биотехнологии «Вектор» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека, р. п. Кольцово, Новосибирская область, Россия
✉ e-mail: yudin@bionet.nsc.ru

Основным принципом индивидуализации вакцинаций является создание достаточной иммунной защиты организма при избегании излишней иммунизации. Отсюда возникает необходимость разработки методов прогнозирования величины иммунного ответа еще до проведения вакцинации. Длину теломер можно рассматривать как перспективный прогностический параметр для оценки иммунной реакции пациента при проведении вакцинации. Целью работы был анализ возможной ассоциации длины теломер лейкоцитов с уровнем специфических антител после вакцинации против клещевого энцефалита. В исследование было включено 55 мужчин и 40 женщин, ранее не вакцинированных против клещевого энцефалита и не имевших контактов с клещами. Вакцинацию проводили вакциной «ЭнцеВир». Через месяц после вакцинации анализировали уровень специфических антител IgG против вируса клещевого энцефалита с использованием тест-системы «ВектоВКЭ-IgG-стрип», а также длину теломер лейкоцитов с помощью количественной ПЦР в реальном времени. По уровню ответа на вакцинацию пациентов разделяли на нереагировавших (уровень IgG 0–100 МЕ/мл), низкореагировавших (уровень IgG 101–200 МЕ/мл) и высокореагировавших (уровень IgG выше 200 МЕ/мл). Длина теломер как минимум на уровне тенденции ($p < 0.1$) зависела как от ответа на вакцинацию, так и от возраста, уровня образования и наличия психоэмоционального стресса. С помощью общей линейной модели была выявлена ассоциация длины теломер с ответом на вакцинацию против клещевого энцефалита на уровне тенденции ($p < 0.1$) только у женщин. При попарном сравнении длина теломер у высокореагировавших женщин была достоверно выше, чем у нереагировавших. Таким образом, нами обнаружена ассоциация длины теломер лейкоцитов и уровня специфических антител после вакцинации против клещевого энцефалита у женщин. По-видимому, длину теломер лейкоцитов в крови можно рассматривать в качестве перспективного маркера для прогноза пролиферативного ответа лимфоцитов и величины иммунологической реакции пациента в ответ на вакцинацию.
Ключевые слова: клещевой энцефалит; вакцинация; IgG антитела; длина теломер; пол; биомаркер.

Для цитирования: Юдин Н.С., Белявская В.А., Максимов В.Н., Иванощук Д.Е., Орлов П.С., Воевода М.И. Ассоциация длины теломер лейкоцитов и уровня специфических антител после вакцинации против клещевого энцефалита. Вавиловский журнал генетики и селекции. 2019;23(8):1026-1031. DOI 10.18699/VJ19.580

Association between leukocyte telomere length and specific antibody levels after vaccination against tick-borne encephalitis

N.S. Yudin[1, 2, 3] ✉, V.A. Belyavskaya[4], V.N. Maksimov[1, 2, 3], D.E. Ivanoshchuk[1, 2, 3], P.S. Orlov[1, 2, 3], M.I. Voevoda[1, 2, 3]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Research Institute of Internal and Preventive Medicine – Branch of the Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[4] State Research Center of Virology and Biotechnology "Vector" of the Federal Service for Surveillance in Consumer Rights Protection and Human Well-being, Koltsovo, Novosibirsk region, Russia
✉ e-mail: yudin@bionet.nsc.ru

The primary objective of personalized vaccination is to induce an efficient immune defense while avoiding excessive immunization. Hence, it necessitates the development of methods for predicting the magnitude of the immune response prior to vaccination. Telomere length can be considered as a promising prognostic parameter for assessing the immune response to vaccination. The aim of the work was to analyze the possible association between leukocyte telomere length and specific antibody levels after vaccination against tick-borne encephalitis. The study included 55 men and 40 women who had not previously been vaccinated against tick-borne encephalitis and had no contacts with ticks. Vaccination was carried out with the EnceVir vaccine. One month after vaccination, the level of specific IgG antibodies against tick-borne encephalitis virus was analyzed using the VektoVKE-IgG-strip test system and leukocyte telomere length was measured

© Юдин Н.С., Белявская В.А., Максимов В.Н., Иванощук Д.Е., Орлов П.С., Воевода М.И., 2019

using real-time quantitative PCR. According to the intensity of vaccine-elicited immune responses, patients were divided into three groups: unresponsive (IgG level 0–100 IU/ml), slightly responsive (IgG level 101–200 IU/ml) and highly responsive (IgG level above 200 IU/ml). The telomere length, at least at trend level ($p < 0.1$), correlated with the response to vaccination as well as age, educational level and the presence of emotional stress. Using a general linear model, an association between telomere length and immune response to vaccination against tick-borne encephalitis at trend level ($p < 0.1$) was found only in women. Using a pairwise comparison, it was found that telomere length was significantly higher in highly responsive women than in unresponsive women. Hence, an association between leukocyte telomere length and specific antibody levels after vaccination against tick-borne encephalitis was identified in women. Therefore, peripheral blood leukocyte telomere length can be considered as a promising marker for predicting lymphocyte proliferative responses and the magnitude of vaccine-elicited cellular immune responses.
Key words: tick-borne encephalitis; vaccination; IgG antibodies; telomere length; gender; biomarker.

For citation: Yudin N.S., Belyavskaya V.A., Maksimov V.N., Ivanoshchuk D.E., Orlov P.S., Voevoda M.I. Association between leukocyte telomere length and specific antibody levels after vaccination against tick-borne encephalitis. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1026-1031. DOI 10.18699/VJ19.580 (in Russian)

Введение

Клещевой энцефалит – природно-очаговая вирусная инфекция, характеризующаяся лихорадкой, интоксикацией и поражением серого вещества головного мозга (энцефалит) и/или мозговых оболочек (менингит) (Valarcher et al., 2015). Возбудителем клещевого энцефалита является РНК-содержащий вирус, относящийся к семейству Flaviviridae, который передается человеку через укус инфицированных клещей из рода *Ixodes* (Gritsun et al., 2003; Юдин и др., 2018). Клещевой энцефалит распространен в России, Прибалтике, Скандинавии, Восточной и Центральной Европе, Китае и Японии (Steffen, 2019). Случаи заболевания все чаще отмечаются на новых территориях, что свидетельствует о миграции вируса (Jääskeläinen et al., 2016). Клинический спектр заболевания варьирует от легких лихорадочных до тяжелых форм, сопровождающихся поражением оболочек и вещества мозга. Смертность от клещевого энцефалита для вирусов европейского и сибирского субтипов составляет 0.5–2 %. Это заболевание может привести к стойким неврологическим и психиатрическим осложнениям у 30–40 % пациентов (Haglund, Günther, 2003; Karelis et al., 2012; Veje et al., 2016).

Наиболее эффективным способом профилактики клещевого энцефалита, формирующим длительный активный иммунитет, остается вакцинация (Воробьева и др., 2006). Например, в Австрии, где охват населения вакцинацией достиг 86 %, начиная с 2001 г. случаи заболевания отмечены только среди непривитого населения (Heinz, Kunz, 2004). В России ежегодно от клещевого энцефалита вакцинируется не менее 2.5–3 млн человек (Онищенко, 2000). Однако уровень современных знаний о механизме развития иммунитета и возможности его прогнозирования, а также о наличии потенциальных побочных эффектов позволяет говорить о необходимости дифференцированного подхода к проведению профилактических вакцинаций (индивидуализации вакцинаций) (Медуницын, 2004). Основной принцип индивидуализации вакцинаций – создание достаточной иммунной защиты организма при избегании излишней иммунизации. Например, некоторые пациенты могут реагировать высокой выработкой антител уже на первичную вакцинацию и, таким образом, не требуют повторной ревакцинации. Поэтому по этическим и экономическим причинам избыточная иммунизация недопустима. Одним из подходов к индивидуализации вакцинации является выбор дозы вакцины в зависимости от ожидаемого уровня иммунного ответа. Отсюда возникает необходимость в разработке методов прогнозирования величины иммунологических реакций пациента еще до проведения вакцинации. Известно, что индивидуальная вариабельность выработки и длительной циркуляции антител после вакцинации в значительной степени обусловлена генетическими особенностями индивида (Haralambieva et al., 2010; Gröndahl-Yli-Hannuksela et al., 2016; Yudin et al., 2018).

Теломеры – концевые участки хромосом, состоящие из сотен повторов $(TTAGGG)_n$ и теломер-связывающих белков, которые предохраняют ДНК от деградации (Blackburn et al., 2015). Поскольку ДНК-полимераза не может реплицировать линейные концы хромосом, каждое деление клеток приводит к потере небольшого участка теломерной ДНК. В делящихся клетках утраченные фрагменты теломер восстанавливаются с помощью фермента теломеразы (Shawi, Autexier, 2008). Тем не менее критическое укорочение теломер, которое накапливается в ходе митозов, приводит к прекращению деления клетки или апоптозу (Harley et al., 1990).

Поскольку эффективный адаптивный иммунный ответ основан на клональной экспансии большого количества эффекторных лимфоцитов, поддержание длины теломер играет важную роль в регуляции пролиферации лимфоцитов (Hodes et al., 2002; Weng, 2012). Так, было показано, что активированные В-клетки герминативных центров лимфатических узлов имели повышенную длину теломер и повышенный уровень активности теломеразы по сравнению с неактивированными (наивными) В-клетками (Weng et al., 1997). Однако другие исследователи не смогли воспроизвести этот эффект (Son et al., 2003). Длина теломер в наивных Т-клетках превышала этот параметр в Т-клетках памяти у одного и того же индивида, независимо от возраста (Weng et al., 1995). Длина теломер лимфоцитов может существенно различаться между В- и Т-клетками (Lin et al., 2015). Хроническое воздействие высоких уровней антигенов у пациентов с хроническими вирусными инфекциями (вирусы Эпштейна–Барр, гепатита B/C/D, ВИЧ и т. д.) может приводить к прогрессирующему укорочению теломер антигенспецифических Т-клеток, что, в свою очередь, ослабляет противовирусный иммунитет (Bellon, Nicot, 2017). В целом эти работы указывают на критическую роль длины теломер при активации Т и В-лимфоцитов.

Таким образом, длину теломер можно рассматривать как перспективный прогностический параметр для оценки иммунного ответа пациента при проведении вакцинации. Однако в литературе имеются лишь две работы, в которых длину теломер анализировали после вакцинации, – против гриппа (Najarro et al., 2015) и опоясывающего герпеса (Verschoor et al., 2017). Целью нашей работы был анализ возможной ассоциации длины теломер лейкоцитов с уровнем специфических антител после вакцинации против клещевого энцефалита.

Материалы и методы

В исследование было включено 55 мужчин и 40 женщин в возрасте от 15 до 30 лет (средний возраст 24.9 ± 0.5), ранее не вакцинированных против клещевого энцефалита и не имевших, по результатам анкетирования, контактов с клещами. Испытуемые заполняли анкеты самостоятельно, включая информацию о возрасте, поле, социальном положении (в баллах от 1 до 6), образовании (в баллах от 1 до 5), занимаемой должности (в баллах от 1 до 6), наличии психоэмоционального стресса (да/нет), курении (в баллах от 1 до 3), употреблении алкоголя (в баллах от 1 до 4), наличии аллергии (да/нет), частоте респираторных заболеваний в предшествующий вакцинации период (в баллах от 1 до 3), наличии хронического заболевания (тонзиллит, бронхит, пиелонефрит и т. д.) (да/нет). Вакцинацию проводили вакциной для профилактики клещевого энцефалита «ЭнцеВир» (НПО «Микроген», Томск) по рекомендованной производителем схеме. Все пациенты дали добровольное информированное согласие на участие в эксперименте. Исследование было одобрено этическим комитетом Научно-исследовательского института терапии и профилактической медицины – филиала ИЦиГ СО РАН.

Через месяц после вакцинации у пациентов брали пробы венозной крови. Уровень специфических антител IgG против вируса клещевого энцефалита (ВКЭ) анализировали методом иммуноферментного анализа с использованием тест-системы «ВектоВКЭ-IgG-стрип» (ЗАО «Вектор-Бест», р. п. Кольцово, Новосибирская область). По уровню ответа на вакцинацию пациентов делили на нереагировавших (уровень IgG против ВКЭ 0–100 МЕ/мл, в среднем 40.0 ± 5.8 (n = 22)), низкореагировавших (101–200 МЕ/мл, в среднем 159.2 ± 6.9 (n = 19)) и высокореагировавших (выше 200 МЕ/мл, в среднем 246.4 ± 2.4 (n = 54)).

Геномную ДНК выделяли из крови стандартным методом протеолитической обработки с последующей экстракцией фенолом (Sambrook, Russell, 2006). Исследование длины теломер лейкоцитов (ДТЛ) проводили с помощью количественной ПЦР в реальном времени (qPCR) (Cawthon, 2002). В качестве однокопийного референсного гена был взят ген β-гемоглобина. Отдельные количественные реакции для теломер и гена β-гемоглобина ставили в парных 96-луночных плашках в идентичных позициях. Каждая плашка включала серию разведений контрольного образца ДНК (0.78, 3.13, 12.5, 25, 100 нг), которые были использованы для создания калибровочной кривой. В каждую индивидуальную реакцию брали по 10 нг ДНК. Реакционная смесь для анализа теломер содержала следующие реагенты: 270 нМ tel1 праймер (5′-GGTTTTTGAGGGTGAGGGTGAGGGTGAGGGTGAGGGT-3′), 900 нМ tel2 праймер (5′-CACCAACTTCATCCACGTTCACC-3′), 0.2X SYBR Green I, 5 мМ дитиотреитол, 1 % диметилсульфоксид, по 0.2 мМ каждого дезоксинуклеозидтрифосфата, 1.5 мМ $MgCl_2$, 1.25 ед. ДНК-полимеразы («СибЭнзим», Новосибирск) в конечном объеме 15 мкл буфера для ПЦР. Режим амплификации: денатурация при температуре 95 °C (10 мин), затем 25 циклов, включающих денатурацию при 95 °C (15 с), отжиг при 54 °C (30 с) и элонгацию при 72 °C (90 с).

Реакционная смесь для анализа гена β-гемоглобина содержала следующие реагенты: 300 нМ HBG1 праймер (5′-GCTTCTGACACAACTGTGTTCACTAGC-3′), HBG2 праймер (5′-CACCAACTTCATCCACGTTCACC-3′), 0.2X SYBR Green I, 5 мМ дитиотреитол, 1 % диметилсульфоксид, по 0.2 мМ каждого дезоксинуклеозидтрифосфата, 1.5 мМ $MgCl_2$, 1.25 ед. ДНК-полимеразы («СибЭнзим», Новосибирск) в конечном объеме 15 мкл буфера. ПЦР проводили в следующем режиме: денатурация при температуре 95 °C (3 мин), затем 35 циклов, включающих денатурацию при 95 °C (15 с), отжиг при температуре 58 °C (20 с) и элонгацию при 72 °C (20 с).

Обе реакции ставили на системе для ПЦР в реальном времени StepOnePlus (Applied Biosystems, США). Для каждого образца ДНК ПЦР выполняли не менее трех раз. Данные обрабатывали с помощью штатного программного обеспечения прибора. Проводили контроль качества и расчет отношения T/S (отношение сигналов теломер к однокопийному гену), чтобы получить относительную ДТЛ. Если кривые амплификации образца в трех репликах различались > 0.5 цикла, то такой образец исключали из дальнейшего анализа. Каждая плашка включала три контрольных образца с заранее определенной ДТЛ. Сопоставимость между плашками контролировали по относительным интенсивностям сигналов контрольных образцов.

Данные по длине теломер проверяли на соответствие нормальному распределению с помощью критерия Колмогорова–Смирнова. Поскольку данные не соответствовали нормальному распределению, то для дальнейшего статистического анализа их нормализовали с помощью логарифмической трансформации. Для наглядности в статье данные представлены с помощью нетрансформированных среднего арифметического и стандартной ошибки.

Для оценки ассоциации между длиной теломер и отдельными независимыми факторами, способными принципиально повлиять на результаты исследования (возраст, пол, социальное положение, образование, должность, психоэмоциональный стресс, курение, употребление алкоголя, частота респираторных заболеваний, наличие аллергии и хронического заболевания, ответ на вакцинацию), мы использовали одномерные линейный регрессионный и дисперсионный анализы. Связь длины теломер с ответом на вакцинацию против КЭ также проверяли с помощью метода общей линейной модели (general linear model, GLM). Попарные сравнения проводили путем post hoc анализа с использованием критерия Фишера. Критическое значение уровня статистической значимости при проверке нулевых гипотез принимали равным 0.05. Обработку результатов проводили с помощью пакета Statistica версии 8.0.

Результаты

Протективный уровень IgG против ВКЭ (выше 100 МЕ/мл) отмечался у 80.0 % женщин и 74.5 % мужчин, что хорошо соответствует данным других авторов, полученным при проведении вакцинопрофилактики клещевого энцефалита среди взрослого населения (Билалова, 2009). Это позволило нам сформировать репрезентативные группы нереагировавших (уровень IgG против ВКЭ 40.0 ± 5.8, n = 22), низкореагировавших (уровень IgG против ВКЭ 159.2 ± 6.9, n = 19) и высокореагировавших (уровень IgG против ВКЭ 246.4 ± 2.4, n = 54) индивидов.

Мы применили одномерные регрессионный и дисперсионный анализы, чтобы определить, какие факторы могли влиять на длину теломер, и обнаружили, что на этот параметр как минимум на уровне тенденции ($p < 0.1$) оказывали влияние возраст, уровень образования, наличие психоэмоционального стресса и ответ на вакцинацию (табл. 1). С учетом этих результатов при анализе GLM в модель вводили ответ на вакцинацию как независимую величину, а возраст, уровень образования и психоэмоциональный стресс – как ковариаты. Методом GLM нами была выявлена ассоциация длины теломер с ответом на вакцинацию против клещевого энцефалита на уровне тенденции ($p < 0.1$) в общей группе обоих полов, а при разделении по полу – только у женщин (табл. 2). При попарном сравнении длина теломер у высокореагировавших женщин была достоверно выше, чем у нереагировавших. Эти различия сохранялись также при объединении мужчин и женщин в одну группу.

Таблица 1. Одномерные регрессионный и дисперсионный анализы ассоциации длины теломер с ответом на вакцинацию и другими независимыми факторами, способными повлиять на результаты исследования

Фактор	Критерий	p
Регрессионный анализ		
Возраст*	**β = –0.178**	**0.084**
Социальное положение	β = 0.097	0.350
Образование	**β = 0.231**	**0.024**
Должность	β = –0.075	0.466
Курение	β = 0.154	0.136
Употребление алкоголя	β = –0.076	0.464
Респираторные заболевания	β = 0.012	0.906
Дисперсионный анализ		
Пол	F = 1.461	0.230
Психоэмоциональный стресс	**F = 4.927**	**0.029**
Аллергия	F = 2.638	0.108
Хроническое заболевание	F = 2.183	0.143
Ответ на вакцинацию	**F = 2.568**	**0.082**

* Параметры, выделенные жирным шрифтом, оказывали влияние на длину теломер как минимум на уровне тенденции ($p < 0.1$). β – значение стандартизованного регрессионного коэффициента при регрессионном анализе; F – значение критерия Фишера при дисперсионном анализе.

Обсуждение

Длина теломер в нашем эксперименте, по данным одномерных регрессионного и дисперсионного анализов как минимум на уровне тенденции ($p < 0.1$), зависела как от ответа на вакцинацию, так и от возраста, уровня образования и наличия психоэмоционального стресса (см. табл. 1). Зависимое от хронологического возраста укорочение длины теломер хорошо изучено у человека и, по-видимому, является результатом комбинированного воздействия оксидативного стресса, воспаления и повторяющегося деления клеток, а также возрастных заболеваний (Rizvi et al., 2014). Известно, что хронический стресс приводит также к укорочению длины теломер и подавлению активности теломеразы в клетках иммунной системы (Спивак и др., 2015; de Punder et al., 2018). Положительная корреляция длины теломер с уровнем образования в нашем исследовании, вероятно, связана с наличием значительных психоэмоциональных нагрузок у студентов и снижением их после окончания обучения.

Как правило, женщины имеют более длинные теломеры, чем мужчины, и с возрастом эти различия усиливаются (Gardner et al., 2014). Хотя пол не оказывал существенного влияния на длину теломер по результатам одномерного дисперсионного анализа (см. табл. 1), мы анализировали ассоциацию длины теломер с ответом на вакцинацию как в общей группе, так и отдельно у мужчин и женщин. Оказалось, что женщины с высокой реакцией на вакцинацию имели достоверно более длинные теломеры, чем нереагировавшие. У мужчин этот эффект отсутствовал, но сохранялся при объединении обоих полов в одну группу. По-видимому, этот эффект связан с меньшим распространением среди женщин курения и злоупотребления алкоголем, которые приводят к укорочению длины теломер (Freitas-Simoes et al., 2016; Astuti et al., 2017). Другие исследователи обнаружили, что ассоциация длины теломер с сахарным диабетом (Wang et al., 2016) или социальным стрессом (Willis et al., 2018) также зависит от пола. Ранее нами были выявлены противоположные корреляции длины теломер с объемом талии у мужчин и женщин (Максимов и др., 2016).

Таблица 2. Ассоциация длины теломер с ответом на вакцинацию против клещевого энцефалита

Группа	Длина теломер лейкоцитов*			p
	Нереагировавшие	Низкореагировавшие	Высокореагировавшие	
Мужчины	1.229 ± 0.108 (14)	1.169 ± 0.117 (12)	1.348 ± 0.075 (29)	0.453
Женщины	1.075 ± 0.182 (8)	1.420 ± 0.195 (7)	**1.514 ± 0.103 (25)****	0.096
Оба пола	1.173 ± 0.096 (22)	1.262 ± 0.104 (19)	**1.425 ± 0.061 (54)*****	0.076

* Среднее ± стандартная ошибка (число наблюдений); ** $p = 0.035$ и *** $p = 0.028$ по сравнению с группой нереагировавших по GLM.

Наши данные хорошо соответствуют полученным ранее результатам о положительной ассоциации длины теломер с ответом В- и Т-клеток на вакцинацию против гриппа у пожилых субъектов (Najarro et al., 2015). Однако в работе (Verschoor et al., 2017) длина теломер не коррелировала с иммунным ответом после вакцинации против опоясывающего герпеса. Отсутствие ассоциации в последнем случае может быть связано с тем, что это исследование проводили на лицах преклонного возраста (80–102 года). Уровень специфических антител после вакцинации, особенно в пожилом возрасте, может зависеть от множества факторов: генетической предрасположенности (Haralambieva et al., 2010; Gröndahl-Yli-Hannuksela et al., 2016; Yudin et al., 2018), характера питания и физических нагрузок (Sagawa et al., 2011), типа и дозы вакцины (Cools et al., 2009) и сопутствующих заболеваний (Muszkat et al., 2003).

Конкретный фундаментальный механизм, лежащий в основе наблюдаемой зависимости между длиной теломер лейкоцитов и величиной иммунного ответа на вакцинацию, в настоящее время не вполне понятен. Известно, что длина теломер может регулировать экспрессию генов задолго до того, как теломеры достаточно укорачиваются, чтобы вызвать клеточный ответ на повреждение ДНК (Laberthonniere et al., 2019). На культурах клеток человека было показано, что «выпетливание» хромосом может сближать «длинные» теломеры с генами на расстоянии до 10 млн п. н. друг от друга, а при укороченных теломерах эти локусы остаются физически разделенными (так называемый теломерный эффект положения на длинных расстояниях, TRE-OLD) (Robin et al., 2014). В частности, авторы обнаружили, что посредством TRE-OLD длина теломер может регулировать экспрессию генов иммунной системы (ген системы комплемента *C1S* и ген сигнального пути интерферона *ISG15*). A.K. Mukherjee с коллегами (2018) показали, что белки шелтерина (белковый комплекс, защищающий теломеры от механизмов репарации ДНК и регулирующий активность теломеразы) способны также регулировать экспрессию генов вне теломерных районов. Например, связывание белка шелтерина TRF2 с промоторами генов на расстоянии до 60 млн п. н. от теломерной ДНК зависело от длины теломер.

Заключение

Таким образом, нами обнаружена ассоциация длины теломер лейкоцитов и уровня специфических антител после вакцинации против клещевого энцефалита у женщин. Хотя длина теломер лимфоцитов может различаться между В- и Т-клетками (Lin et al., 2015), а также разновидностями В- и Т-клеток (Weng et al., 1995, 1997), по-видимому, длину теломер лейкоцитов в крови можно рассматривать в качестве перспективного маркера для прогноза пролиферативного ответа лимфоцитов и величины иммунологической реакции пациента в ответ на вакцинацию.

Список литературы / References

Билалова Г.П. Вопросы практического применения вакцины «Энцевир». Сиб. мед. журн. 2009;2:86-91.
[Bilalova G.P. Issues of the practical using of Encevir vaccine. Sibirskiy Meditsinskiy Zhurnal = Siberian Medical Journal (Tomsk). 2009;2:86-91. (in Russian)]

Воробьева М.С., Ращепкина М.Н., Павлова Л.И., Быстрицкий Л.Д., Ставицкая Н.Х., Ильченко Т.Э., Билалова Г.П., Мищенко И.А., Шарова О.И. Вакцинопрофилактика клещевого энцефалита на современном этапе и препараты для ее реализации. Бюлл. сиб. медицины, 2006;5(S1):63-71.
[Vorob'yeva M.S., Rashchepkina M.N., Pavlova L.I., Bystritskiy L.D., Stavitskaya N.H., Il'chenko T.E., Bilalova G.P., Mishchenko I.A., Sharova O.I. The present vaccinoprophilaxis of tick-borne encephalitis and vaccines for its provision. Byulleten' Sibirskoy Meditsiny = Bulletin of Siberian Medicine. 2006;5(S1):63-71. (in Russian)]

Максимов В.Н., Малютина С.К., Орлов П.С., Иванощук Д.Е., Воропаева Е.Н., Бобак М., Воевода М.И. Длина теломер лейкоцитов как маркер старения и фактор риска развития возрастзависимых заболеваний у человека. Усп. геронтол. 2016;29(5):702-708.
[Maksimov V.N., Malyutina S.K., Orlov P.S., Ivanoshchuk D.E., Voropayeva E.N., Bobak M., Voevoda M.I. Length telomere leukocytes as aging markers and risk factors for age-related diseases in humans. Uspekhi Gerontologii = Advances in Gerontology. 2016;29(5):702-708. (in Russian)]

Медуницын Н.В. Вакцинология. М.: Триада-Х, 2004.
[Medunitsin N.V. Vaccinology. Moscow: Triada-X Publ., 2004. (in Russian)]

Онищенко Г.Г. Распространение вирусных природноочаговых инфекций в Российской Федерации и меры по их профилактике. Эпидемиол. и инф. бол. 2000;4:4-8.
[Onishchenko G.G. The spread of viral natural focal infections in the Russian Federation and measures for their prevention. Epidemiologiya i Infektsionnyye Bolezni = Epidemiology and Infectious Diseases. 2000;4:4-8. (in Russian)]

Спивак И.М., Михельсон В.М., Спивак Д.Л. Длина теломер, активность теломеразы, стресс и старение. Усп. геронтол. 2015;28(3): 441-448.
[Spivak I.M., Mikhelson V.M., Spivak D.L. Telomere length, telomerase activity, stress and aging. Uspekhi Gerontologii = Advances in Gerontology. 2015;28(3):441-448. (in Russian)]

Юдин Н.С., Бархаш А.В., Максимов В.Н., Игнатьева Е.В., Ромащенко А.Г. Генетическая предрасположенность человека к заболеваниям, вызываемым вирусами семейства Flaviviridae. Молекуляр. биология. 2018;52(2):190-209. DOI 10.7868/S0026898418020039.
[Yudin N.S., Barkhash A.V., Maksimov V.N., Ignatiyeva Ye.V., Romashchenko A.G. Human genetic predisposition to diseases caused by viruses of the Flaviviridae family. Molecular Biology (Moscow). 2018;52(2):165-181. DOI 10.1134/S0026893317050223.]

Astuti Y., Wardhana A., Watkins J., Wulaningsih W.; PILAR Research Network. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. Environ. Res. 2017;158:480-489. DOI 10.1016/j.envres.2017.06.038.

Bellon M., Nicot C. Telomere dynamics in immune senescence and exhaustion triggered by chronic viral infection. Viruses. 2017;9(10). DOI 10.3390/v9100289.

Blackburn E.H., Epel E.S., Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science. 2015;350(6265):1193-1198. DOI 10.1126/science.aab3389.

Cawthon R.M. Telomere measurement by quantitative PCR. Nucleic Acids Res. 2002;30(10):e47.

Cools H.J., Gussekloo J., Remmerswaal J.E., Remarque E.J., Kroes A.C. Benefits of increasing the dose of influenza vaccine in residents of long-term care facilities: a randomized placebo-controlled trial. J. Med. Virol. 2009;81(5):908-914. DOI 10.1002/jmv.21456.

de Punder K., Heim C., Wadhwa P.D., Entringer S. Stress and immunosenescence: The role of telomerase. Psychoneuroendocrinology. 2018;101:87-100. DOI 10.1016/j.psyneuen.2018.10.019.

Freitas-Simoes T.M., Ros E., Sala-Vila A. Nutrients, foods, dietary patterns and telomere length: Update of epidemiological studies and randomized trials. Metabolism. 2016;65(4):406-415. DOI 10.1016/j.metabol.2015.11.004.

Gardner M., Bann D., Wiley L., Cooper R., Hardy R., Nitsch D., Martin-Ruiz C., Shiels P., Sayer A.A., Barbieri M., Bekaert S., Bischoff C., Brooks-Wilson A., Chen W., Cooper C., Christensen K., De Meyer T., Deary I., Der G., Diez Roux A., Fitzpatrick A., Hajat A., Halaschek-Wiener J., Harris S., Hunt S.C., Jagger C., Jeon H.S., Kaplan R., Kimura M., Lansdorp P., Li C., Maeda T., Mangino M., Nawrot T.S., Nilsson P., Nordfjall K., Paolisso G., Ren F., Riabowol K., Robertson T., Roos G., Staessen J.A., Spector T., Tang N., Unryn B., van der Harst P., Woo J., Xing C., Yadegarfar M.E., Park J.Y., Young N., Kuh D., von Zglinicki T., Ben-Shlomo Y.; Halcyon study team. Gender and telomere length: systematic review and meta-analysis. Exp. Gerontol. 2014;51:15-27. DOI 10.1016/j.exger.2013.12.004.

Gritsun T.S., Lashkevich V.A., Gould E.A. Tick-borne encephalitis. Antiviral Res. 2003;57(1-2):129-146.

Gröndahl-Yli-Hannuksela K., Vahlberg T., Ilonen J., Mertsola J., He Q. Polymorphism of IL-10 gene promoter region: association with T cell proliferative responses after acellular pertussis vaccination in adults. Immunogenetics. 2016;68(9):733-741. DOI 10.1007/s00251-016-0923-0.

Haglund M., Günther G. Tick-borne encephalitis – pathogenesis, clinical course and long-term follow-up. Vaccine. 2003;21(Suppl. 1): S11-S18.

Haralambieva I.H., Dhiman N., Ovsyannikova I.G., Vierkant R.A., Pankratz V.S., Jacobson R.M., Poland G.A. 2′-5′-Oligoadenylate synthetase single-nucleotide polymorphisms and haplotypes are associated with variations in immune responses to rubella vaccine. Hum. Immunol. 2010;71(4):383-391. DOI 10.1016/j.humimm.2010.01.004.

Harley C.B., Futcher A.B., Greider C.W. Telomeres shorten during ageing of human fibroblasts. Nature. 1990;345(6274):458-460.

Heinz F.X., Kunz C. Tick-borne encephalitis and the impact of vaccination. Arch. Virol. Suppl. 2004;18:201-205.

Hodes R.J., Hathcock K.S., Weng N.P. Telomeres in T and B cells. Nat. Rev. Immunol. 2002;2(9):699-706.

Jääskeläinen A., Tonteri E., Pieninkeroinen I., Sironen T., Voutilainen L., Kuusi M., Vaheri A., Vapalahti O. Siberian subtype tick-borne encephalitis virus in *Ixodes ricinus* in a newly emerged focus, Finland. Ticks Tick Borne Dis. 2016;7(1):216-223. DOI 10.1016/j.ttbdis.2015.10.013.

Karelis G., Bormane A., Logina I., Lucenko I., Suna N., Krumina A., Donaghy M. Tick-borne encephalitis in Latvia 1973–2009: epidemiology, clinical features and sequelae. Eur. J. Neurol. 2012;19(1): 62-68. DOI 10.1111/j.1468-1331.2011.03434.x.

Laberthonniere C., Magdinier F., Robin J.D. Bring it to an end: does telomeres size matter? Cells. 2019;8(1):pii: E30. DOI 10.3390/cells8010030.

Lin Y., Damjanovic A., Metter E.J., Nguyen H., Truong T., Najarro K., Morris C., Longo D.L., Zhan M., Ferrucci L., Hodes R.J., Weng N.P. Age-associated telomere attrition of lymphocytes *in vivo* is co-ordinated with changes in telomerase activity, composition of lymphocyte subsets and health conditions. Clin. Sci. (Lond.). 2015;128(6): 367-377. DOI 10.1042/CS20140481.

Mukherjee A.K., Sharma S., Sengupta S., Saha D., Kumar P., Hussain T., Srivastava V., Roy S.D., Shay J.W., Chowdhury S. Telomere length-dependent transcription and epigenetic modifications in promoters remote from telomere ends. PLoS Genet. 2018;14(11):e1007782. DOI 10.1371/journal.pgen.1007782.

Muszkat M., Friedman G., Dannenberg H.D., Greenbaum E., Lipo M., Heymann Y., Zakay-Rones Z., Ben-Yehuda A. Response to influenza vaccination in community and in nursing home residing elderly: relation to clinical factors. Exp. Gerontol. 2003;38(10):1199-1203.

Najarro K., Nguyen H., Chen G., Xu M., Alcorta S., Yao X., Zukley L., Metter E.J., Truong T., Lin Y., Li H., Oelke M., Xu X., Ling S.M., Longo D.L., Schneck J., Leng S., Ferrucci L., Weng N.P. Telomere length as an indicator of the robustness of B- and T-cell response to influenza in older adults. J. Infect. Dis. 2015;212(8):1261-1269. DOI 10.1093/infdis/jiv202.

Rizvi S., Raza S.T., Mahdi F. Telomere length variations in aging and age-related diseases. Curr. Aging Sci. 2014;7(3):161-167.

Robin J.D., Ludlow A.T., Batten K., Magdinier F., Stadler G., Wagner K.R., Shay J.W., Wright W.E. Telomere position effect: regulation of gene expression with progressive telomere shortening over long distances. Genes Dev. 2014;28(22):2464-2476. DOI 10.1101/gad.251041.114.

Sagawa M., Kojimahara N., Otsuka N., Kimura M., Yamaguchi N. Immune response to influenza vaccine in the elderly: association with nutritional and physical status. Geriatr. Gerontol. Int. 2011;11(1): 63-68. DOI 10.1111/j.1447-0594.2010.00641.x.

Sambrook J., Russell D.W. The Condensed Protocols from Molecular Cloning: a Laboratory Manual, Cold Spring Harbor; New York: Cold Spring Harbor Laboratory Press, 2006.

Shawi M., Autexier C. Telomerase, senescence and ageing. Mech. Ageing Dev. 2008;129(1-2):3-10. DOI 10.1016/j.mad.2007.11.007.

Son N.H., Joyce B., Hieatt A., Chrest F.J., Yanovski J., Weng N.P. Stable telomere length and telomerase expression from naïve to memory B-lymphocyte differentiation. Mech. Ageing Dev. 2003;124(4): 427-432.

Steffen R. Tick-borne encephalitis (TBE) in children in Europe: Epidemiology, clinical outcome and comparison of vaccination recommendations. Ticks Tick Borne Dis. 2019;10(1):100-110. DOI 10.1016/j.ttbdis.2018.08.003.

Valarcher J.F., Hagglund S., Juremalm M., Blomqvist G., Renstrom L., Zohari S., Leijon M., Chirico J. Tick-borne encephalitis. Rev. Sci. Tech. 2015;34(2):453-466.

Veje M., Nolskog P., Petzold M., Bergström T., Lindén T., Peker Y., Studahl M. Tick-borne encephalitis sequelae at long-term follow-up: a self-reported case-control study. Acta Neurol. Scand. 2016; 134(6):434-441. DOI 10.1111/ane.12561.

Verschoor C.P., Lelic A., Parsons R., Evelegh C., Bramson J.L., Johnstone J., Loeb M.B., Bowdish D.M.E. Serum C-reactive protein and congestive heart failure as significant predictors of herpes zoster vaccine response in elderly nursing home residents. J. Infect. Dis. 2017;216(2):191-197. DOI 10.1093/infdis/jix257.

Wang J., Dong X., Cao L., Sun Y., Qiu Y., Zhang Y., Cao R., Covasa M., Zhong L. Association between telomere length and diabetes mellitus: A meta-analysis. J. Int. Med. Res. 2016;44(6):1156-1173. DOI 10.1177/0300060516667132.

Weng N.P. Telomeres and immune competency. Curr. Opin. Immunol. 2012;24(4):470-475. DOI 10.1016/j.coi.2012.05.001.

Weng N.P., Granger L., Hodes R.J. Telomere lengthening and telomerase activation during human B cell differentiation. Proc. Natl. Acad. Sci. USA. 1997;30;94(20):10827-10832.

Weng N.P., Levine B.L., June C.H., Hodes R.J. Human naive and memory T lymphocytes differ in telomeric length and replicative potential. Proc. Natl. Acad. Sci. USA. 1995;92(24):11091-11094.

Willis M., Reid S.N., Calvo E., Staudinger U.M., Factor-Litvak P. A scoping systematic review of social stressors and various measures of telomere length across the life course. Ageing Res. Rev. 2018;47:89-104. DOI 10.1016/j.arr.2018.07.006.

Yudin N.S., Igoshin A.V., Lutova S.L., Gon Ya., Voevoda M.I., Belyavskaya V.A. Association between polymorphisms in genes encoding 2′-5′-oligoadenylate synthetases and the humoral immune response upon vaccination against tick-borne encephalitis. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2018;22(4):445-451. DOI 10.18699/VJ18.381. (in Russian)

ORCID ID
N.S. Yudin orcid.org/0000-0002-1947-5554
D.E. Ivanoshchuk orcid.org/0000-0002-0403-545X
M.I. Voevoda orcid.org/0000-0001-9425-413X

Благодарности. Исследование выполнено за счет гранта Российского научного фонда (проект № 16-15-00127).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 03.12.2018. После доработки 06.08.2019. Принята к публикации 07.08.2019.

БИОИНФОРМАТИКА И СИСТЕМНАЯ БИОЛОГИЯ
Оригинальное исследование / Original article
УДК 519.246.87:517.968.233

Вавиловский журнал генетики и селекции. 2019;23(8):1032-1036
DOI 10.18699/VJ19.584

Principal component analysis and its generalizations for any type of sequence (PCA-Seq)

V.M. Efimov[1, 2, 3, 4] ✉, K.V. Efimov[5], V.Y. Kovaleva[2]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Institute of Systematics and Ecology of Animals, SB RAS, Novosibirsk, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
[4] Tomsk State University, Tomsk, Russia
[5] Moscow Institute of Physics and Technology (State University), Moscow, Russia
✉ e-mail: efimov@bionet.nsc.ru

In the 1940s, Karhunen and Loève proposed a method for processing a one-dimensional numeric time series by converting it into multidimensional by shifts. In fact, a one-dimensional number series was decomposed into several orthogonal time series. This method has many times been independently developed and applied in practice under various names (EOF, SSA, Caterpillar, etc.). Nowadays, the name 'SSA' (Singular Spectral Analysis) is the most often used. It turned out that it is universal, applicable to any time series without requiring stationary assumptions, automatically decomposes time series into a trend, cyclic components and noise. By the beginning of the 1980s, Takens had shown that for a dynamical system such a method makes it possible to obtain an attractor from observing only one of these variables, thereby bringing the method to a powerful theoretical basis. In the same years, the practical benefits of phase portraits became clear. In particular, it was used in the analysis and forecast of animal abundance dynamics. In this paper we propose to extend SSA to a one-dimensional sequence of any type of elements, including numbers, symbols, figures, etc., and, as a special case, to a molecular sequence. Technically, the problem is solved using an algorithm like SSA. The sequence is cut by a sliding window into fragments of a given length. Between all fragments, the matrix of Euclidean distances is calculated. This is always possible. For example, the square root of the Hamming distance between fragments is a Euclidean distance. For the resulting matrix, the principal components are calculated by the principal-coordinate method (PCo). Instead of a distance matrix, one can use a matrix of any similarity/dissimilarity indexes and apply methods of multidimensional scaling (MDS). The result will always be PCs in some Euclidean space. We called this method 'PCA-Seq'. It is certainly an exploratory method, as is its particular case SSA. For any sequence, including molecular, PCA-Seq without any additional assumptions allows presenting its principal components in a numerical form and visualizing them in the form of phase portraits. A long history of SSA application for numerical data gives all reason to believe that PCA-Seq will be not less useful in the analysis of non-numerical data, especially in hypothesizing. PCA-Seq is implemented in the freely distributed Jacobi 4 package (http://jacobi4.ru/).
Key words: time series; SVD; PCA; PCo; MDS; SSA; molecular sequences; *p*-distance.

For citation: Efimov V.M., Efimov K.V., Kovaleva V.Y. Principal component analysis and its generalizations for any type of sequence (PCA-Seq). Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1032-1036. DOI 10.18699/VJ19.584

Метод главных компонент и его обобщения для последовательности любого типа (PCA-Seq)

В.М. Ефимов[1, 2, 3, 4] ✉, К.В. Ефимов[5], В.Ю. Ковалева[2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт систематики и экологии животных Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Новосибирский государственный университет, Новосибирск, Россия
[4] Томский государственный университет, Томск, Россия
[5] Московский физико-технический институт (государственный университет), Москва, Россия
✉ e-mail: efimov@bionet.nsc.ru

В 1940-х гг. К. Карунен и М. Лоев предложили метод обработки одномерного числового временного ряда через его преобразование в многомерный путем сдвига несколько раз подряд и разложения на несколько ортогональных временных рядов методом главных компонент (PCA). Предложенный метод ранее независимо возникал и применялся на практике под разными названиями (EOF, SSA, Гусеница и т. д.). Оказалось, что он универсальный, применим к любому временному ряду и, не требуя предположения стационарности, автоматически разлагает его на тренд, циклические составляющие и шум. В наши дни чаще всего используется название SSA (сингулярный спектральный анализ). В начале 1980-х гг. Ф. Такенс показал, что для динамической системы сдвиги только одной наблюдаемой переменной позволяют построить аттрактор всей системы, и тем самым подвел под SSA мощную теоретическую базу. Тогда же выяснилась практическая польза фазовых портретов, что было применено, в частности, при анализе и прогнозе динамики численности животных. В настоящей работе предлагается распространить SSA на одномерную последовательность элементов любого типа, включая числа, символы, фигуры и т. д., и в качестве частного случая –

© Efimov V.M., Efimov K.V., Kovaleva V.Y., 2019

на молекулярную последовательность. Технически проблема решается практически тем же алгоритмом, что и SSA. Последовательность режется скользящим окном на фрагменты заданной длины. Между всеми фрагментами вычисляется матрица евклидовых расстояний. Это всегда возможно. Например, квадратный корень из *p*-дистанции (дистанции Хэмминга) является евклидовым расстоянием. Для полученной матрицы методом главных координат (PCo) вычисляются главные компоненты. Вместо расстояний можно использовать любые индексы сходства/различия и применить методы многомерного шкалирования (MDS). В итоге все равно будут получены главные компоненты в некотором евклидовом пространстве. Мы назвали этот метод PCA-Seq. Это, безусловно, разведочный метод, как и его частный случай SSA. Для любой последовательности, в том числе молекулярной, PCA-Seq без всяких дополнительных предположений позволяет получить ее главные компоненты в числовом виде и визуализировать их в виде графиков и фазовых портретов. Многолетний опыт применения SSA для числовых данных дает все основания полагать, что PCA-Seq окажется не менее полезным при анализе нечисловых данных, особенно при выдвижении гипотез. PCA-Seq реализован в свободно распространяемом пакете Jacobi 4 (http://jacobi4.ru/).
Ключевые слова: временные ряды; SVD; PCA; PCo; MDS; SSA; молекулярные последовательности; *p*-дистанция.

Introduction

In the 1940s, Karhunen and Loève proposed a method for processing a one-dimensional numerical time series by shifting it several times and decomposing into several orthogonal time series by a multidimensional method of principal components (PCA) (Karhunen, 1947; Loève, 1948). In the 1980s, Takens showed for a dynamic system, that shifts of only one observed variable allow constructing an attractor of the entire system, thereby bringing the method to a powerful theoretical basis (Takens, 1981).

The method was independently developed and applied in practice under various names (EOF, SSA, Caterpillar, etc.), including by us for the analysis of animals abundance dynamics (Efimov, Galaktionov, 1983; Efimov et al., 1988, 2003), and for other topics (Golyandina et al., 2001, 2018; Golyandina, Zhigljavsky, 2013). Today the name 'SSA' (Singular Spectral Analysis) is the most often used. The method can be extended for a sequence of any type of elements, including numbers, symbols, figures, etc. and, as a special case, for a molecular sequence (Efimov et al., 2018). This is the point of this article.

Material and methods

Algorithm. Let there be a sequence $Q = \{q_1, q_2, \ldots, q_N\}$ of any type of elements. Choose a lag L, $N > L > 1$. Denote by Q_i the fragment Q of length L terminated by the element q_i, $Q_i = (q_{i-L+1}, q_{i-L+2}, \ldots, q_{i-1}, q_i)$, $i = L, \ldots, N$. Compute the matrix of Euclidean distances $D = (d_{ij} = d(Q_i, Q_j))$ between all fragments (this is always possible, for example, using the number of unmatched elements, but not only that). Apply the method of principal coordinates (PCo) to D and obtain its principal components PCs (Gower, 1966). Call this method 'PCA-Seq'.

The usual method of finding principal components consists in the following (Jolliffe, Cadima, 2016). Let X be a centered matrix of objects' coordinates in a certain Euclidean space. We can apply to X the singular value decomposition (SVD): $X = PSV^T$, where P, V^T are orthogonal matrices, and S is the diagonal matrix of X singular values. It is possible to apply SVD to a symmetric matrix XX^T: $XX^T = P\Lambda P^T$, where P is the same orthogonal matrix as for X, and Λ is a diagonal matrix of the matrix XX^T singular values. But $XX^T = PSV^T \times VSP^T = PSSP^T = PS^2P^T$. Consequently, $S^2 = \Lambda$ and $S = \Lambda^{1/2}$. That is, the matrix of singular values of the matrix XX^T is the matrix of eigenvalues of the matrix X. Therefore, it is necessary to calculate the principal components by the formula $U = P\Lambda^{1/2}$. This is very useful in practice if the number of objects is significantly lower than the number of traits that are becoming more common in biological research, especially molecular ones.

More than half a century ago, Gower (Gower, 1966) found that if we calculate the matrix D of Euclidean distances between rows X, square these distances, double center and multiply them by –1/2, then we obtain the XX^T matrix. Applying SVD to it, we obtain principal components. For this reason, Gower called this method the 'principal coordinates (PCo) analysis'. However, it follows from the results of Gower that the matrix X itself is not needed and may not even exist in numerical form. To calculate the principal components of a certain set of objects, it is enough to have a matrix of Euclidean distances between them obtained no matter which way. If we calculate the Euclidean matrix of distances between the rows of the matrix of principal components, then it will coincide with the initial matrix of Euclidean distances D. This property can be used to verify the calculations.

PCo is quite often used for dissimilarity matrices, for which it is unknown whether they are Euclidean distances between objects or not. In the case of non-Euclidean distances, some diagonal values of the matrix Λ will be negative. Small negative diagonal values can sometimes arise due to the accumulation of computational errors. All such "components", as well as zero ones, should be excluded from consideration.

Instead of a distance matrix, one can use a matrix of any coefficients of similarity/dissimilarity. In this case, it is necessary to apply methods of multidimensional scaling (MDS). The results will always be the PCs in some Euclidean space (Gower, 1966). PCo has another name: metric multidimensional scaling abbreviated as MMDS or simply MDS. It is more correct to call all the multidimensional scaling methods 'MDS', and apply 'MMDS' to PCo only.

Data. The amino acid sequence of the *Homo sapiens Cytb* gene was used (Q0ZFD6_HUMAN, Swiss Model repository) (Table 1). The sequence (length N = 380) contains two chains at positions 19–204, 259–359 and nine transmembrane helices at positions 30–56, 77–98, 113–133, 140–158, 178–200, 229–246, 288–308, 320–339, and 345–372 (https://swissmodel.expasy.org/repository/uniprot/Q0ZFD6).

Processing. Denote $N-L+1$ by N_L. For L = 2, … , 24, the sequence Q0ZFD6 represented as Seq_L matrix of size $N_L \times L$ (Table 2 with L = 8, as an example). For each Seq_L, the matrix A_L of size $N_L \times 20$ – each amino acid content in the fragment and the H_L vector of length N_L – the fraction of fragment positions coinciding with transmembrane helices are additionally calculated. For all matrices Seq_L, matrices

Table 1. The amino acid sequence of the *Homo sapiens Cytb* gene
(Q0ZFD6_HUMAN, Swiss Model repository, https://swissmodel.expasy.org/repository/uniprot/Q0ZFD6)

>tr|Q0ZFD6|Q0ZFD6_HUMAN Cytochrome b OS=Homo sapiens GN=CYB

MTPMRKTNPLMKLINHSFIDLPTPSNISAWWNFGSLLGACLILQITTGLFLAMHYSPDASTAFSSIAHITRDVNYGWIIRYLHANGASMFFICLFLHIGRGLYYGS
FLYSETWNIGIILLLATMATAFMGYVLPWGQMSFWGATVITNLLSAIPYIGTDLVQWIWGGYSVDSPTLTRFFTFHFILPFIIAALATLHLLFLHETGSNNPLGITSH
SDKITFHPYYTIKDALGLLLFLLSLMTLTLFSPDLLGDPDNYTLANPLNTPPHIKPEWYFLFAYTILRSVPNKLGGVLALLLSILILAMIPILHMSKQQSMMFRPLSQ
SLYWLLAADLLILTWIGGQPVSYPFTIIGQVASVLYFTTILILMPTISLI**ENKMLKWA**

Note: The top line is the sequence identifier, the lower line is the sequence itself. The first and last fragments of length 8 are highlighted in bold type (see Table 2).

Table 2. Takens embedding transformation
of the amino acid sequence from Table 1

i	q_{i-7}	q_{i-6}	q_{i-5}	q_{i-4}	q_{i-3}	q_{i-2}	q_{i-1}	q_i
8	**M**	**T**	**P**	**M**	**R**	**K**	**T**	**N**
9	T	P	M	R	K	T	N	P
10	P	M	R	K	T	N	P	L
11	M	R	K	T	N	P	L	M
12	R	K	T	N	P	L	M	K
13	K	T	N	P	L	M	K	L
14	T	N	P	L	M	K	L	I
15	N	P	L	M	K	L	I	N
16	P	L	M	K	L	I	N	H
17	L	M	K	L	I	N	H	S
18	M	K	L	I	N	H	S	F
19	K	L	I	N	H	S	F	I
20	L	I	N	H	S	F	I	D
21	I	N	H	S	F	I	D	L
22	N	H	S	F	I	D	L	P
23	H	S	F	I	D	L	P	T
24	S	F	I	D	L	P	T	P
25	F	I	D	L	P	T	P	S
26	I	D	L	P	T	P	S	N
27	D	L	P	T	P	S	N	I
28	L	P	T	P	S	N	I	S
29	P	T	P	S	N	I	S	A
30	T	P	S	N	I	S	A	W
…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…
379	I	E	N	K	M	L	K	W
380	**E**	**N**	**K**	**M**	**L**	**K**	**W**	**A**

Note: The first and last rows are the first and last fragments of the sequence, the same as in Table 1. The number of a fragment is defined by the number of its last amino acid, therefore the table begins with row 8.

of Euclidean distances between the fragments are calculated (square root of the *p*-distance (Hamming distance) between a couple of fragments is used as the Euclidean distance (Efimov et al., 2013)).

For all matrices of Euclidean distances, its PCs are calculated by the method of PCo (Gower, 1966). The matrices PC-Seq_L, A_L and the vector H_L were combined into one matrix, and for it, the matrix of Pearson correlation coefficients was calculated between all columns. Only the correlation coefficients exceeding a threshold of 0.316 in absolute magnitude ($r^2 \sim 0.1$, i. d. 10 %; $p < 10^{-8}$) were considered. Jacobi 4 package was used for calculations (Polunin et al., 2014).

Results

For the first principal component PC1-Seq_L, the correlation exceeding the threshold was found with a fraction of helix positions ($0.370 \le r \le 0.547$ in the range $4 \le L \le 18$; $r_{max} = 0.547$ for L = 12), leucine content in the fragment ($0.95 \le r$ in the range $2 \le L \le 17$; $r_{12} = 0.974$), proline content ($0.331 \le r \le 0.364$ in the range $9 \le L \le 14$) and tyrosine content ($0.318 \le r \le 0.351$ in the range $14 \le L \le 18$), what is more, correlations PC1-Seq_L with the contents of proline and tyrosine have inverse signs in relation of correlations with the fraction of helix positions and the leucine content in the fragments. The graph of PC1-Seq_L against the background of the fraction of helix positions is shown in Fig. 1, the dynamics of the leucine content against the same background is in Fig. 2, the scatterplot of PC1-Seq_L vs the leucine content in the fragment is shown in Fig. 3 (all for L = 12).

Discussion

The good correlation of the first principal component with a fraction of helix positions means that the similarity of the fragments depends on how much the fragments intersect with α-helices. It is known that hydrophobic amino acids are most often found in α-helices, and hydrophilic ones are outside them. Leucine is a hydrophobic amino acid, and indeed it appears in α-helices of humans more often than other amino acids. This explains the high correlation between the first principal component and the leucine content in the fragments. Note that we did not specifically look for any information about the amino acids content in the fragments or about the secondary structure of the sequence. The only thing that we investigated was how much the fragments coincided with each other by amino acids in total for all L positions. If we had set another measure of similarity, then perhaps we would have discovered some other regularity. In this case, this one is found.

PCA-Seq is certainly an exploratory method, as is its particular case SSA. For any sequence, including molecular, PCA-Seq without any additional assumptions allows obtaining its principal components in a numerical form and visualizing them in the form of phase portraits. Today SSA for numerical series is a huge scientific field with applications in various sciences. There is no doubt that the analysis of non-numeric sequences will become a scientific field, no less in scope than SSA.

It should be noted that the approach of calculation is used in the standard SSA through a covariation (correlation) matrix

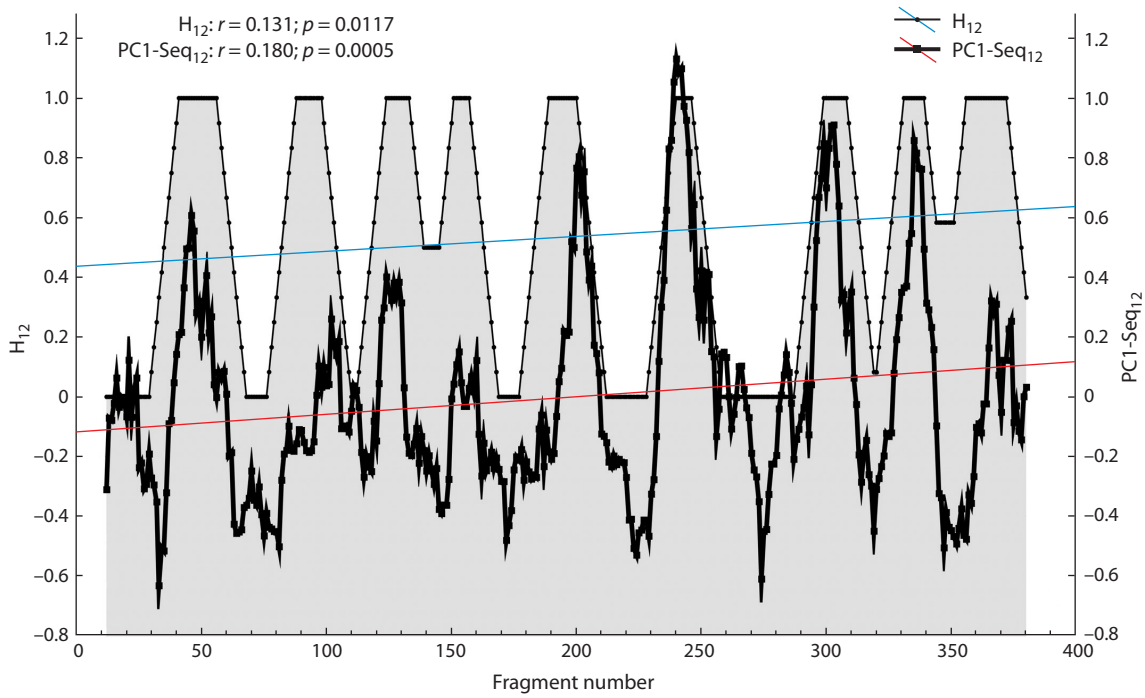


Fig. 1. PC1-Seq_{12} is first principal component of the *Homo sapiens Cytb* amino acid sequence (Q0ZFD6_HUMAN, Swiss Model repository) and H_{12} is the fraction of helix position in sequence fragments of length L = 12 ($r = 0.547$, N = 369, $p < 10^{-9}$).

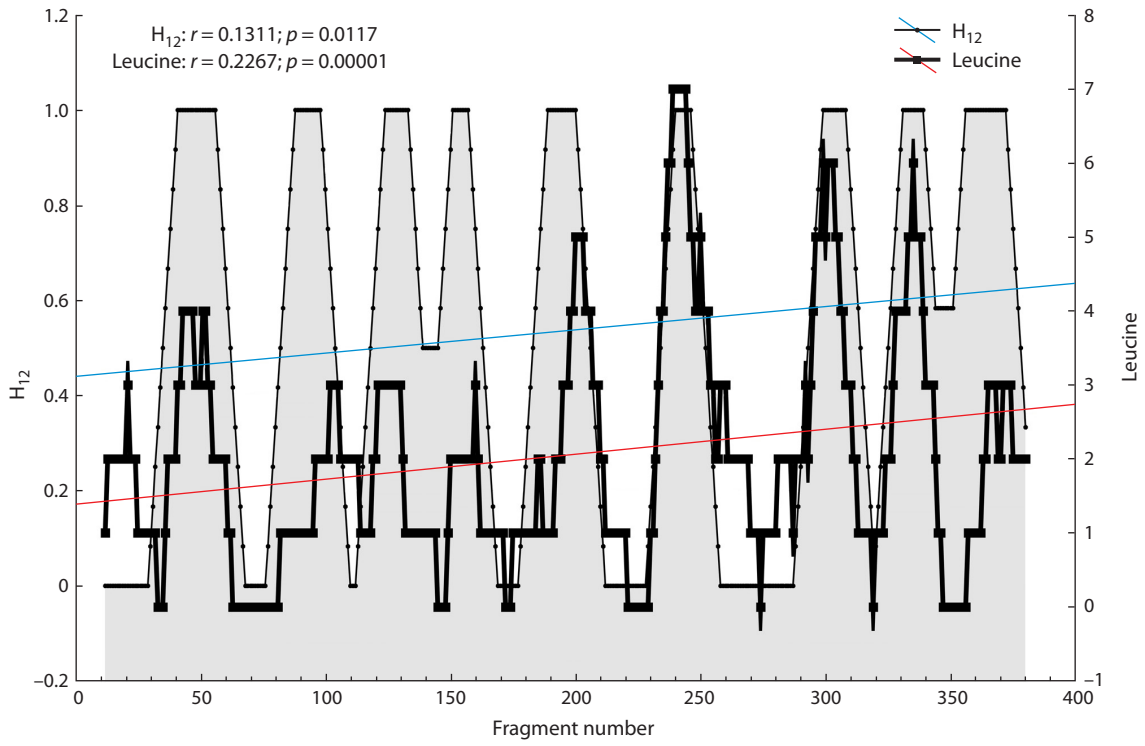


Fig. 2. Content of leucine in *Homo sapiens Cytb* amino acid sequence fragments (Q0ZFD6_HUMAN, Swiss Model repository) and H_{12} is the fraction of helix position in sequence fragments of length L = 12 ($r = 0.499$, N = 369, $p < 10^{-9}$).

only, and the MDS methods, including PCo, despite more than half a century-long history, are almost unknown. This gives reason to hope that PCA-Seq can be useful in the analysis of real data, especially in hypothesizing. PCA-Seq is a particular case of the geometric approach, in which any similarity/dissimilarity relations between objects are modeled by the distance between points in a certain Euclidean space. In this case, the objects are fragments of any sequence by length L, including non-numeric, in particular molecular one. Orthogonal rotations of the entire set of points leave the distances

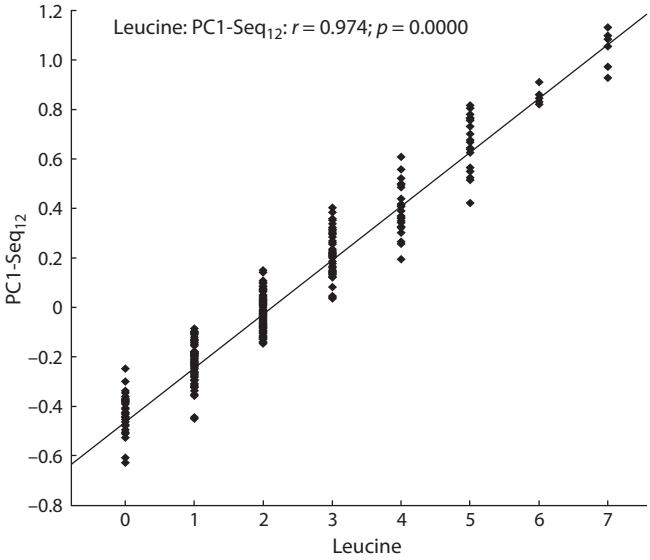


Fig. 3. Content of leucine in *Homo sapiens Cytb* amino acid sequence fragments (Q0ZFD6_HUMAN, Swiss Model repository) vs first principal component PC1-Seq_{12} (L = 12, r = 0.974, N = 369, $p < 10^{-9}$).

between them unchanged. This allows calculating such axes, in the projection on which the maximum variances of the set of points are reached. These axes always exist, they are exactly the principal components. By construction, they are not statistically correlated with each other.

This does not mean at all that they are meaningfully independent. In particular, for a time series, it is a general rule that their PCs, despite being uncorrelated, break up into couples in which one component is derivative of another. When one component shifts from another by a quarter of a period, the correlation appears again. In successful cases, this allows predicting the future values of one component from the already known values of another and thus predicting, to some extent, the initial series. The sine/cosine couple is a telling example. Moreover, it is possible that there is a third component, as a rule, a trend, which, despite the lack of correlation with the first two, modulates their amplitude. Thus, it is a part of a general interconnected complex. In 3D phase space such components form a funnel.

PCA-Seq is implemented in the freely distributed Jacobi 4 package (http://jacobi4.ru/).

Conclusion

PCA-Seq is promising for processing molecular sequences – and then some.

References

Efimov V.M., Galaktionov Y.K. On the possibility of predicting cyclic changes in the abundance of mammals. Zhurnal Obshchey Biologii = Journal of General Biology. 1983;3:343-352. (in Russian)

Efimov V.M., Galaktionov Y.K., Galaktionova T.A. Reconstruction and prognosis of water vole population dynamics on the basis of tularemia morbidity among Novosibirsk oblast residents. Doklady. Biological Sciences. 2003;388(1/6):59-61.

Efimov V.M., Galaktionov Y.K., Shushpanova N.F. Analysis and Prediction of Time Series by the Principal Component Method. Novosibirsk: Nauka Publ., 1988. (in Russian)

Efimov V.M., Kovaleva V.Y., Efimov K.V. Principal Component Analysis for any type Sequences (PCA-Seq). In: Mathematical Modeling and High-Performance Computing in Bioinformatics, Biomedicine and Biotechnology (MM-HPC-BBB-2018): Proc. of the 3rd Int. Symp. Novosibirsk, 21–24 Aug 2018. Novosibirsk, 2018;20.

Efimov V.M., Melchakova M.A., Kovaleva V.Y. Geometric properties of evolutionary distances. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2013;17(4/1):714-723. (in Russian)

Golyandina N., Korobeynikov A., Zhigljavsky A. Singular Spectrum Analysis with R. (Ser. Use R!) Berlin; Heidelberg: Springer Verlag, 2018.

Golyandina N., Nekrutkin V., Zhigljavsky A.A. Analysis of Time Series Structure: SSA and Related Techniques. Chapman and Hall/CRC, 2001.

Golyandina N., Zhigljavsky A. Singular Spectrum Analysis for Time Series. Springer Science & Business Media, 2013.

Gower J.C. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika. 1966;53(3/4):325-338.

Jolliffe I.T., Cadima J. Principal component analysis: a review and recent developments. Phil. Trans. R. Soc. A. 2016;374:20150202.

Karhunen K. Über lineare methoden in der wahrscheinlich-keitsrechnung. Ann. Acad. Sci. Fennicea. 1947;Ser. A137.

Loève M. Fonctions Aléatoires de second order. In: Lévy P. (Ed.). Processus Stochastiques et Movement Brownien. Paris: Hermann, 1948.

Polunin D.A., Shtaiger I.A., Efimov V.M. Development of software system JACOBI 4 for multivariate analysis of microarray data. Vestnik Novosibirskogo Gosudarstvennogo Universiteta. Seriya Informatsyonnye Tekhnologii = Vestnik NSU. Information Technology. 2014;12(2):90-98. (in Russian)

Takens F. Detecting strange attractors in turbulence. In: Dynamical Systems and Turbulence. Warwick, 1980. Berlin; Heidelberg: Springer, 1981;366-381.

Acknowledgements. Supported by Russian Foundation for Basic Research (# 19-07-00658). The authors are grateful to D.A. Afonnikov, P.N. Menshanov and two anonymous reviewers for useful discussion and constructive comments.

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

Received September 19, 2018. Revised June 28, 2019. Accepted July 2, 2019.

БИОИНФОРМАТИКА И СИСТЕМНАЯ БИОЛОГИЯ
Оригинальное исследование / Original article
УДК 612.172.1:575.113

Вавиловский журнал генетики и селекции. 2019;23(8):1037-1046
DOI 10.18699/VJ19.585

In silico поиск генов, контролирующих ишемическую болезнь сердца

И.В. Зоркольцева[1] ✉, Н.М. Белоногова[1], Г.Р. Свищёва[1, 2], А.В. Кириченко[1], Т.И. Аксенович[1, 3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ e-mail: zor@bionet.nsc.ru

К настоящему времени с помощью крупномасштабных полногеномных исследований обнаружено более 100 локусов, ассоциированных с ишемической болезнью сердца (ИБС). Для некоторых из нескольких сотен генов, лежащих в этих локусах, была показана их роль в патогенезе болезни. Тем не менее основные генетические механизмы и конкретные гены, контролирующие заболевание, все еще полностью не известны. Данное исследование посвящено *in silico* поиску новых генов, контролирующих ИБС. Проведен региональный анализ ассоциаций, при котором все полиморфные варианты внутри гена анализируются одновременно. Материалом для анализа служили результаты полногеномного анализа ассоциаций, депонированные в открытых базах данных MICAD (120575 человек, 85112 маркеров) и UK Biobank (337199 человек, 10894597 маркеров). Использовался программный пакет sumFREGAT, в котором реализован широкий спектр новых методов для тестирования генных ассоциаций с помощью суммарных статистик. Всего было обнаружено 88 генов. Из них 44 являются уже известными для ИБС генами. Кроме того, в известных локусах нами идентифицировано 28 дополнительных генов, которые можно рассматривать в качестве новых генов-кандидатов. 16 генов (*AGPAT4*, *ARHGEF12*, *BDP1*, *DHX58*, *EHBP1*, *FBF1*, *HSPB9*, *NPBWR2*, *PDLIM5*, *PLCB3*, *PLEKHM2*, *POU2F3*, *PRKD2*, *TMEM136*, *TTC29* и *UTP20*), обнаруженных нами вне известных локусов, являются новыми. Информация о функциональной роли этих генов позволяет рассматривать многие из них в качестве кандидатов для ИБС. Среди набора идентифицированных генов 41 ген не имел значимых сигналов полногеномного анализа ассоциаций и был идентифицирован только благодаря одновременному рассмотрению всех вариантов внутри гена в рамках регионального анализа ассоциаций. Полученные результаты демонстрируют, что региональный анализ ассоциаций представляет собой мощный инструмент для поиска новых генов. Он применим для анализа различных признаков и болезней с использованием накопленных в мире огромных объемов данных, полученных с помощью полногеномного анализа ассоциаций. Такие исследования общедоступны, поскольку не требуют дополнительных материальных затрат.
Ключевые слова: ишемическая болезнь сердца; полногеномный анализ ассоциаций; региональный анализ ассоциаций; суммарные статистики; *in silico* картирование.

Для цитирования: Зоркольцева И.В., Белоногова Н.М., Свищёва Г.Р., Кириченко А.В., Аксенович Т.И. *In silico* поиск генов, контролирующих ишемическую болезнь сердца. Вавиловский журнал генетики и селекции. 2019;23(8): 1037-1046. DOI 10.18699/VJ19.585

In silico mapping of coronary artery disease genes

I.V. Zorkoltseva[1] ✉, N.M. Belonogova[1], G.R. Svishcheva[1, 2], A.V. Kirichenko[1], T.I. Axenovich[1, 3]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Vavilov Institute of General Genetics, RAS, Moscow, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
✉ e-mail: zor@bionet.nsc.ru

To date, more than 100 loci associated with coronary artery disease (CAD) have been detected in large-scale genome-wide studies. For some of the several hundreds of genes located in these loci, roles in the pathogenesis of the disease have been shown. However, the genetic mechanisms and specific genes controlling this disease are still not fully understood. This study is aimed at *in silico* search for new CAD genes. We performed a gene-based association analysis, where all polymorphic variants within a gene are analyzed simultaneously. The analysis was based on the results of the genome-wide association studies (GWAS) available from the open databases MICAD (120,575 people, 85,112 markers) and UK Biobank (337,199 people, 10,894,597 markers). We used the sumFREGAT package implementing a wide range of new methods for gene-based association analysis using summary statistics. We found 88 genes demonstrating significant gene-based associations. Forty-four of the identified genes were already known as CAD genes. Furthermore, we identified 28 additional genes in the known CAD loci. They can be considered as new candidate genes. Finally, we identified sixteen new genes (*AGPAT4*, *ARHGEF12*, *BDP1*, *DHX58*, *EHBP1*, *FBF1*, *HSPB9*, *NPBWR2*, *PDLIM5*, *PLCB3*, *PLEKHM2*, *POU2F3*, *PRKD2*, *TMEM136*, *TTC29* and *UTP20*) outside the known loci. Information about the functional role of these genes allows us to consider many of them as candidates for CAD. The 41 identified genes did not have significant GWAS signals and they were identified only due to simultaneous consideration of all variants within the gene in the

© Зоркольцева И.В., Белоногова Н.М., Свищёва Г.Р., Кириченко А.В., Аксенович Т.И., 2019

framework of gene-based analysis. These results demonstrate that gene-based association analysis is a powerful tool for gene mapping. The method can utilize huge amounts of GWAS results accumulated in the world to map different traits and diseases. This type of studies is widely available, as it does not require additional material costs.
Key words: coronary artery disease; gene-based association analysis; genome-wide association analysis; summary statistics; *in silico* mapping.

For citation: Zorkoltseva I.V., Belonogova N.M., Svishcheva G.R., Kirichenko A.V., Axenovich T.I. *In silico* mapping of coronary artery disease genes. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019; 23(8):1037-1046. DOI 10.18699/VJ19.585 (in Russian)

Введение

Ишемическая болезнь сердца (ИБС) – одна из основных причин потери трудоспособности и преждевременной смерти во всем мире. ИБС является следствием атеросклероза коронарных артерий, которые питают миокард и снабжают его кислородом. Согласно «Международной статистической классификации болезней, травм и причин смерти» (http://apps.who.int/classifications/icd10/browse/2010/en#/I51.6), основные формы ИБС включают стенокардию, инфаркт миокарда (ИМ), кардиосклероз, нарушение сердечного ритма и проводимости, сердечную недостаточность и внезапную коронарную смерть. Клинические проявления ИБС зависят от ее формы: ИМ и внезапная коронарная смерть развиваются в течение нескольких минут, а постинфарктный кардиосклероз – хроническое состояние, которое стабильно при адекватно подобранном лечении. ИБС часто сопровождается гипертонией, диабетом второго типа, гиперхолестеринемией, так как эти заболевания имеют общие факторы риска и патогенетические механизмы (Ehret et al., 2011; Johnson et al., 2011). Различные метаболические и физиологические расстройства, имеющиеся при перечисленных заболеваниях, также содействуют повышению риска ИБС. Развитие ИБС часто проходит на фоне нездорового образа жизни, связанного с курением, неправильным питанием, ожирением, недостатком физической активности и др. (Yusuf et al., 2004).

Известно, что наследуемость ИБС составляет 40–50 % (Peden, Farrall, 2011), но активный поиск молекулярных механизмов, реализующих наследственную предрасположенность к ИБС, длительное время оставался малоуспешным. Многочисленные молекулярно-генетические исследования были направлены на изучение полиморфизмов в функционально значимых генах, принимающих участие в процессах, связанных с патогенезом ИБС: обмене липидов, тромбообразовании, эндотелиальной дисфункции, регуляции ренин-ангиотензиновой системы, воспалительных реакций. Однако лишь для небольшого числа генов, главным образом связанных с метаболизмом липопротеинов низкой плотности (ЛПНП), удалось установить связь с риском развития ИБС (Brown, Goldstein, 1986; Abifadel et al., 2003; Cohen et al., 2006).

Новая эра в изучении ИБС началась, когда благодаря развитию технологий генотипирования стал доступен полногеномный анализ ассоциаций (ПГАА), ставший основным инструментом для выяснения генетической природы сложных признаков. Результаты первых полногеномных исследований ИБС были опубликованы в 2007 г. Три группы исследователей независимо друг от друга обнаружили локус 9p21.3, ассоциированный с ИБС (Helgadottir et al., 2007; McPherson et al., 2007; Samani et al., 2007). Связь между 9p21.3 и ИБС является одной из самых сильных генетических ассоциаций, наблюдаемых в ПГАА (Schunkert et al., 2011), при этом данный локус расположен в так называемой генной пустыне – регионе, в котором нет генов, кодирующих известные белки. Точный механизм, с помощью которого 9p21.3 влияет на ИБС, до сих пор остается неясным.

Спустя два года в нескольких относительно небольших индивидуальных исследованиях удалось выявить двенадцать дополнительных локусов (Erdmann et al., 2009; Gudbjartsson et al., 2009; Samani et al., 2009; Tregouet et al., 2009). Существенный прогресс был достигнут, когда стали проводить метаанализ, объединяя результаты анализов отдельных исследований. Это стало возможным благодаря созданию международных консорциумов по изучению ИБС и ИМ, например Myocardial Infarction Genetics Consortium (MIGen, 6042 чел.), Coronary Artery Disease (C4D) Genetics Consortium (68065 чел.), IBC 50K CAD Consortium (180183 чел.), Coronary ARtery DIsease Genome wide Replication and Meta-analysis consortium (CARDIoGRAM, 86995 чел.).

К настоящему времени в общей сложности известно около 300 однонуклеотидных полиморфных вариантов (single nucleotide polymorphism, SNP), ассоциированных с ИБС, локализованных в более чем 100 локусах (Schunkert et al., 2011; The CARDIoGRAMplusC4D Consortium, 2013; Nikpay et al., 2015; Nioi et al., 2016; Stitziel et al., 2016; Howson et al., 2017; Klarin et al., 2017; Nelson et al., 2017; Webb et al., 2017; van der Harst, Verweij, 2018), которые совместно объясняют около 40 % генетической компоненты ИБС (Nikpay et al., 2015; Hartiala et al., 2017; Nelson et al., 2017; Vilne, Schunkert, 2018).

Трудно объяснить генетическую структуру заболевания только на основе результатов ПГАА, так как подавляющее большинство ассоциированных с болезнью SNP лежит в некодирующих регионах генома. Они не связаны с изменениями структуры и функции белка, а вероятнее всего, оказывают влияние на уровень экспрессии генов. Кроме того, идентифицированные с помощью ПГАА SNP служат своего рода флагом для всех SNP, находящихся в неравновесии по сцеплению с ними и которые также могут быть задействованы в контроле данного заболевания. Вследствие этих причин конкретные гены, непосредственно влияющие на риск развития ИБС, все еще остаются неизвестными.

Несколько лет назад был предложен подход, в основе которого лежит региональный анализ ассоциаций (РАА), или анализ ассоциаций на генном уровне (gene-based association analysis), при котором все полиморфные варианты

внутри гена анализируются одновременно (Li, Leal, 2008; Eichler et al., 2010). Это позволяет установить статистически значимую ассоциацию между геном и признаком даже в том случае, когда ни один из вариантов внутри гена не дает значимого сигнала. При использовании РАА в анализ можно вовлечь не только распространенные, но и редкие генетические варианты, тем самым увеличивая статистическую мощность анализа. Кроме того, РАА сразу позволяет интерпретировать значимый сигнал ассоциации как указание на роль структуры соответствующего белка в контроле изучаемого признака или болезни.

Применение РАА для анализа ИБС до недавнего времени было затруднено тем, что он требовал огромных выборок, содержащих информацию об индивидуальных фенотипах и генотипах всех членов выборки. Такие данные никогда не депонируются в открытые базы данных, в отличие от результатов ПГАА и метаанализа, представленных набором суммарных статистик для каждого исследованного генетического варианта. В последних работах было показано, что РАА можно проводить *in silico*, используя суммарные статистики, не требуя индивидуальных фенотипов и генотипов (de Leeuw et al., 2015; Bakshi et al., 2016; Wang et al., 2017; Svishcheva et al., 2019).

Целью нашей работы является поиск новых генов, участвующих в контроле ИБС, путем *in silico* картирования с использованием регионального анализа ассоциаций.

Материалы и методы

Суммарные статистики. Использовались два набора суммарных статистик, полученных для ИБС с помощью ПГАА и депонированных в общедоступных базах данных. Первый набор, MICAD (Myocardial Infarction Genetics and CARDIoGRAM Exome), представляет собой результаты метаанализа 20 отдельных когорт. Детальная информация о каждой из них представлена в оригинальной статье (Stitziel et al., 2016). Метаанализ включает 42 335 человек с ИБС и/или ИМ и 78 240 здоровых людей европейского происхождения. Генотипирование осуществлялось с использованием панелей Illumina HumanExome BeadChip array и Illumina OmniExome, дополнительная информация о которых размещена по адресу http://genome.sph.umich.edu/wiki/Exome_Chip_Design. Суммарные статистики (*p*-value, размер эффекта (β)) для 85 112 SNP были загружены с сайта www.cardiogramplusc4d.org.

Второй набор данных, UKbb, представляет результаты ПГАА хронической ишемической болезни сердца из проекта UK Biobank Resource, включающего 8755 больных и 328 444 здоровых человека. Генотипирование осуществлялось на основе метабочипа, специально созданного для ИБС. Подробное описание критериев включения и характеристики фенотипа дано в оригинальных публикациях (Klarin et al., 2017; van der Harst, Verweij, 2018). Суммарные статистики для 10 894 597 SNP были загружены с сайта www.ukbiobank.ac.uk.

Таким образом, анализируемые выборки различались между собой как по структуре (метаанализ/индивидуальное исследование, число больных и здоровых), так и по набору генотипированных маркеров.

Региональный анализ ассоциаций проводили с использованием программного пакета sumFREGAT, созданного коллективом авторов на языке программирования R. Свободный доступ к нему открыт по адресу: https://CRAN.R-project.org/package=sumFREGAT. Пакет предназначен для тестирования генных ассоциаций с использованием суммарных статистик. В нем реализован широкий спектр методов для регионального анализа ассоциаций, построенных на обобщенной регрессионной модели с аддитивными генетическими эффектами (Svishcheva, 2019). Эти методы включают: метод множественной линейной регрессии (MLR); методы, использующие функциональный анализ данных (FLM) и анализ главных компонент (PCA); метод коллапсинга или Бурден-тест (BT); метод SKAT, основанный на анализе компонент дисперсии и использующий принцип ядерного сглаживания, а также метод SKAT-O, объединяющий BT и SKAT. Перечисленные методы отличаются стратегией комбинирования суммарных статистик и являются эффективным инструментом для РАА. Кроме них в пакете реализованы еще три метода, использующие два традиционных подхода для объединения суммарных статистик. Один метод основан на сумме хи-квадратов (sum χ^2) и в двух методах применен минимум *p*-value (min *p*-value и SimpleM).

Подробно методы и их имплементация описаны в нашей работе (Svishcheva et al., 2019), где также проведена апробация каждого метода на выборках различной структуры с использованием реальных данных, в том числе MICAD и UKbb.

Для проведения анализа на генном уровне с помощью пакета sumFREGAT достаточно задать в качестве входных данных суммарные статистики ПГАА, вычисленные для генетических вариантов исследуемого гена, а именно *p*-value и размер эффекта (β), и матрицы корреляций между генотипами этих вариантов. Для подсчета матриц корреляций мы использовали референсную панель 1000 Genomes Phase 3, включающую 503 индивида европейского происхождения. Границы генов устанавливали в соответствии с GRCh37/hg19. Маркер приписывали к определенному гену, если он находится в его экзоне или интроне.

Уровень значимости определяли с помощью поправки Бонферрони. Сигнал ассоциации считали значимым, если значение *p*-value для гена не превышало порогового уровня $4.58 \cdot 10^{-6}$ (0.05/10 912 генов) для выборки MICAD и $2.78 \cdot 10^{-6}$ (0.05/17 975 генов) для выборки UKbb. Мы считали, что ген достоверно ассоциирован с болезнью, если хотя бы в одном из методов значение *p*-value достигало порогового уровня.

Биоинформатический анализ. Анализ представленности функциональных групп генов выполняли с помощью FUMA GWAS (Functional Mapping and Annotation of Genome-Wide Association Studies) (http://fuma.ctglab.nl/), используя стандартные настройки, предлагаемые сервисом.

Известные локусы ИБС. На начало данного исследования было известно 97 локусов, ассоциированных с ИБС. Их список приведен в работе (Klarin et al., 2017). Каждый из локусов в этом списке представлен одним или несколькими лидирующими SNP, удовлетворяющими следующим условиям: сигнал ПГАА достигает полногеномного уровня значимости $5.00 \cdot 10^{-8}$ в оригинальном исследовании и реплицирован на независимой выборке

или в метаанализе. Для каждого локуса в списке приводится также набор генов, которые содержат эти SNP и/или являются ближайшими к ним. В дальнейшем такие локусы и гены мы называем известными.

Результаты

Анализируемые выборки UKbb и MICAD были генотипированы на основе двух различных платформ, поэтому число генов и число SNP в них сильно варьировало (табл. 1). Гены, содержащие менее двух SNP, были исключены из анализа.

Для выборки UKbb мы получили статистически значимый сигнал ассоциации (p-value $< 2.78 \cdot 10^{-6}$) хотя бы в одном из тестов для 49 генов. Результаты РАА для этих генов отражены в табл. 2, где для каждого гена приведено наименьшее среди всех тестов значение p-value. Аналогичная таблица приведена для выборки MICAD (табл. 3). Она содержит список из 61 гена, сигнал ассоциации которых достиг порогового уровня $4.58 \cdot 10^{-6}$ хотя бы в одном из тестов.

Всего при анализе выборок UKbb и MICAD было идентифицировано 88 генов, ассоциированных с ИБС, 22 из них были определены на обеих выборках (см. рисунок).

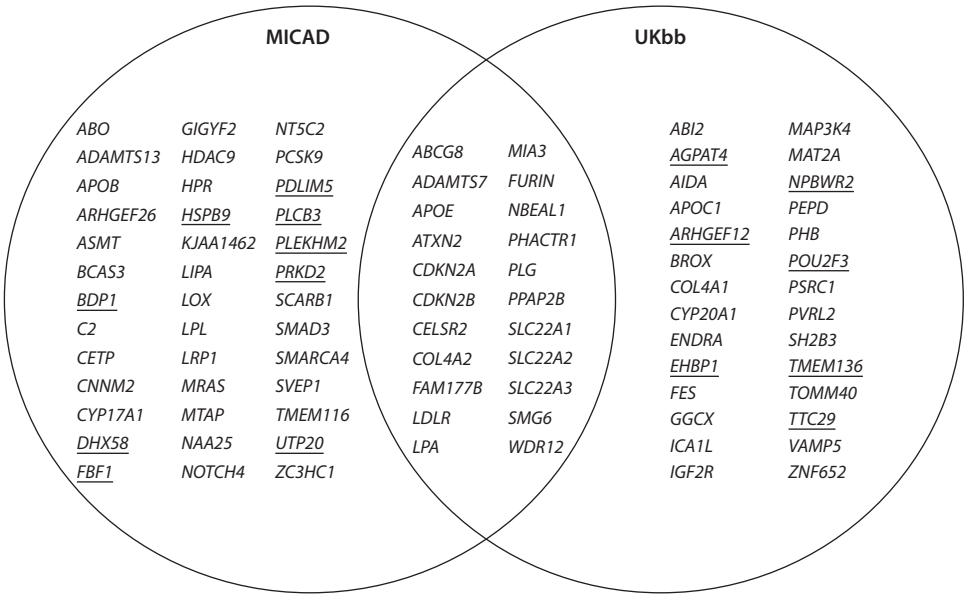
Мы разделили все идентифицированные гены на две категории в зависимости от наличия в них значимого SNP-сигнала. Гены со значимым SNP-сигналом могут быть идентифицированы уже на уровне ПГАА, в то время как гены, не включающие такие сигналы, можно идентифицировать только с использованием РАА. Наибольший интерес представляет вторая категория, к которой относится 41 ген (45.5 % от всех идентифицированных) (табл. 4). Три из них (*COL4A2*, *FURIN* и *SMG6*) были обнаружены с помощью РАА на обеих выборках.

Кроме того, мы сопоставили локализацию идентифицированных генов с известными локусами. Мы определили известный локус как регион ±500 тыс. п. н. от позиции полногеномно значимого SNP, маркирующего этот локус. Регионы, находящиеся от него на расстоянии более 500 тыс. п. н., считали новыми. Оказалось, что 72 из 88 генов (81.8 %) лежат в уже известных локусах, ассоциированных с ИБС. 44 из этих генов входят в список известных генов-кандидатов ИБС, а остальные 28 генов могут рассматриваться как новые гены-кандидаты ИБС, лежащие в известных локусах. 16 генов (*AGPAT4*, *ARHGEF12*, *BDP1*, *DHX58*, *EHBP1*, *FBF1*, *HSPB9*, *NPBWR2*, *PDLIM5*, *PLCB3*, *PLEKHM2*, *POU2F3*, *PRKD2*, *TMEM136*, *TTC29* и *UTP20*) были идентифицированы нами за пределами известных локусов, что позволяет считать их новыми потенциальными генами для ИБС (см. табл. 4). Эти гены расположены в 13 новых локусах на хромосомах 1, 2, 4, 5, 6, 11, 12, 17, 19 и 20.

Анализ представленности функциональных групп генов был проведен для всего набора идентифицированных генов. Значительная часть этих групп включала метаболические пути, связанные с обменом липидов, например белок-липидный комплекс, связывание липопротеинов,

Таблица 1. Характеристика выборок UKbb и MICAD

Выборка	Платформа	Число SNP в гене				Число генов	Общее число SNP
		min	mean	median	max		
MICAD	Exome-chip	2	5.3	4	231	10912	57815
UKbb	Metabochip для ИБС	2	237.4	85	21287	17975	4267351



Гены, ассоциированные с ИБС на двух выборках. Подчеркиванием выделены 16 новых генов.

Таблица 2. Результаты PAA выборки UKbb

Хромосома	Ген	Позиция		Число SNP	*p*-value	Метод	Лидирующий SNP	
		Начало	Конец				*p*-value	SNP
1	*PPAP2B*	56960419	57045257	325	2.96E–07	PCA	6.94E–08	rs17114036
	CELSR2	109792641	109818378	99	1.58E–14	min *p*-val	**4.10E–16**	rs660240
	PSRC1	109822176	109825790	14	2.07E–06	MLR. FLM	2.66E–04	rs35358959
	MIA3	222791444	222841351	170	5.85E–09	sum χ^2	**2.28E–09**	rs2133189
	AIDA	222841355	222886602	107	6.14E–08	SKAT-O	**3.89E–09**	rs36079339
	BROX	222885906	222908529	49	1.19E–07	SKAT-O	**3.97E–08**	rs12043288
	FAM177B	222910558	222926932	44	2.01E–07	sum χ^2	**2.90E–08**	rs34885880
2	*ABCG8*	44066103	44105947	248	3.91E–07	PCA	1.97E–07	rs116426890
	EHBP1	62900986	63273622	739	9.67E–07	SKAT	4.23E–05	rs13410889
	MAT2A	85766101	85772403	14	9.96E–08	PCA	**1.39E–08**	rs2028900
	GGCX	85771978	85788657	60	2.01E–07	SKAT-O	8.43E–08	rs7605975
	VAMP5	85811531	85820511	29	1.67E–06	BT	1.64E–07	rs55971080
	ICA1L	203637873	203736708	292	2.07E–08	sum χ^2	**4.60E–09**	rs72932711
	WDR12	203745323	203776949	110	1.49E–08	sum χ^2	**7.47E–09**	rs6722332
	NBEAL1	203879597	204091101	485	3.12E–07	min *p*-val	**2.53E–09**	rs72934573
	CYP20A1	204103164	204170563	188	8.33E–07	PCA	5.65E–08	rs115600411
4	***TTC29***	147628179	147867034	758	3.43E–07	SimpleM	**1.89E–09**	rs534388792
	EDNRA	148402069	148466106	171	1.68E–09	FLM	**8.29E–09**	rs77249653
6	*PHACTR1*	12716888	13290476	1980	1.99E–11	min *p*-val	**1.52E–13**	rs9349379
	IGF2R	160390131	160527583	473	3.94E–12	PCA	**7.58E–14**	rs3777402
	SLC22A1	160542821	160580147	139	5.54E–12	min *p*-val	**1.33E–13**	rs2282143
	SLC22A2	160637794	160679963	146	1.97E–07	min *p*-val	**4.29E–09**	rs184091076
	SLC22A3	160769405	160873613	361	4.47E–50	MLR	**2.65E–27**	rs4709431
	LPA	160952506	161097478	410	2.85E–97	MLR	**1.03E–48**	rs55730499
	PLG	161123225	161175086	203	1.69E–52	MLR	**1.10E–32**	rs4252185
	MAP3K4	161412816	161538417	461	7.22E–11	min *p*-val	**1.03E–12**	rs141766382
	AGPAT4	161551057	161695107	555	7.08E–07	min *p*-val	**7.81E–09**	rs148349043
9	*CDKN2A*	21967751	21994490	62	1.82E–17	MLR	**4.21E–14**	rs3731239
	CDKN2B	22002902	22009312	15	1.20E–33	PCA	**1.19E–32**	rs3217992
11	***POU2F3***	120107349	120190653	305	1.28E–06	PCA	2.01E–04	rs138488120
	TMEM136	120195838	120204397	27	2.27E–06	MLR. FLM	2.86E–06	rs1893261
	ARHGEF12	120207264	120360645	398	1.08E–06	sum χ^2	3.07E–07	rs12417256
12	*SH2B3*	111843720	111889427	85	8.79E–07	PCA	2.35E–07	rs3184504
	ATXN2	111890018	112037480	258	1.27E–06	PCA	1.78E–07	rs4766578
13	*COL4A1*	110801310	110959496	863	4.36E–11	PCA	3.56E–06	rs9521642
	COL4A2	110959631	111165374	1252	1.79E–06	FLM	5.46E–06	rs9515203
15	*ADAMTS7*	79051545	79103805	202	1.66E–09	PCA	**1.47E–08**	rs2904223
	FURIN	91411885	91426687	37	4.91E–07	sum χ^2	6.27E–08	rs8039305
	FES	91427665	91439006	38	2.58E–08	sum χ^2	1.28E–07	rs7177338
17	*SMG6*	1963133	2207069	846	1.85E–06	BT	1.11E–06	rs67457628
	ZNF652	47366568	47439835	209	3.05E–06	sum χ^2	1.13E–06	rs62076439
	PHB	47481410	47492267	36	2.63E–07	sum χ^2	**2.30E–08**	rs7502499
19	*LDLR*	11200037	11244506	220	2.14E–08	PCA	**2.62E–08**	rs6511720
	PEPD	33877855	34012799	754	1.64E–07	min *p*-val	**1.91E–09**	rs145436496
	PVRL2	45349393	45392485	221	4.61E–09	FLM	9.24E–08	rs6857
	TOMM40	45394477	45406946	61	6.74E–07	PCA	7.47E–08	rs2075650
	APOE	45409039	45412650	9	4.00E–15	PCA	**2.64E–09**	rs429358
	APOC1	45417577	45422606	11	3.99E–09	SKAT	**4.01E–08**	rs12721051
20	***NPBWR2***	62737183	62738184	6	1.92E–07	SKAT-O	1.06E–07	rs13036542

Примечание. Здесь и в табл. 3: позиция гена указана в соответствии с GRCh37/hg19; новые гены выделены жирным шрифтом; *p*-value, не превышающие полногеномного уровня значимости $5.0 \cdot 10^{-8}$, указаны жирным шрифтом.

Таблица 3. Результаты PAA выборки MICAD

Хромосома	Ген	Позиция		Число SNP	*p*-value	Метод	Лидирующий SNP	
		Начало	Конец				*p*-value	SNP
1	***PLEKHM2***	16010632	16061264	4	2.47E–06	SKAT	3.72E–06	rs12091750
	PCSK9	55505149	55530526	5	3.21E–11	MLR	**8.03E–10**	rs11591147
	PPAP2B	56960419	57045257	5	2.79E–10	SKAT-O	**4.55E–10**	rs9970807
	CELSR2	109792641	109818378	18	1.43E–22	SKAT	**1.13E–23**	rs12740374
	MIA3	222791444	222841351	11	9.03E–14	min *p*-val	**1.22E–14**	rs17465637
	FAM177B	222910558	222926932	3	1.75E–10	SKAT	**1.12E–09**	rs2378607
2	*APOB*	21224301	21266945	33	2.60E–06	MLR	1.09E–04	rs533617
	ABCG8	44066103	44105947	10	1.30E–09	SKAT	**4.88E–09**	rs4299376
	WDR12	203745323	203776949	2	1.24E–15	PCA	**1.25E–15**	rs35212307
	NBEAL1	203879597	204091101	14	3.11E–14	min *p*-val	**2.26E–15**	rs2351524
	GIGYF2	233562015	233725287	5	8.17E–09	PCA	**1.46E–08**	rs1801251
3	*MRAS*	138066490	138124377	2	1.83E–08	min *p*-val	**1.83E–08**	rs2306374
	ARHGEF26	153838792	153975616	3	4.92E–09	sum χ^2	**8.28E–09**	rs12493885
4	***PDLIM5***	95373008	95589378	6	2.30E–06	sum χ^2	5.97E–07	rs2452600
5	***BDP1***	70751442	70863649	18	8.26E–07	BT	3.84E–03	rs12187098
	LOX	121398890	121414206	2	1.40E–07	SKAT	1.72E–07	rs1800449
6	*PHACTR1*	12716888	13290476	7	5.17E–35	min *p*-val	**8.61E–36**	rs9349379
	C2	31865562	31913449	18	1.02E–06	SimpleM	7.87E–08	rs3130683
	NOTCH4	32162620	32191844	33	2.41E–06	MLR	8.98E–06	rs394657
	SLC22A1	160542821	160580147	15	2.47E–11	min *p*-val	**1.78E–12**	rs2282143
	SLC22A2	160637794	160679963	3	2.18E–10	min *p*-val	**1.09E–10**	rs2279463
	SLC22A3	160769405	160873613	5	7.41E–26	MLR	**6.98E–21**	rs1810126
	LPA	160952506	161097478	17	1.16E–28	MLR	**9.74E–24**	rs3798220
	PLG	161123225	161175086	11	8.76E–10	PCA	**2.62E–10**	rs4252120
7	*HDAC9*	18126572	19039135	6	1.92E–06	min *p*-val	3.20E–07	rs2023938
	ZC3HC1	129658126	129691233	4	1.30E–12	min *p*-val	**4.11E–13**	rs11556924
8	*LPL*	19796582	19824770	9	6.78E–09	MLR	6.99E–07	rs264
9	*MTAP*	21802635	21941040	4	2.34E–07	min *p*-val	7.76E–08	rs7023954
	CDKN2A	21967751	21994490	2	8.11E–07	PCA	8.46E–07	rs3731249
	CDKN2B	22002902	22009312	2	1.67E–37	PCA	**1.47E–33**	rs1063192
	SVEP1	113127528	113342160	35	2.91E–06	min *p*-val	1.05E–07	rs111245230
	ABO	136130563	136150630	22	2.42E–18	MLR	**9.31E–17**	rs507666
	ADAMTS13	136279459	136324525	14	9.78E–07	PCA	1.63E–07	rs4962153
10	*KIAA1462*	30301729	30348488	8	8.68E–09	SKAT-O	**3.94E–09**	rs3739998
	LIPA	90973326	91011660	6	1.22E–15	min *p*-val	**3.04E–16**	rs2246942
	CYP17A1	104590288	104597290	3	1.41E–06	SKAT-O	1.85E–06	rs1004467
	AS3MT	104629183	104661656	4	2.97E–11	SKAT-O	**2.16E–08**	rs11191447
	CNNM2	104678075	104838344	2	2.86E–09	sum χ^2	**4.46E–08**	rs12413409
	NT5C2	104847774	104953063	2	5.56E–08	min *p*-val	**2.78E–08**	rs11191580
11	***PLCB3***	64018995	64036924	6	3.75E–07	MLR	3.98E–05	rs12146487
12	*LRP1*	57522282	57607142	17	6.10E–08	PCA	**1.78E–08**	rs11172113
	UTP20	101673905	101780397	6	2.42E–07	BT	3.31E–05	rs2305858
	ATXN2	111890018	112037480	7	1.67E–13	PCA	**2.50E–13**	rs653178
	TMEM116	112369086	112451023	2	2.91E–07	PCA	2.91E–07	rs3752630
	NAA25	112464493	112546826	9	4.50E–08	min *p*-val	**8.20E–09**	rs17696736
	SCARB1	125262174	125348519	5	1.48E–09	min *p*-val	**3.69E–10**	rs11057830
13	*COL4A2*	110959631	111165374	13	6.00E–07	SKAT	9.09E–05	rs3809346
15	*SMAD3*	67358036	67487533	3	2.15E–06	min *p*-val	8.77E–07	rs17293632
	ADAMTS7	79051545	79103805	7	8.87E–07	min *p*-val	4.43E–07	rs1994016
	FURIN	91411885	91426687	4	1.64E–06	PCA	1.64E–06	rs17514846
16	*CETP*	56995835	57017757	7	8.80E–07	SKAT	1.12E–06	rs1532624
	HPR	72097125	72111145	2	7.55E–08	PCA	7.57E–08	rs2000999
17	*SMG6*	1963133	2207069	11	7.76E–07	SKAT-O	3.17E–06	rs903160
	DHX58	40253422	40264751	23	2.97E–06	min *p*-val	5.93E–07	rs2074158
	HSPB9	40274756	40275371	2	2.41E–06	SKAT-O	3.01E–06	rs1122326
	BCAS3	58755172	59470199	9	2.06E–07	PCA	7.25E–08	rs8080784
	FBF1	73905925	73937119	6	1.48E–06	SKAT	9.13E–07	rs1135889
19	*SMARCA4*	11071598	11172958	4	1.45E–10	SKAT	**1.02E–10**	rs1122608
	LDLR	11200037	11244506	3	6.84E–19	min *p*-val	**2.28E–19**	rs6511720
	APOE	45409039	45412650	3	5.17E–19	PCA	**9.97E–16**	rs7412
	PRKD2	47177573	47220384	11	2.60E–06	SKAT	2.05E–05	rs425105

Таблица 4. Локализация ассоциированных с ИБС генов, идентифицированных РАА

Локализация	Гены со значимым сигналом ПГАА	Гены без значимого сигнала ПГАА
Известные гены-кандидаты в известных локусах	***ABCG8***, *ABO*, ***ADAMTS7***, *APOC1*, ***APOE***, *ARHGEF26*, ***CDKN2B***, *CNNM2*, *EDNRA*, *GIGYF2*, *KIAA1462*, ***LDLR***, *LIPA*, ***LPA***, *LPL*, *LRP1*, ***MIA3***, *MRAS*, *NT5C2*, *PCSK9*, ***PHACTR1***, ***PLG***, ***PPAP2B***, *SCARB1*, ***SLC22A3***, ***WDR12***, *ZC3HC1*	*APOB*, *BCAS3*, *C2*, *CETP*, *COL4A1*, ***COL4A2***, *CYP17A1*, *FES*, ***FURIN***, *GGCX*, *HDAC9*, *LOX*, *SH2B3*, *SMAD3*, ***SMG6***, *SVEP1*, *VAMP5*
Новые гены-кандидаты в известных локусах	*AIDA*, *AS3MT*, ***ATXN2***, *BROX*, ***CDKN2A***, ***CELSR2***, ***FAM177B***, *ICA1L*, *IGF2R*, *MAP3K4*, *MAT2A*, *NAA25*, ***NBEAL1***, *PEPD*, *PHB*, ***SLC22A1***, ***SLC22A2***, *SMARCA4*	*ADAMTS13*, *CYP20A1*, *HPR*, *MTAP*, *NOTCH4*, *PSRC1*, *PVRL2*, *TMEM116*, *TOMM40*, *ZNF652*
Новые гены в новых локусах	*AGPAT4*, *TTC29*	*ARHGEF12*, *BDP1*, *DHX58*, *EHBP1*, *FBF1*, *HSPB9*, *NPBWR2*, *PDLIM5*, *PLCB3*, *PLEKHM2*, *POU2F3*, *PRKD2*, *TMEM136*, *UTP20*

Примечание. Гены, идентифицированные на обеих выборках MICAD и UKbb, указаны жирным шрифтом.

Таблица 5. Онтологии новых идентифицированных генов

Генная онтология	Число генов		Значение *p*	Гены
	в генной онтологии	из набора		
Метаболизм нейтральных липидов	83	7	9.39E–08	*PCSK9*, *CETP*, *APOE*, *APOC1*, *APOB*, ***AGPAT4***, *LPL*
Катаболические процессы в организме	947	17	9.42E–07	*LIPA*, *NT5C2*, ***PLCB3***, *SCARB1*, *COL4A1*, *COL4A2*, *ADAMTS7*, *FURIN*, *HPR*, *PEPD*, *APOE*, ***PRKD2***, *APOB*, *SLC22A3*, *LPL*, *MTAP*, *CDKN2A*
Метаболизм алкоголя	347	10	3.68E–06	*PCSK9*, *LIPA*, *CYP17A1*, ***PLCB3***, *SCARB1*, *CETP*, *LDLR*, *APOE*, *APOC1*, *APOB*
Формирование сосудов	468	11	8.17E–06	*PPAP2B*, *LRP1*, *COL4A1*, *COL4A2*, *BCAS3*, *APOE*, ***PRKD2***, *APOB*, *EDNRA*, *LOX*, *NOTCH4*
Развитие циркуляторной системы	786	14	1.07E–05	*PPAP2B*, *LRP1*, *COL4A1*, *COL4A2*, *SMAD3*, *BCAS3*, *APOE*, ***PRKD2***, *APOB*, ***PDLIM5***, *EDNRA*, *LOX*, *NOTCH4*, *HDAC9*
Регуляция миграции эндотелиальных клеток	114	6	1.28E–05	*PPAP2B*, *SCARB1*, *BCAS3*, *APOE*, ***PRKD2***, *HDAC9*
Морфогенез кровеносных сосудов	364	9	3.88E–05	*LRP1*, *COL4A1*, *COL4A2*, *BCAS3*, *APOE*, ***PRKD2***, *APOB*, *EDNRA*, *NOTCH4*
Метаболический процесс органических гидроксисоединений	481	10	6.12E–05	*PCSK9*, *LIPA*, *CYP17A1*, ***PLCB3***, *SCARB1*, *CETP*, *LDLR*, *APOE*, *APOC1*, *APOB*

Примечание. Новые гены выделены жирным шрифтом.

транспорт фосфолипидов и др. Этот результат хорошо согласуется с липопротеидной теорией развития ИБС. Онтологии, включающие новые гены, приведены в табл. 5. Как видно, во все функциональные группы, обнаруженные для новых генов, входят также известные гены. Большая часть этих групп относится к формированию кровеносных сосудов.

Обсуждение

В данном исследовании было проведено *in silico* картирование ИБС с использованием методов РАА. При анализе двух выборок европейского происхождения MICAD и UKbb нами идентифицировано 88 генов, ассоциированных с ИБС, 16 из которых являются новыми.

Научная новизна нашего подхода обеспечена как спецификой материала, так и новизной статистических методов анализа. Анализ был выполнен на материале, представленном суммарными статистиками, такими как *p*-value и размер эффекта (β), являющимися результатом ПГАА. Использование суммарных статистик для РАА обладает рядом существенных преимуществ перед традиционным подходом, основанным на материале, представленном индивидуальными фенотипами и генотипами. Во-первых, при использовании суммарных статистик не возникает этических проблем, связанных с доступом к персональной информации, что позволяет размещать их в базах данных с открытым доступом (Pasaniuc, Price, 2017). К настоящему времени существуют уже десятки таких баз данных (Zheng et al., 2017). Во-вторых, такой подход позволяет вовлечь в анализ материал, на порядки превышающий по объему каждое индивидуальное исследование, так как суммарные статистики часто являются результатом метаанализа, объединяющего ресурсы отдельных исследований. Кроме того, анализ суммарных статистик освобождает исследователя от проблем, связанных с подразделенностью выборки, наличием в ней скрытого родства, отклонением распределения анализируемого признака от нормального, наличием выбросов и др., поскольку эти проблемы обычно устранены при проведении ПГАА (Yang et al., 2018).

Новизна методического подхода заключалась в следующем. Во-первых, мы использовали РАА, который позволил нам обнаружить 40 генов, не имеющих значимых сигналов

ПГАА в анализируемых выборках. При этом для трех генов, *COL4A2*, *FURIN* и *SMG6*, связь с ИБС была ранее установлена на других выборках, и они уже вошли в список генов-кандидатов в известных локусах (Klarin et al., 2017). То, что мы их обнаружили с помощью РАА, служит положительным контролем нашего анализа. Во-вторых, мы использовали широкий спектр методов объединения суммарных статистик, реализованных в созданном ранее авторами этой статьи пакете sumFREGAT (Svishcheva et al., 2019). Кроме методов минимума *p*-value и суммы хи-квадратов, этот пакет содержит ряд мощных методов, не реализованных ни в одном из существующих пакетов для РАА. Применение всего набора методов позволяет существенно увеличить мощность анализа и получить максимально полную информацию о генетической природе заболевания.

Анализ ассоциаций на генном уровне позволил приоритизировать гены в известных локусах. Традиционно в качестве генов-кандидатов рассматривают гены, включающие самый сильный сигнал ПГАА и/или ближайшие к нему гены. Однако из-за неравновесия по сцеплению к развитию заболевания могут быть причастны и более удаленные от лидирующего SNP гены. Мы рассмотрели регион 1 млн п.н. вокруг позиции SNP, маркирующего известный локус, так как показано, что свыше 99 % генетических вариантов, находящихся в неравновесии по сцеплению с этим SNP, локализованы в окне размером 1 cM (Bulik-Sullivan et al., 2015). Помимо 44 генов из числа известных, в эти регионы попали 28 новых генов, которые можно считать приоритетными для дальнейшего исследования известных локусов. Некоторые из них уже появлялись в поле зрения ученых в связи с участием продуктов этих генов в различных метаболических процессах. Например, в исследовании (Samani et al., 2008) авторы предположили, что локус вблизи генов *PSRC1* и *CELSR2* повышает риск развития коронарного атеросклероза посредством влияния на уровень липопротеинов низкой плотности. Важную роль в метаболизме лекарственных средств играют гены *SLC22A1* и *SLC22A2*, которые кодируют белки-транспортеры, отвечающие за попадание в клетки органических катионов. Мутации в гене *SLC22A1*, снижающие его экспрессию, приводят к целому ряду различных негативных последствий на энергетический обмен (Zhou et al., 2009). Известно также, что ключевым ферментом для метаболизма лекарств и синтеза холестерина, стероидов и других липидов является цитохром P450, кодируемый семейством генов *CYP*, к которому принадлежит *CYP20A1* (Zanger, Schwab, 2013).

Идентифицированные нами 16 генов попали в 13 новых локусов, расположенных на 10 хромосомах. Гены *AGPAT4*, *PLCB3*, *PRKD2*, *PDLIM5* были кластеризованы в группы, ассоциированные с несколькими биологическими процессами, которые раньше уже были ассоциированы с ИБС и ИМ. Известно, что белок, кодируемый геном *AGPAT4*, играет важную роль в поддержании нормального уровня докозагексаеновой кислоты (основного структурного компонента липидов в мембране клеток головного мозга) (Eto et al., 2014). Гены *PLCB3* и *PRKD2* участвуют в процессах пролиферации и миграции эндотелиальных клеток, необходимых для ангиогенеза (Bhattacharya et al., 2009). Ген *PDLIM5* участвует в процессе экспансии кардиомиоцитов и связан с развитием дилатационной кардиомиопатии (Bang et al., 2014).

Двенадцать новых генов, *BDP1*, *ARHGEF12*, *DHX58*, *EHBP1*, *FBF1*, *HSPB9*, *NPBWR2*, *PLEKHM2*, *POU2F3*, *TMEM136*, *TTC29* и *UTP20*, не вошли в состав статистически значимых функциональных кластеров. Тем не менее известно, что ген *UTP20* контролирует кальцификацию коронарной артерии (O'Donnell et al., 2007), а *EHBP1* связан с уровнями липидов (Willer et al., 2013). Мутации в гене *PLEKHM2* приводят к аномальной локализации лизосом и нарушению механизмов аутофагии и являются причиной рецессивной дилатационной кардиомиопатии и некомпактной кардиомиопатии левого желудочка (Muhammad et al., 2015). Полученные нами данные о причастности этих генов к ИБС могут способствовать лучшему пониманию этиологии и патогенеза этой болезни.

Заключение

Основной задачей современного здравоохранения является переход к персонализированной медицине, в основе которой лежит учет индивидуальных особенностей пациентов, в частности наличие генетических маркеров болезни. Очевидно, чем полнее список таких маркеров, тем эффективнее профилактика и лечение болезни. Полученные нами результаты открывают новые возможности в этом направлении. Они демонстрируют, что РАА служит мощным инструментом для поиска новых генов, позволяющим извлекать новую информацию из накопленных в мире огромных объемов данных, полученных с помощью ПГАА, без дополнительных материальных затрат.

Список литературы / References

Abifadel M., Varret M., Rabes J.P., Allard D., Ouguerram K., Devillers M., Cruaud C., Benjannet S., Wickham L., Erlich D., Derre A., Villeger L., Farnier M., Beucler I., Bruckert E., Chambaz J., Chanu B., Lecerf J.M., Luc G., Moulin P., Weissenbach J., Prat A., Krempf M., Junien C., Seidah N.G., Boileau C. Mutations in *PCSK9* cause autosomal dominant hypercholesterolemia. Nat. Genet. 2003; 34:154-156. DOI 10.1038/ng1161.

Bakshi A., Zhu Z., Vinkhuyzen A.A., Hill W.D., McRae A.F., Visscher P.M., Yang J. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Sci. Rep. 2016;6:32894. DOI 10.1038/srep32894.

Bang C., Batkai S., Dangwal S., Gupta S.K., Foinquinos A., Holzmann A., Just A., Remke J., Zimmer K., Zeug A., Ponimaskin E., Schmiedl A., Yin X., Mayr M., Halder R., Fischer A., Engelhardt S., Wei Y., Schober A., Fiedler J., Thum T. Cardiac fibroblast-derived microRNA passenger strand-enriched exosomes mediate cardiomyocyte hypertrophy. J. Clin. Invest. 2014;124:2136-2146. DOI 10.1172/JCI70577.

Bhattacharya R., Kwon J., Li X., Wang E., Patra S., Bida J.P., Bajzer Z., Claesson-Welsh L., Mukhopadhyay D. Distinct role of PLCβ3 in VEGF-mediated directional migration and vascular sprouting. J. Cell Sci. 2009;122:1025-1034. DOI 10.1242/jcs.041913.

Brown M.S., Goldstein J.L. A receptor-mediated pathway for cholesterol homeostasis. Science. 1986;232:34-47.

Bulik-Sullivan B.K., Loh P.R., Finucane H.K., Ripke S., Yang J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson N., Daly M.J., Price A.L., Neale B.M. LD Score regression distinguishes confounding from polygenicity in ge-

nome-wide association studies. Nat. Genet. 2015;47:291-295. DOI 10.1038/ng.3211.

Cohen J.C., Boerwinkle E., Mosley T.H., Jr., Hobbs H.H. Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. N. Engl. J. Med. 2006;354:1264-1272. DOI 10.1056/NEJMoa054013.

de Leeuw C.A., Mooij J.M., Heskes T., Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. 2015;11:e1004219. DOI 10.1371/journal.pcbi.1004219.

Ehret G.B., Munroe P.B., Rice K.M., Bochud M., Johnson A.D., Chasman D.I., Smith A.V., … van Duijn C.M., Newton-Cheh C., Levy D., Caulfield M.J., Johnson T. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478:103-109. DOI 10.1038/nature10405.

Eichler E.E., Flint J., Gibson G., Kong A., Leal S.M., Moore J.H., Nadeau J.H. Missing heritability and strategies for finding the underlying causes of complex disease. Nat. Rev. Genet. 2010;11:446-450. DOI 10.1038/nrg2809.

Erdmann J., Grosshennig A., Braund P.S., Konig I.R., Hengstenberg C., Hall A.S., Linsel-Nitschke P., … Deloukas P., Thompson J.R., Ziegler A., Samani N.J., Schunkert H. New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat. Genet. 2009; 41:280-282. DOI 10.1038/ng.307.

Eto M., Shindou H., Shimizu T. A novel lysophosphatidic acid acyltransferase enzyme (LPAAT4) with a possible role for incorporating docosahexaenoic acid into brain glycerophospholipids. Biochem. Biophys. Res. Commun. 2014;443:718-724. DOI 10.1016/j.bbrc.2013.12.043.

Gudbjartsson D.F., Holm H., Gretarsdottir S., Thorleifsson G., Walters G.B., Thorgeirsson G., Gulcher J., Mathiesen E.B., Njolstad I., Nyrnes A., Wilsgaard T., Hald E.M., Hveem K., Stoltenberg C., Kucera G., Stubblefield T., Carter S., Roden D., Ng M.C., Baum L., So W.Y., Wong K.S., Chan J.C., Gieger C., Wichmann H.E., Gschwendtner A., Dichgans M., Kuhlenbaumer G., Berger K., Ringelstein E.B., Bevan S., Markus H.S., Kostulas K., Hillert J., Sveinbjornsdottir S., Valdimarsson E.M., Lochen M.L., Ma R.C., Darbar D., Kong A., Arnar D.O., Thorsteinsdottir U., Stefansson K. A sequence variant in *ZFHX3* on 16q22 associates with atrial fibrillation and ischemic stroke. Nat. Genet. 2009;41:876-878. DOI 10.1038/ng.417.

Hartiala J., Schwartzman W.S., Gabbay J., Ghazalpour A., Bennett B.J., Allayee H. The genetic architecture of coronary artery disease: current knowledge and future opportunities. Curr. Atheroscler. Rep. 2017;19:6. DOI 10.1007/s11883-017-0641-6.

Helgadottir A., Thorleifsson G., Manolescu A., Gretarsdottir S., Blondal T., Jonasdottir A., Jonasdottir A., Sigurdsson A., Baker A., Palsson A., Masson G., Gudbjartsson D.F., Magnusson K.P., Andersen K., Levey A.I., Backman V.M., Matthiasdottir S., Jonsdottir T., Palsson S., Einarsdottir H., Gunnarsdottir S., Gylfason A., Vaccarino V., Hooper W.C., Reilly M.P., Granger C.B., Austin H., Rader D.J., Shah S.H., Quyyumi A.A., Gulcher J.R., Thorgeirsson G., Thorsteinsdottir U., Kong A., Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007;316:1491-1493. DOI 10.1126/science.1142842.

Howson J.M.M., Zhao W., Barnes D.R., Ho W.K., Young R., Paul D.S., Waite L.L., … Nordestgaard B.G., Assimes T.L., Danesh J., Butterworth A.S., Saleheen D. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. Nat. Genet. 2017;49:1113-1119. DOI 10.1038/ng.3874.

Johnson A.D., Newton-Cheh C., Chasman D.I., Ehret G.B., Johnson T., Rose L., Rice K., Verwoert G.C., Launer L.J., Gudnason V., Larson M.G., Chakravarti A., Psaty B.M., Caulfield M., van Duijn C.M., Ridker P.M., Munroe P.B., Levy D., on behalf of the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium, Global BPgen Consortium, Women's Genome Health Study. Association of hypertension drug target genes with blood pressure and hypertension in 86,588 individuals. Hypertension. 2011;57:903-910. DOI 10.1161/HYPERTENSIONAHA.110.158667.

Klarin D., Zhu Q.M., Emdin C.A., Chaffin M., Horner S., McMillan B.J., Leed A., Weale M.E., Spencer C.C.A., Aguet F., Segre A.V., Ardlie K.G., Khera A.V., Kaushik V.K., Natarajan P., Consortium C.A.D., Kathiresan S. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat. Genet. 2017;49:1392-1397. DOI 10.1038/ng.3914.

Li B., Leal S.M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am. J. Hum. Genet. 2008;83:311-321. DOI 10.1016/j.ajhg.2008.06.024.

McPherson R., Pertsemlidis A., Kavaslar N., Stewart A., Roberts R., Cox D.R., Hinds D.A., Pennacchio L.A., Tybjaerg-Hansen A., Folsom A.R., Boerwinkle E., Hobbs H.H., Cohen J.C. A common allele on chromosome 9 associated with coronary heart disease. Science. 2007;316:1488-1491. DOI 10.1126/science.1142447.

Muhammad E., Levitas A., Singh S.R., Braiman A., Ofir R., Etzion S., Sheffield V.C., Etzion Y., Carrier L., Parvari R. *PLEKHM2* mutation leads to abnormal localization of lysosomes, impaired autophagy flux and associates with recessive dilated cardiomyopathy and left ventricular noncompaction. Hum. Mol. Genet. 2015;24:7227-7240. DOI 10.1093/hmg/ddv423.

Nelson C.P., Goel A., Butterworth A.S., Kanoni S., Webb T.R., Marouli E., Zeng L., … Farrall M., Danesh J., Samani N.J., Watkins H., Deloukas P. Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat. Genet. 2017;49: 1385-1391. DOI 10.1038/ng.3913.

Nikpay M., Goel A., Won H.H., Hall L.M., Willenborg C., Kanoni S., Saleheen D., … McPherson R., Deloukas P., Schunkert H., Samani N.J., Farrall M. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat. Genet. 2015;47:1121-1130. DOI 10.1038/ng.3396.

Nioi P., Sigurdsson A., Thorleifsson G., Helgason H., Agustsdottir A.B., Norddahl G.L., Helgadottir A., … Holm H., Gudbjartsson D., Sulem P., Thorsteinsdottir U., Stefansson K. Variant *ASGR1* associated with a reduced risk of coronary artery disease. N. Engl. J. Med. 2016;374:2131-2141. DOI 10.1056/NEJMoa1508419.

O'Donnell C.J., Cupples L.A., D'Agostino R.B., Fox C.S., Hoffmann U., Hwang S.J., Ingellson E., Liu C., Murabito J.M., Polak J.F., Wolf P.A., Demissie S. Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study. BMC Med. Genet. 2007;8(Suppl.1):S4. DOI 10.1186/1471-2350-8-S1-S4.

Pasaniuc B., Price A.L. Dissecting the genetics of complex traits using summary association statistics. Nat. Rev. Genet. 2017;18:117-127. DOI 10.1038/nrg.2016.142.

Peden J.F., Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum. Mol. Genet. 2011;20:R198-205. DOI 10.1093/hmg/ddr384.

Samani N.J., Braund P.S., Erdmann J., Gotz A., Tomaszewski M., Linsel-Nitschke P., Hajat C., Mangino M., Hengstenberg C., Stark K., Ziegler A., Caulfield M., Burton P.R., Schunkert H., Tobin M.D. The novel genetic variant predisposing to coronary artery disease in the region of the *PSRC1* and *CELSR2* genes on chromosome 1 associates with serum cholesterol. J. Mol. Med. (Berl). 2008;86:1233-1241. DOI 10.1007/s00109-008-0387-2.

Samani N.J., Deloukas P., Erdmann J., Hengstenberg C., Kuulasmaa K., McGinnis R., Schunkert H., Soranzo N., Thompson J., Tiret L., Ziegler A. Large scale association analysis of novel genetic loci for coronary artery disease. Arterioscler. Thromb. Vasc. Biol. 2009;29:774-780. DOI 10.1161/ATVBAHA.108.181388.

Samani N.J., Erdmann J., Hall A.S., Hengstenberg C., Mangino M., Mayer B., Dixon R.J., Meitinger T., Braund P., Wichmann H.E., Barrett J.H., Konig I.R., Stevens S.E., Szymczak S., Tregouet D.A., Iles M.M., Pahlke F., Pollard H., Lieb W., Cambien F., Fischer M., Ouwehand W., Blankenberg S., Balmforth A.J., Baessler A., Ball S.G., Strom T.M., Braenne I., Gieger C., Deloukas P., Tobin M.D., Ziegler A., Thompson J.R., Schunkert H., WTCCC and the Cardiogenics Consortium. Genomewide association analysis of

coronary artery disease. N. Engl. J. Med. 2007;357:443-453. DOI 10.1056/NEJMoa072366.

Schunkert H., Konig I.R., Kathiresan S., Reilly M.P., Assimes T.L., Holm H., Preuss M., … O'Donnell C.J., McPherson R., Erdmann J., Consortium C.A., Samani N.J. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat. Genet. 2011;43:333-338. DOI 10.1038/ng.784.

Stitziel N.O., the Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Variants in *ANGPTL4* and the risk of coronary artery disease. N. Engl. J. Med. 2016;375:2306. DOI 10.1056/NEJMc1607380.

Svishcheva G.R. A generalized model for combining dependent SNP-level summary statistics and its extensions to statistics of other levels. Sci. Rep. 2019;9:5461. DOI 10.1038/s41598-019-41827-5.

Svishcheva G.R., Belonogova N.M., Zorkoltseva I.V., Kirichenko A.V., Axenovich T.I. Gene-based association tests using GWAS summary statistics. Bioinformatics. 2019;35(19):3701-3708. DOI 10.1093/bioinformatics/btz172.

The CARDIoGRAMplusC4D Consortium (Deloukas P., Kanoni S., Willenborg C., Farrall M., Assimes T.L., Thompson J.R., Ingelsson E., … Palmer C.N., Roberts R., Watkins H., Schunkert H., Samani N.J.) Large-scale association analysis identifies new risk loci for coronary artery disease. Nat. Genet. 2013;45:25-33. DOI 10.1038/ng.2480.

Tregouet D.A., Konig I.R., Erdmann J., Munteanu A., Braund P.S., Hall A.S., Grosshennig A., Linsel-Nitschke P., Perret C., DeSuremain M., Meitinger T., Wright B.J., Preuss M., Balmforth A.J., Ball S.G., Meisinger C., Germain C., Evans A., Arveiler D., Luc G., Ruidavets J.B., Morrison C., van der Harst P., Schreiber S., Neureuther K., Schafer A., Bugert P., El Mokhtari N.E., Schrezenmeir J., Stark K., Rubin D., Wichmann H.E., Hengstenberg C., Ouwehand W., Ziegler A., Tiret L., Thompson J.R., Cambien F., Schunkert H., Samani N.J. Genome-wide haplotype association study identifies the *SLC22A3-LPAL2-LPA* gene cluster as a risk locus for coronary artery disease. Nat. Genet. 2009;41:283-285. DOI 10.1038/ng.314.

van der Harst P., Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ. Res. 2018;122:433-443. DOI 10.1161/CIRCRESAHA.117.312086.

Vilne B., Schunkert H. integrating genes affecting coronary artery disease in functional networks by multi-omics approach. Front. Cardiovasc. Med. 2018;5:89. DOI 10.3389/fcvm.2018.00089.

Wang M., Huang J., Liu Y., Ma L., Potash J.B., Han S. COMBAT: A Combined Association Test for genes using summary statistics. Genetics. 2017;207:883-891. DOI 10.1534/genetics.117.300257.

Webb T.R., Erdmann J., Stirrups K.E., Stitziel N.O., Masca N.G., Jansen H., Kanoni S., … Samani N.J., Schunkert H., Deloukas P., Kathiresan S., for the Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Systematic evaluation of pleiotropy identifies 6 further loci associated with coronary artery disease. J. Am. Coll. Cardiol. 2017;69:823-836. DOI 10.1016/j.jacc.2016.11.056.

Willer C.J., Schmidt E.M., Sengupta S., Peloso G.M., Gustafsson S., Kanoni S., Ganna A., … Deloukas P., Kathiresan S., Mohlke K.L., Ingelsson E., Abecasis G.R. Discovery and refinement of loci associated with lipid levels. Nat. Genet. 2013;45:1274-1283. DOI 10.1038/ng.2797.

Yang J., Chen S., Abecasis G., IAMDGC. Improved score statistics for meta-analysis in single-variant and gene-level association studies. Genet. Epidemiol. 2018;42:333-343. DOI 10.1002/gepi.22123.

Yusuf S., Hawken S., Ounpuu S., Dans T., Avezum A., Lanas F., McQueen M., Budaj A., Pais P., Varigos J., Lisheng L., on behalf of the INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet. 2004;364:937-952. DOI 10.1016/S0140-6736(04)17018-9.

Zanger U.M., Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol. Ther. 2013;138:103-141. DOI 10.1016/j.pharmthera.2012.12.007.

Zheng J., Erzurumluoglu A.M., Elsworth B.L., Kemp J.P., Howe L., Haycock P.C., Hemani G., Tansey K., Laurin C., Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, Pourcain B.S., Warrington N.M., Finucane H.K., Price A.L., Bulik-Sullivan B.K., Anttila V., Paternoster L., Gaunt T.R., Evans D.M., Neale B.M. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics. 2017;33:272-279. DOI 10.1093/bioinformatics/btw613.

Zhou K., Donnelly L.A., Kimber C.H., Donnan P.T., Doney A.S., Leese G., Hattersley A.T., McCarthy M.I., Morris A.D., Palmer C.N., Pearson E.R. Reduced-function *SLC22A1* polymorphisms encoding organic cation transporter 1 and glycemic response to metformin: a GoDARTS study. Diabetes. 2009;58:1434-1439. DOI 10.2337/db08-0896.

Благодарности. Работа выполнена при поддержке Минобрнауки (0324-2019-0040) и РФФИ (18-04-00076).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 02.04.2019. После доработки 15.07.2019. Принята к публикации 19.07.2019.

БИОИНФОРМАТИКА И СИСТЕМНАЯ БИОЛОГИЯ
Оригинальное исследование / Original article
УДК 616.72-002.772:577.2

Вавиловский журнал генетики и селекции. 2019;23(8):1047-1058
DOI 10.18699/VJ19.586

Кандидатные SNP-маркеры ревматоидного полиартрита, которые могут достоверно изменять сродство TATA-связывающего белка к промоторам генов человека

И.В. Чадаева[1, 2], Д.А. Рассказов[1], Е.Б. Шарыпова[1], И.А. Драчкова[1], Е.А. Ощепкова[1], Л.К. Савинкова[1], П.М. Пономаренко[3], М.П. Пономаренко[1, 2] ✉, Н.А. Колчанов[1, 2], В.А. Козлов[4]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Университет Ла-Верна, Ла-Верн, Калифорния, США
[4] Научно-исследовательский институт фундаментальной и клинической иммунологии, Новосибирск, Россия
✉ e-mail: pon@bionet.nsc.ru

Ревматоидный полиартрит (РА) – аутоиммунное заболевание с наличием аутоантител (например, антитела к антигенам цитруллированных белков), и провоспалительных цитокинов, таких как TNF-α и IL-6, которые принимают участие в индукции хронического синовита, эрозии костной ткани с последующей деформацией. Иммунопатогенез базируется на механизмах полома иммунной толерантности к собственным антигенам, что характеризуется повышением активности Т-клеток-эффекторов, обуславливающих симптоматику РА. В то же время на фоне такой повышенной активности эффекторных лимфоцитов регистрируется снижение активности ряда регуляторных клеток, включая регуляторные Т-клетки (Treg) и клетки-супрессоры миелоидного происхождения. Имеются основания полагать, что именно изменение активности клеток-супрессоров является ведущим звеном в патогенезе развития РА. Поэтому говорят лишь о периодах ослабления (ремиссии) РА. Вследствие более мощной женской иммунной системы, по сравнению с мужской, риск развития РА у женщин втрое выше, он снижается при лактации и растет при беременности, а также после менопаузы пропорционально уровню половых гормонов. Считают, что риск развития РА на 50 % зависит от условий и образа жизни, тогда как оставшиеся 50 % – от генетической предрасположенности. Поэтому РА отвечает главной идее постгеномной предиктивно-превентивной персонализированной медицины: дать шанс тем, кто хотел бы снизить риск заболеваний, приведя образ своей жизни в соответствие с данными по своему расшифрованному геному. Это очень важно, поскольку врачи относят РА к наиболее частым причинам инвалидизации. С использованием ранее созданного нами Web-сервиса SNP_TATA_Z-tester (http://beehive.bionet.nsc.ru/cgi-bin/mgs/tatascan_fox/start.pl) исследовано 227 вариантов однонуклеотидного полиморфизма (SNP) промоторов генов человека. В итоге было предсказано 43 кандидатных SNP-маркера РА, которые способны изменять сродство TATA-связывающего белка (TBP) к промоторам этих генов.
Ключевые слова: TATA-связывающий белок (TBP); сайт TBP-связывания (TATA-бокс); промотор; сродство TBP/промотор; ген; однонуклеотидный полиморфизм (SNP); экспрессия гена; достоверное изменение; SNP-маркер; ревматоидный полиартрит (РА).

Для цитирования: Чадаева И.В., Рассказов Д.А., Шарыпова Е.Б., Драчкова И.А., Ощепкова Е.А., Савинкова Л.К., Пономаренко П.М., Пономаренко М.П., Колчанов Н.А., Козлов В.А. Кандидатные SNP-маркеры ревматоидного полиартрита, которые могут достоверно изменять сродство TATA-связывающего белка к промоторам генов человека. Вавиловский журнал генетики и селекции. 2019;23(8):1047-1058. DOI 10.18699/VJ19.586

Candidate SNP-markers of rheumatoid arthritis that can significantly alter the affinity of the TATA-binding protein for human gene promoters

I.V. Chadaeva[1, 2], D.A. Rasskazov[1], E.B. Sharypova[1], I.A. Drachkova[1], E.A. Oshchepkova[1], L.K. Savinkova[1], P.M. Ponomarenko[3], M.P. Ponomarenko[1, 2] ✉, N.A. Kolchanov[1, 2], V.A. Kozlov[4]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] University of La Verne, La Verne, California, USA
[4] Research Institute of Fundamental and Clinical Immunology, Novosibirsk, Russia
✉ e-mail: pon@bionet.nsc.ru

Rheumatoid polyarthritis (RA) is an autoimmune disease with autoantibodies, including antibodies to citrullant antigens and proinflammatory cytokines, such as TNF-α and IL-6, which are involved in the induction of chronic synovitis, bone erosion, followed by deformity. Immunopathogenesis is based on the mechanisms of the breakdown of im-

© Чадаева И.В., Рассказов Д.А., Шарыпова Е.Б., Драчкова И.А., Ощепкова Е.А., Савинкова Л.К., Пономаренко П.М., Пономаренко М.П., Колчанов Н.А., Козлов В.А., 2019

mune tolerance to its own antigens, which is characterized by an increase in the activity of T-effector cells, causing RA symptomatology. At the same time, against the background of such increased activity of effector lymphocytes, a decrease in the activity of a number of regulatory cells, including regulatory T-cells (Treg) and myeloid suppressor cells, is recorded. There is reason to say that it is the change in the activity of suppressor cells that is the leading element in RA pathogenesis. That is why only periods of weakening (remission) of RA are spoken of. According to the more powerful female immune system compared to the male one, the risk of developing RA in women is thrice as high, this risk decreases during breastfeeding and grows during pregnancy as well as after menopause in proportion to the level of sex hormones. It is believed that 50 % of the risk of developing RA depends on the conditions and lifestyle, while the remaining 50 % is dependent on genetic predisposition. That is why, RA fits the main idea of postgenomic predictive-preventive personalized medicine that is to give a chance to those who would like to reduce his/her risk of diseases by bringing his/her conditions and lifestyle in line with the data on his/her genome sequenced. This is very important, since doctors consider RA as one of the most frequent causes of disability. Using the Web service SNP_TATA_Z-tester (http://beehive.bionet.nsc.ru/cgi-bin/mgs/tatascan_fox/start.pl), 227 variants of single nucleotide polymorphism (SNP) of the human gene promoters were studied. As a result, 43 candidate SNP markers for RA that can alter the affinity of the TATA-binding protein (TBP) for the promoters of these genes were predicted.
Key words: TATA binding protein (TBP); TBP-binding site (TATA box); promoter; TBP-promoter affinity; gene; single nucleotide polymorphism (SNP); gene expression; significant change; SNP-marker; rheumatoid polyarthritis (RA).

For citation: Chadaeva I.V., Rasskazov D.A., Sharypova E.B., Drachkova I.A., Oshchepkova E.A., Savinkova L.K., Ponomarenko P.M., Ponomarenko M.P., Kolchanov N.A., Kozlov V.A. Candidate SNP-markers of rheumatoid arthritis that can significantly alter the affinity of the TATA-binding protein for human gene promoters. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1047-1058. DOI 10.18699/VJ19.586 (in Russian)

Введение

Ревматоидный полиартрит (РА) является аутоиммунным заболеванием с наличием аутоантител, таких как антитела к антигенам цитруллированных белков, и провоспалительных цитокинов, например TNF-α и IL-6, принимающих участие в индукции хронического синовита, эрозии костной ткани с последующей деформацией. В основе иммунопатогенеза лежат механизмы полома иммунной толерантности к собственным антигенам, которые характеризуются повышением активности Т-клеток-эффекторов, обуславливающим симптоматику РА. За повышение активности эффекторов Т-клеток отвечают главным образом две субпопуляции клеток-помощников: Th1 и Th17. Вместе с тем на фоне такой повышенной активности эффекторных лимфоцитов регистрируется снижение активности ряда регуляторных клеток, включая регуляторные Т-клетки (Трег) и клетки-супрессоры миелоидного происхождения. Есть основания считать, что именно изменение активности клеток-супрессоров – ведущее звено в патогенезе развития РА. Имеющиеся данные свидетельствуют о снижении супрессорной активности тех же регуляторных Т-клеток у больных РА. С другой стороны, у пациентов с РА отмечается повышение резистентности у Т-клеток эффекторов к супрессорному действию регуляторных Т-клеток. В целом, рассматривая патогенез этого тяжелейшего заболевания, следует отметить участие в нем многих факторов, относящихся к нарушениям механизмов регуляции функционирования иммунокомпетентных клеток, в первую очередь, цитокинов и их рецепторов, хемокинов и их рецепторов, факторов транскрипции.

Из-за аутоиммунной природы диагноз РА, однажды поставленный пациенту, не может быть снят окончательно (Smolen et al., 2016) даже после успешного лечения, поскольку иммунные клетки памяти в крови способны сохранять информацию об антигенах длительное время (Sarkander et al., 2016). Поэтому обычно практикующие врачи говорят лишь о периодах ослабления (ремиссии) РА (Smolen, Aletaha, 2015). Вследствие более мощной иммунной системы у женщин риск развития РА у них втрое выше по сравнению с мужчинами (Krasselt, Baerwald, 2017) и растет с уровнем половых гормонов после менопаузы и во время беременности (Ho, Weinshilboum, 2017), но снижается в период лактации (Karlson et al., 2004). Ретроспективный клинико-фармакологический метаанализ 29 880 больных РА в сравнении с 73 758 условно здоровыми добровольцами выявил 42 новых локуса и 98 генов в качестве кандидатных мишеней для терапевтического воздействия при РА (Okada et al., 2014). Принято считать, что РА у индивида обусловлен на 50 % генетической предрасположенностью (Nair et al., 2017), тогда как остальные 50 % – историей, образом и условиями его жизни, включая вредные привычки (Malm et al., 2016), прежде всего курение (Erlandsson et al., 2016), микробиом как отражение пищевого рациона (Sato et al., 2017), перенесенные болезни (Scott et al., 2010; Sakkas et al., 2017) и условия окружающей среды (Klareskog et al., 2006). В связи с этим РА отвечает идее постгеномной предиктивно-превентивной персонализированной медицины (Ginsburg, Willard, 2009) – предоставить возможность людям, готовым приложить для этого усилия, снизить риск заболеваний, корректируя образ жизни в соответствии с данными своего расшифрованного генома (Trovato, 2014). Это особенно важно, так как врачи относят РА к классу хронических воспалений, которые чаще всего выступают причиной инвалидизации (Köller, Nöbauer-Huhmann, 2009). Кроме того, фибробласт-подобные синовиоциты приобретают способность к гиперпролиферации на поздних стадиях развития РА, что приводит к синовиальной гиперплазии, т. е. нарушению формы и подвижности сустава (Lim, Bae, 2011). Наконец, при лейкопении ювенильный РА у детей до 16 лет достоверно часто перерождается в лейкемию (Jones et al., 2006).

Основой предиктивно-превентивной персонализированной медицины является самый крупный научный проект современности «1000 геномов», в рамках кото-

рого уже расшифровано более 10 тыс. индивидуальных геномов людей (Telenti et al., 2016) и обнаружено более 100 млн однонуклеотидных отличий (SNP) каждого из них от референсного генома человека как биологического вида, собранных в базу данных dbSNP (Sherry et al., 2001). Создан также банк данных dbWGFP обо всех 8.6 млрд потенциально возможных SNP в референсном геноме человека. Эти данные о каждом из SNP необходимы для их использования в предиктивно-превентивной персонализированной медицине (Wu et al., 2016).

Поскольку размер популяции (7.5 млрд человек) все еще численно меньше 8.6 млрд SNP в их геномах, то в рамках дилеммы Холдейна (Haldane, 1957) и теории нейтральной эволюции (Kimura, 1968) абсолютное большинство этих SNP должно быть нейтральным. Поэтому их предварительное исключение быстрыми и дешевыми компьютерными методами *in silico* до начала проведения дорогостоящих, трудоемких и длительных исследований *in vitro* экспериментальными и *in vivo* клиническими методами становится задачей-вызовом для современной биоинформатики.

Сравнение частот заданных SNP с использованием клинических стандартов и протоколов на когортах больных и здоровых добровольцев (Varzari et al., 2016) – единственно приемлемый путь выявления SNP-маркеров для заболеваний человека с минимальным риском ошибочных врачебных рекомендаций на основе этих маркеров. Такие клинические SNP-маркеры собирают в базе данных OMIM (Amberger et al., 2015). В этой базе данных известные биомедицинские SNP-маркеры были картированы чаще всего в белок-кодирующих районах генов, где им соответствуют дефекты структуры и/или функции белков, не поддающиеся коррекции; крайне редко – в регуляторных районах генов (Zerbino et al., 2015), где они изменяют лишь уровень экспрессии этих генов, что корректируемо медикаментозно и/или изменением образа жизни. Многие регуляторные клинические SNP-маркеры попали в район [–70; –20] промоторов генов, где они проявляются как вариации сродства TATA-связывающего белка (TBP) к этим промоторам при реорганизации хроматина от плотной упаковки в нуклеосомы к прединициаторным комплексам транскрипции этих генов (Ponomarenko et al., 2013). Связывание TBP с промоторами абсолютно необходимо для первичного запуска экспрессии любого белок-кодирующего гена, поскольку двойные нокауты по гену *TBP* у мышей (Martianov et al., 2002) и у данио-рерио (Muller et al., 2001) развиваются лишь до бластулы, когда исчерпан запас материнских TBP. В сравнении с нормой эти клинические SNP-маркеры изменяют уровень экспрессии соответствующего гена пропорционально сродству TBP к промотору этого гена (Mogno et al., 2010).

Ранее мы вывели уравнение трехшагового связывания TBP/ДНК (Ponomarenko et al., 2008), верифицировали это уравнение как в наших экспериментах (Savinkova et al., 2013; Arkova et al., 2014; Drachkova et al., 2014), так и с использованием независимых экспериментальных и клинических данных других авторов (Ponomarenko et al., 2009, 2010; Suslov et al., 2010a, b). На этой основе мы создали Web-сервис SNP_TATA_Comparator (Ponomarenko et al., 2015) для применения указанного уравнения связывания TBP с ДНК в процессе анализа референсного генома человека (Telenti et al., 2016). С помощью SNP_TATA_Comparator ранее уже был сделан прогноз кандидатных SNP-маркеров агрессивности, женского репродуктивного потенциала, предрасположенности к социальному подчинению и доминированию, а также ожирения, нарушений циркадного ритма, аутоиммунных заболеваний и болезни Альцгеймера (Arkova et al., 2015; Chadaeva et al., 2016, 2018, 2019; Ponomarenko M. et al., 2016; Ponomarenko P. et al., 2016, 2017). В настоящей работе мы представляем обновление SNP_TATA_Z-tester этого Web-сервиса (Ponomarenko et al., 2015) и результаты его применения к исследованию РА в качестве продолжения серии наших биоинформатических прогнозов.

Материалы и методы

Исследовали последовательности $S = \{s_{-90}\ldots s_i\ldots s_{-1}\}$ длиной 90 п. о. относительно старта транскрипции, TSS (Transcription Start Site, $TSS \equiv s_{1;0} \equiv s_{2;0}$; где $s_i \in \{a, c, g, t\}$) промоторов генов с известными SNP-маркерами заболеваний человека, связанных с РА вследствие изменения сродства TBP к этим промоторам. Их нормальные варианты брали из референсного генома человека (Telenti et al., 2016), минорные – из базы данных dbSNP (Sherry et al., 2001) либо из оригинальных статей (Kavlie et al., 2003). Акронимы генов, идентификаторы SNP в базах данных dbSNP (Sherry et al., 2001) или OMIM (Amberger et al., 2015) и фрагменты ДНК вокруг SNP, а также сопутствующие РА заболевания и литературные ссылки приведены в таблице.

Все рассматриваемые SNP анализировали с помощью обновления SNP_TATA_Z-tester нашего Web-сервиса (Ponomarenko et al., 2015), где был изменен лишь ввод анцестрального аллеля (нормы). Алгоритм работы представлен в Приложении[1] (Hahn et al., 1989; Bucher, 1990; Karas et al., 1996; Ponomarenko et al., 2008, 2009; Waardenberg et al., 2015). Применение SNP_TATA_Z-tester на примере SNP-маркера rs2276109 гена *MMP12* человека для пониженного риска широкого круга аутоиммунных заболеваний (Ponomarenko M. et al., 2016) вследствие дефицита матриксной металлоэластазы макрофага для преодоления межклеточного матрикса в процессе его аутоиммунной атаки на клетку-мишень (Starodubtseva et al., 2011) показано на рисунке, *б*. Два верхних окна на рисунке содержат входные данные SNP_TATA_Z-tester: последовательности ДНК анцестрального (вверху) и минорного (внизу) аллелей промотора *MMP12* (круги выделяют различие между ними), в строке Decision окна Results – прогноз deficiency : significant, что согласуется с известным влиянием SNP rs2276109 на аутоиммунные заболевания (Ponomarenko M. et al., 2016). Этот результат представлен также в таблице.

Далее для SNP, предсказанного как достоверно изменяющий сродство TBP к промотору, проводили поиск по ключевым словам в базе данных PubMed (Lu, 2011) для клинически найденных физиологических маркеров РА, которые отвечают влиянию этого SNP на экспрессию соответствующего гена. Блок-схема поиска дана в Приложении. С его помощью мы нашли ретроспективный

[1] Приложение см. по адресу:
http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx22.pdf

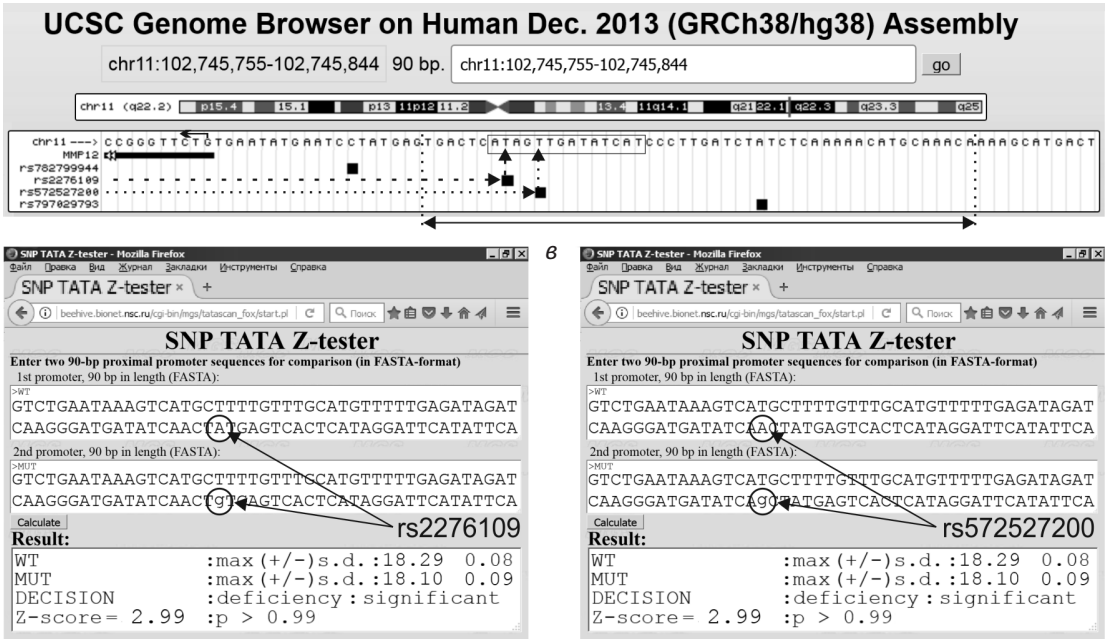
Кандидатные SNP-маркеры ревматоидного полиартрита (РА), предсказанные в этой работе в качестве способных изменять сродство ТВР к промоторам генов человека

Ген	dbSNP (Sherry et al., 2001)	5′-фланк	wt/mut	3′-фланк	K_D, нМ wt/mut	Δ	Z	α	Факт, *прогноз (эта работа)*	Литературный источник
MMP12	rs2276109	gatatcaact	a/g	tgagtcactc	11/14	↓	3	0.01	Пониженные риски псориаза, астмы, системной склеродермы и РА	Hunninghake et al., 2009; Manetti et al., 2010; Starodubtseva et al., 2011;
	rs572527200	gatgatatca	a/g	ctatgagtca	11/14	↓	3	10^{-2}	*← РА*	*Liu et al., 2004*
DHFR	rs10168	ctgcacaaat	g/a	gggacgaggg	15/9	↑	9	10^{-6}	Устойчивость к метотрексату при лейкозе	Al-Shakfa et al., 2009
	rs750793297	tgcacaaatg	g/t	ggacgagggg	15/13	↑	3	0.01	*Устойчивость к метотрексату при РА*	*Bennike et al., 2017*
	rs766799008	ctgcacaaat	a/g	tggggacgag	15/19	↓	3	10^{-3}	*Выше биоактивость метотрексата при РА*	
	rs764508464	ctgcacaaat	a/–	tggggacgag	15/37	↓	17	10^{-6}		
	rs754122321	ctcgcctgca	c/g	aaatggggac	15/25	↓	9	10^{-3}		
NOS2	–51t→c (Burgner et al., 2003)	gtataaatac	t/c	tcttggctgc	2/1	↑	3	0.01	Устойчивость к малярии и эпилепсия *→ РА, ген-мишень при терапии РА*	Burgner et al., 2003; Clark et al., 2003; Gonzalez-Martinez et al., 2009; *Ohtsuka et al., 2002*
F7	–33a→c (Kavlie et al., 2003)	ccttggaggc	a/c	gagaactttg	53/62	↓	3	0.01	Умеренные кровотечения	Kavlie et al., 2003
	rs749691733	agaactttgc	c/t	cgtcagtccc	53/66	↓	4	10^{-3}	*→ РА, F7-терапия против РА*	*Zuber-Jerger et al., 2005; Drobiecki et al., 2013*
F3	rs563763767	ccctttatag	c/t	gcgcggggca	3/2	↑	6	10^{-6}	Инфаркт миокарда, тромбоз *→ РА*	Arnaud et al., 2000; *Lyberg et al., 1982*
MBL2	rs72661131	tctatttcta	t/c	atagcctgca	2/4	↓	12	10^{-6}	Вариабельный иммунодефицит, преэклампсия, инсульт	Boldt et al., 2006; Sziller et al., 2007; Cervera et al., 2010
	rs562962093	atctatttct	a/g	tatagcctgc	2/5	↓	15	10^{-6}	*→ восприимчивость к инфекциям при РА*	*Nisihara et al., 2016*
	rs567653539	tttctatata	g/a	cctgcaccca	2/1	↑	12	10^{-6}	*→ риск сердечно-сосудистых болезней при РА*	*Troelsen et al., 2010*
APOA1	–35a→c (Matsunaga et al., 1999)	tgcagacata	a/c	ataggccctg	3/4	↓	5	10^{-6}	Стеатоз печени *→ РА с сердечно-сосудистым осложнением*	Matsunaga et al., 1999; *Kokkonen et al., 2017*
TPI1	rs1800202	gcgctctata	t/g	aagtgggcag	1/4	↓	17	10^{-6}	Гемолитическая анемия и нейромышечные заболевания	Vives-Corrons et al., 1978; Watanabe et al., 1996
	rs781835924	cgcggcgctc	t/c	atataagtgg	1/2	↓	10	10^{-6}	*Мимикрия РА у детей*	*Richardson et al., 1984*
SOD1	rs7277748	ggtctggcct	a/g	taaagtagtc	2/7	↓	17	10^{-6}	ALS *→ усиливает боли при РА*	Niemann et al., 2007 *Staron et al., 2012*
INS	rs5505	agatcactgt	c/t	cttctgccat	53/44	↑	4	10^{-3}	Неонатальный сахарный диабет	Landrum et al., 2014
	rs563207167	tcagccctgc	c/t	tgtctcccag	53/44	↑	4	10^{-3}	*→ РА*	*Abrahamson, 1952*
IL1B	rs1143627	ttttgaaagc	c/t	ataaaaacag	5/2	↑	15	10^{-6}	Рак печени, легких и желудка, гастрит, ожирение	Ponomarenko et al., 2015
									→ выше циркадность боли при РА	*Olkkonen et al., 2015*

Окончание таблицы

Ген	dbSNP (Sherry et al., 2001)	5′-фланк	wt/mut	3′-фланк	K_D, нМ wt/mut	Δ	Z	α	Факт, *прогноз (эта работа)*	Литературный источник
HBB и *HBD*$	rs397509430	gggctgggca	t/–	atacaacagt	5/29	↓	34	10^{-6}	Талассемия и устойчивость к малярии	Martiney et al., 1996; *Giakoumi et al., 2005*
	rs33980857	gggctgggca	t/a, g, c	atacaacagt	5/21	↓	27	10^{-6}		
	rs34598529	ggctgggcat	a/g	aaagtcaggg	5/18	↓	24	10^{-6}		
	rs33931746	gctgggcata	a/g, c	aagtcagggc	5/11	↓	14	10^{-6}		
	rs33981098	agggctgggc	a/g, c	taaaagtcag	5/9	↓	10	10^{-6}		
	rs34500389	cagggctggg	c/a, t, g	ataaaagtca	5/6	↓	3	0.01		
	rs35518301$	caggaccagc	a/g	taaaaggcag	4/8	↓	11	10^{-6}		
	rs63750953	ctgggcataa	aa/–	gtcagggcag	5/8	↓	9	10^{-6}	⟶ *PA*	
	rs281864525	tgggcataaa	a/c	gtcagggcag	5/8	↓	7	10^{-6}		
	rs34166473$	aggaccagca	t/c	aaaaggcagg	4/8	↓	18	10^{-6}		
ACKR1	rs2814778	ttggctctta	t/c	cttggaagca	10/12	↓	4	10^{-3}	Устойчивость к малярии и лейкопения, *которая известна как предиктор лейкоза из PA у детей*	Michon et al., 2001; Nalls et al., 2008; *Jones et al., 2006*
LEP	rs201381696	tcgggccgct	a/g	taagaggggc	4/12	↓	17	10^{-6}	Ожирение ⟶ *PA при курении*	Zhang et al., 1994; *Erlandsson et al., 2016*
	rs200487063	tgatcgggcc	g/a	ctataagagg	4/2	↑	6	10^{-6}	⟶ *PA*	*Cao et al., 2016*
	rs34104384	ccgctataag	a/t	ggggcgggca	4/3	↑	4	0.01		
GCG	rs183433761	gctggagagt	a/g	tataaaagca	0.9/1.6	↓	17	10^{-6}	⟶ *PA* (*предотвратимо вегетарианской диетой*)	*McCarty, 1999*
	rs757035851	tatataaaag	cag/–	tgcgccttgg	0.9/1.1	↓	3	10^{-3}		
PTPN22	rs185537537	aggtggttac	g/a	tactaatttc	3.6/2.6	↑	4	10^{-2}	⟵ *индуцируемый цитруллинизацией PA*	*Chang et al., 2016*
	rs375600763	gttacgtact	a/–	atttcctcta	3.6/4.4	↓	2	0.05	⟵ *индуцируемый маннозой PA*	*Sood et al., 2016*
HLA-A	rs776909339	caattcatga	g/a	tgacagtgtg	22/18	↑	4	10^{-3}	⟵ *PA via ингибирование NK-клеток*	*Zhang et al., 2007*
	rs369883779	gtttctccct	t/g	gtttctcaga	37/46	↓	4	10^{-3}	⟵ *PA via ниже аутоиммунная активность*	*Rutten et al., 2014*
	rs9260120	cgcggtcgct	c/c, (t)		12/14	↓	3	10^{-2}		

Примечание. wt – норма (референсный геном человека); mut – минорный аллель; «–» – делеция; K_D – константа диссоциации комплекса ТВР/ДНК (Savinkova et al., 2013); Δ – изменение: роста (↑), спада (↓); Z – Z-статистика; α – значимость (α = 1 – *p*, где *p* – вероятность, показанная на рисунке). Патологии: ALS – боковой амиотрофический склероз; PA – ревматоидный полиартрит. Риск развития PA: выше (⟶ PA), ниже (⟵ PA). Гены: *ACKR1* – гликопротеин D; *APOA1* – аполипопротеин A1; *DHFR* – дигидрофолатредуктаза; *F3* – тромбопластин; *F7* – проконвертин; *GCG* – глюкагон; *HBB* и *HBD* – β- и δ-цепи гемоглобина соответственно; *HLA-A* – главный комплекс гистосовместимости, I класс, α-цепь; *IL1B* – интерлейкин 1β; *INS* – инсулин; *LEP* – лептин; *MBL2* – маннозосвязывающий лектин; *MMP12* – металлоэластаза макрофагов; *NOS2* – индуцибельная синтаза оксида азота; *PTPN22* – лимфоидная фосфатаза; *SOD1* – Cu/Zn-супероксиддисмутаза; *TPI1* – триозофосфатизомераза. $ – SNP в промоторе гена *HBD*.



Пример применения Web-сервиса SNP_TATA_Z-tester к анализу района [–70; –20] (↔) промотора гена *MMP12* человека с известным SNP-маркером rs2276109 (штриховая стрелка) пониженного риска ревматоидного полиартрита (Liu et al., 2004) из-за дефицита металлоэластазы макрофага вследствие снижения сродства TATA-связывающего белка (TBP) к сайту TBP-связывания (рамка).

а – визуализация промотора гена *MMP12* человека с помощью Web-сервиса UCSC Genome Browser (Haeussler et al., 2015); *б* – rs2276109: строка DECISION в окне Result свидетельствует о согласии результата SNP_TATA_Z-tester с клиническими данными; *в* – анализ неаннотированного SNPrs572527200 (пунктирная стрелка на панели «*а*»), который может вызвать дефицит MMP12 и поэтому может быть кандидатным SNP-маркером снижения риска развития ревматоидного полиартрита.

обзор (Liu et al., 2004) об избытке MMP12 как физиологическом маркере повышения риска развития РА, где предлагается рассматривать SNP rs2276109, вызывающий дефицит MMP12, в качестве кандидатного SNP-маркера пониженного риска развития РА (см. таблицу).

Результаты и обсуждение

Проанализировано 227 SNP в промоторах 18 генов человека: *ACKR1*, *APOA1*, *DHFR*, *F3*, *F7*, *GCG*, *HBB*, *HBD*, *HLA-A*, *IL1B*, *INS*, *LEP*, *MBL2*, *MMP12*, *NOS2*, *PTPN22*, *SOD1* и *TPI1*, среди которых найдено 43 кандидатных SNP-маркера для РА, достоверно изменяющих сродство TBP/промотор (см. таблицу).

Рассмотрим наиболее подробно пример одного из предсказанных кандидатных SNP-маркеров РА для того, чтобы затем более кратко представлять остальные 42 кандидатных SNP-маркера РА благодаря аналогии с этим примером.

Ген *MMP12* несет известный SNP-маркер rs2276109 (рисунок, *а*) снижения риска широкого круга аутоиммунных заболеваний (Ponomarenko M. et al., 2016), таких как псориаз (Starodubtseva et al., 2011), астма (Hunninghake et al., 2009) и системная склеродерма (Manetti et al., 2010), в силу снижения сродства TBP к промотору этого гена и, как следствие, дефицита кодируемой им матриксной металлоэластазы макрофагов, обеспечивающей их доступ к клеткам-мишеням для аутоиммунной атаки. SNP_TATA_Z-tester дал верный прогноз дефицита MMP12 благодаря снижению сродства от 11 нМ TBP к норме этого промотора до 14 нМ к минорному аллелю этого SNP (см. рисунок, *б* и таблицу). Полученный результат указывает на адекватность SNP_TATA_Z-tester для изучения РА в соответствии с ранее показанным результатом для аутоиммунных заболеваний в целом (Ponomarenko M. et al., 2016).

Результатом поиска по ключевым словам в базе данных PubMed (Lu, 2011) был ретроспективный клинический обзор (Liu et al., 2004) об избытке MMP12 как физиологическом маркере повышенного риска развития РА (см. таблицу), который согласуется с эмпирически известным молекулярным механизмом влияния MMP12 на риск аутоиммунных заболеваний (Hunninghake et al., 2009; Manetti et al., 2010; Starodubtseva et al., 2011; Ponomarenko M. et al., 2016). Это позволило впервые предложить SNP rs2276109 в качестве кандидатного маркера снижения риска развития РА (см. таблицу).

Наконец, с помощью SNP_TATA_Z-tester оценены два неаннотированных SNP в районе [–70; –20] этого промотора (см. рисунок, *а*), один из которых, rs572527200, может влиять на экспрессию MMP12 (см. рисунок, *в*) так же, как известный SNP-маркер rs2276109 (см. рисунок, *б*). Поэтому rs572527200 тоже может быть кандидатным SNP-маркером снижения риска развития РА (см. таблицу). Другой неаннотированный SNP rs797029793 промотора гена *MMP12* недостоверно влияет на сродство TBP к нему, в силу чего он был исключен из дальнейшего анализа (данные не показаны). Тем не менее это не означает, что rs797029793 не может быть кандидатным SNP-

маркером РА: в частности, возможно его влияние на сайты связывания транскрипционных факторов, нуклеосомную упаковку промотора или сайты его эпигенетической регуляции (например, метилирования CpG-пар промотора) вне рамок этой работы. Для оценки более широкого спектра манифестаций SNP можно использовать общедоступные Web-сервисы, описанные в обзорах (Deplancke et al., 2016; Gunbin et al., 2018).

Гены *DHFR* (дигидрофолатредуктаза), ***NOS2*** (индуцибельная синтаза оксида азота) и ***F7*** (проконвертин) несут в промоторах известные SNP-маркеры rs10168 для устойчивости к метотрексату при лейкозе (Al-Shakfa et al., 2009), замену T→C в позиции –51 (здесь и далее –51T→C) для эпилепсии (Gonzalez-Martinez et al., 2009) и устойчивости к малярии (Burgner et al., 2003; Clark et al., 2003), а также замену –33A→C для умеренных кровотечений (Kavlie et al., 2003) из-за избытков DHFR и NOS2, а также дефицита F7 соответственно. Прежде всего, уровень DHFR у пациентов был пропорционален устойчивости этих пациентов к метотрексату при терапии РА (Bennike et al., 2017). Поэтому неаннотированный SNP rs750793297 (избыток DHFR) может быть кандидатным SNP-маркером повышенной устойчивости к метотрексату при терапии РА, тогда как неаннотированные SNP rs766799008, rs764508464 и rs754122321 (дефицит DHFR) – восприимчивости метотрексата (см. таблицу). Аналогично, –51T→C (избыток NOS2) в промоторе *NOS2* может быть кандидатным SNP-маркером повышенного риска развития РА, поскольку *NOS2* является геном-мишенью иммунохимиотерапии при РА (Ohtsuka et al., 2002).

Наконец, в порядке обсуждения дефицита проконвертина, который предсказан Web-сервисом SNP_TATA_Z-tester в случае замены –33A→C в промоторе гена *F7* человека (см. таблицу), мы нашли как клинические случаи РА при дефиците проконвертина (Zuber-Jerger et al., 2005), так и случаи применения экзогенного рекомбинантного активированного проконвертина для успешной терапии РА (Drobiecki et al., 2013). Как применение этих клинических наблюдений без каких-либо гипотез о возможных причинно-следственных связях мы предлагаем здесь кандидатный SNP-маркер *F7*: –33A→C предрасположенности к РА (см. таблицу).

Гены *F3* (тромбопластин), ***MBL2*** (маннозосвязывающий лектин) и ***APOA1*** (аполипопротеин A1) содержат известные SNP-маркеры rs563763767 для инфаркта миокарда и тромбоза (Arnaud et al., 2000), rs72661131 для преэклампсии (Sziller et al., 2007), инсульта (Cervera et al., 2010), вариабельного иммунодефицита (Boldt et al., 2006), а также –35A→C для стеатоза печени и ожирения (Matsunaga et al., 1999) в случаях избытка F3, дефицитов MBL2 и APOA1 соответственно.

Результатом поиска по ключевым словам в PubMed (Lu, 2011) были клинические данные (Lyberg et al., 1982) о сгущении крови как физиологическом маркере РА. Поскольку избыток тромбопластина как индуктора свертывания крови вызывает ее сгущение (Arnaud et al., 2000), то в рамках применимости клинических данных (Lyberg et al., 1982) без каких-либо дополнительных гипотез о причинно-следственных отношениях при патогенезе РА впервые предлагается рассмотреть известный клинический SNP-маркер rs563763767 повышенного риска сердечно-сосудистых заболеваний еще и как кандидатный SNP-маркер повышенного риска РА (см таблицу).

В работе (Nisihara et al., 2016) сообщается о дефиците MBL2 как о физиологическом маркере повышенного риска инфекционных заболеваний верхних дыхательных путей и мочевого тракта при РА. Найдены клинические данные (Troelsen et al., 2010) о росте риска сердечно-сосудистых заболеваний при РА в случаях как дефицита, так и избытка MBL2. На их основе мы предсказали кандидатные SNP-маркеры rs72661131, rs562962093 и rs567653539 перечисленных осложнений РА (см. таблицу).

Наконец ретроспективный клинический обзор (Kokkonen et al., 2017) указывает на дефицит APOA1 как физиологический маркер повышенного риска развития РА с сердечно-сосудистыми осложнениями у молодых курильщиков с избыточным весом. В связи с этим известный SNP-маркер –35A→C дефицита APOA1, стеатоза печени и ожирения впервые предложен как кандидатный SNP-маркер риска развития РА и его осложнений у молодых курильщиков (см. таблицу).

Ген *TPI1* содержит известный SNP-маркер rs1800202 для дефицита кодируемой этим геном триозофосфатизомеразы и, как следствие, гемолитической анемии (Watanabe et al., 1996) и нейромышечных заболеваний (Vives-Corrons et al., 1978). Рядом с этим известным клиническим SNP-маркером нашли неаннотированный SNP rs781835924, способный также вызвать дефицит TPI1 и поэтому быть кандидатным SNP-маркером тех же патологий (см. таблицу).

Поиск по ключевым словам в PubMed (Lu, 2011) выявил ретроспективный клинический обзор (Richardson et al., 1984) о способности нейромышечных заболеваний маскировать симптомы ювенильного РА. Поэтому мы впервые предложили кандидатные SNP-маркеры rs1800202 и rs781835924 для РА у детей до 16 лет, симптомы которого схожи с симптомами нейромышечных заболеваний, что может затруднять своевременную диагностику ювенильного РА (см. таблицу).

Гены *SOD1* (Cu/Zn-супероксиддисмутаза), ***INS*** (инсулин) и ***IL1B*** (интерлейкин 1β) несут в промоторах известные клинические SNP-маркеры rs7277748 бокового амиотрофического склероза, ALS (Niemann et al., 2007), rs1143627 для гастрита, ожирения, рака печени, легких, желудка, язвы желудка, психических расстройств (Ponomarenko et al., 2015), а также rs5505 неонатального сахарного диабета (Landrum et al., 2014), которые вызывают дефицит SOD1, избыток IL1B и гиперинсулинемию соответственно. Рядом с клиническим SNP-маркером rs5505 есть неаннотированный SNP rs563207167, способный вызвать гиперинсулинемию и поэтому быть кандидатным SNP-маркером тех же заболеваний (см. таблицу).

Результатом поиска по ключевым словам в PubMed (Lu, 2011) оказались ретроспективные клинические обзоры (Abrahamson, 1952; Staron et al., 2012; Olkkonen et al., 2015) о характерных болевых ощущениях при РА, физиологическими маркерами которых служат дефицит SOD1, избыток IL1B и гиперинсулинемия. Прежде всего, уровень боли при РА достоверно негативно коррелирует с уровнем Cu/Zn-супероксиддисмутазы (Staron et al., 2012), степень

ее циркадных обострений – с уровнем интерлейкина 1β (Olkkonen et al., 2015), а боли из-за дистрофии мышц в силу нарушения метаболизма глюкозы при гиперинсулинемии маскируют РА (Abrahamson, 1952). Поэтому нами дан прогноз кандидатных SNP-маркеров rs7277748 и rs1143627 уровня и циркадности боли при РА соответственно, а также rs5505 и rs563207167 как маркеров усиления боли при РА в случае избытка инсулина (см. таблицу). Это означает, что кандидатный SNP-маркер rs1143627 может способствовать снижению доз обезболивающих препаратов путем ограничения их приема лишь в определенное время суток перед циркадным появлением боли (Olkkonen et al., 2015), тогда как rs5505 и rs563207167 – ослаблению болевых ощущений при РА благодаря применению жирной низкоуглеводной фруктово-овощной диеты без кофеина (Abrahamson, 1952).

Гены *HBB* и ***HBD*** несут известные SNP-маркеры rs397509430, rs33980857, rs34598529, rs33931746, rs33981098, rs34500389 и rs35518301 для дефицита β- и δ-цепей гемоглобина, связанных с устойчивостью к малярии, а также с талассемией (Martiney et al., 1996). Рядом с этими SNP есть неаннотированные rs63750953, rs281864525 и rs34166473, способные также влиять на экспрессию *HBB* и *HBD* и, соответственно, быть кандидатными SNP-маркерами тех же патологий (см. таблицу).

Известен клинический случай РА при талассемии (Giakoumi et al., 2005), который позволяет рассматривать rs397509430, rs33980857, rs34598529, rs33931746, rs33981098, rs34500389, rs35518301, rs63750953, rs281864525 и rs34166473 в качестве кандидатных SNP-маркеров риска РА (см. таблицу).

Ген *ACKR1* (гликопротеин D) имеет в промоторе клинический SNP-маркер rs2814778 устойчивости к малярии (Michon et al., 2001) и повышенного риска лейкопении (Nalls et al., 2008) из-за снижения сродства TBP к промотору этого гена и, соответственно, дефицита его продуктов. Согласно ретроспективному метаанализу педиатров (Jones et al., 2006), лейкопения оказалась самым достоверным среди предикторов перерождения ювенильного РА в лейкемию у детей до 16 лет. Поэтому нами был предсказан кандидатный SNP-маркер rs2814778 повышенного риска перерождения ювенильного РА в лейкемию у детей до 16 лет.

Гены *LEP* (гормон сытости лептин), ***GCG*** (гормон голода глюкагон), ***HLA-A*** (главный комплекс гистосовместимости, I класс, α-цепь) и ***PTPN22*** (лимфоидная фосфатаза) не содержат в промоторах известные SNP-маркеры для изменения сродства TBP к этим промоторам, ассоциированного с заболеваниями человека (Amberger et al., 2015). С использованием нашего Web-сервиса SNP_TATA_Z-tester обнаружены неаннотированные SNP rs200487063, rs34104384, rs185537537 и rs776909339, способные увеличить сродство TBP к промоторам этих генов, а также rs201381696, rs183433761, rs757035851, rs375600763, rs369883779, rs9260120 и rs41560714, которые могут достоверно снижать такое сродство (см. таблицу).

Клинический анализ данных у 126 нынешних и 177 бывших курильщиков в сравнении с данными у 240 некурящих (Erlandsson et al., 2016) установил дефицит лептина и избыточный вес вследствие дефицита LEP (Zhang et al., 1994) в качестве факторов риска развития РА у курильщиков. В связи с этим был предложен кандидатный SNP-маркер rs201381696 риска развития РА у курильщиков (см. таблицу). В свою очередь, ретроспективный метаанализ клинических данных (Cao et al., 2016) выявил избыток лептина как безусловный фактор риска развития РА. На этой основе мы прогнозируем кандидатные SNP-маркеры rs200487063 и rs34104384 повышенного риска развития РА. Представляется, что индивиды-носители кандидатных SNP-маркеров rs201381696, rs200487063 и rs34104384 риска развития РА из-за отклонения уровня лептина от нормы могут снизить этот риск, поддерживая уровень лептина в рамках нормы путем соответствующей диеты.

В свою очередь, обзор диетологов (McCarty, 1999) сообщает о постпрандиальном (т. е. после еды) снижении отношения уровней глюкогона к инсулину в крови как о факторе риска развития РА, который можно снизить за счет вегетарианской диеты. Таким образом, мы предлагаем кандидатные SNP-маркеры rs183433761 и rs757035851 постпрандиального возрастания риска развития РА, который можно понизить вегетарианской диетой (см. таблицу).

Наконец, опубликованы клинические данные об избытке (Chang et al., 2016) и дефиците (Sood et al., 2016) лимфоидной фосфатазы, а также об избытке (Zhang et al., 2007) и дефиците (Rutten et al., 2014) HLA-A как физиологических маркерах РА, соответствующих снижению аутоиммунных повреждений, запускаемых цитруллинацией и маннозой, а также ингибированию NK-клеток натуральных Т-киллеров и снижению уровня аутоиммунной активности в целом. В связи с этим мы впервые предлагаем здесь кандидатные SNP-маркеры rs185537537, rs375600763, rs776909339, rs369883779, rs9260120 и rs41560714 для снижения риска РА (см. таблицу). Наш прогноз согласуется с выводом независимого ретроспективного клинико-фармакологического метаанализа (Okada et al., 2014), который указал на *PTPN22* и *HLA-A* как самые перспективные гены-мишени для заместительной и ингибиторной форм терапии при РА.

Заключение

Исследовано 227 SNP в границах промоторов 18 генов человека, большинство (184) из которых были исключены как недостаточно обоснованные для их дальнейшей экспериментальной и клинической проверки в качестве кандидатных SNP-маркеров РА в рамках современных биологических знаний. В результате было предсказано 43 кандидатных SNP-маркера РА, каждый из которых детально обоснован на молекулярном уровне оценками величин сродства TBP к анцестральным и минорным аллелям соответствующих промоторов, уровнями статистической значимости различия между ними, а также на фенотипическом уровне, в том числе возможности снижения риска РА за счет выбора того или иного образа жизни (например, соответствующей диеты). Эти прогнозы могут удешевить клинический поиск биомедицинских SNP-маркеров с помощью когорт больных и условно здоровых добровольцев благодаря исключению нейтральных SNP.

Список литературы / References

Abrahamson E.M. Hyperinsulinism – a factor in rheumatoid arthritis. Am. J. Dig. Dis. 952;19(1):1-4.

Al-Shakfa F., Dulucq S., Brukner I., Milacic I., Ansari M., Beaulieu P., Moghrabi A., Laverdiere C., Sallan S.E., Silverman L.B., Neuberg D., Kutok J.L., Sinnett D., Krajinovic M. DNA variants in region for noncoding interfering transcript of dihydrofolate reductase gene and outcome in childhood acute lymphoblastic leukemia. Clin. Cancer Res. 2009;15:6931-6938. DOI 10.1158/1078-0432.CCR-09-0641.

Amberger J.S., Bocchini C.A., Schiettecatte F., Scott A.F., Hamosh A. OMIM.org: Online mendelian inheritance in man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res. 2015; 43(Database issue):D789-D798. DOI 10.1093/nar/gku1205.

Arkova O.V., Kuznetsov N.A., Fedorova O.S., Kolchanov N.A., Savinkova L.K. Real-time interaction between TBP and the TATA box of the human triosephosphate isomerase gene promoter in the norm and pathology. Acta Naturae. 2014;6(2):36-40.

Arkova O.V., Ponomarenko M.P., Rasskazov D.A., Drachkova I.A., Arshinova T.V., Ponomarenko P.M., Savinkova L.K., Kolchanov N.A. Obesity-related known and candidate SNP markers can significantly change affinity of TATA-binding protein for human gene promoters. BMC Genomics. 2015;16(Suppl. 13):S5. DOI 10.1186/1471-2164-16-S13-S5.

Arnaud E., Barbalat V., Nicaud V., Cambien F., Evans A., Morrison C., Arveiler D., Luc G., Ruidavets J.B., Emmerich J., Fiessinger J.N., Aiach M. Polymorphisms in the 5′ regulatory region of the tissue factor gene and the risk of myocardial infarction and venous thromboembolism: the ECTIM and PATHROS studies. Etude Cas-Temoins de l'Infarctus duMyocarde. Paris thrombosis case–control study. Arterioscler. Thromb. Vasc. Biol. 2000;20:892-898. DOI 10.1161/01.atv.20.3.892.

Bennike T.B., Ellingsen T., Glerup H., Bonderup O.K., Carlsen T.G., Meyer M.K., Bogsted M., Christiansen G., Birkelund S., Andersen V., Stensballe A. Proteome analysis of rheumatoid arthritis gut mucosa. J. Proteome Res. 2017;16(1):346-354. DOI 10.1021/acs.jproteome.6b00598.

Boldt A.B., Culpi L., Tsuneto L.T., de Souza I.R., Kun J.F., Petzl-Erler M.L. Diversity of the MBL2 gene in various Brazilian populations and the case of selection at the mannose-binding lectin locus. Hum. Immunol. 2006;67(9):722-734. DOI 10.1016/j.humimm.2006.05.009.

Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. 1990;212(4):563-578. DOI 10.1016/0022-2836(90)90223-9.

Burgner D., Rockett K., Ackerman H., Hull J., Usen S., Pinder M., Kwiatkowski D.P. Haplotypic relationship between SNP and microsatellite markers at the NOS2A locus in two populations. Genes Immun. 2003;4(7):506-514. DOI 10.1038/sj.gene.6364022.

Cao H., Lin J., Chen W., Xu G., Sun C. Baseline adiponectin and leptin levels in predicting an increased risk of disease activity in rheumatoid arthritis: A meta-analysis and systematic review. Autoimmunity. 2016;49(8):547-553. DOI 10.1080/08916934.2016.1230847.

Cervera A., Planas A.M., Justicia C., Urra X., Jensenius J.C., Torres F., Lozano F., Chamorro A. Genetically-defined deficiency of mannose-binding lectin is associated with protection after experimental stroke in mice and outcome in human stroke. PLoS One. 2010;5(2):e8433. DOI 10.1371/journal.pone.0008433.

Chadaeva I.V., Ponomarenko M.P., Rasskazov D.A., Sharypova E.B., Kashina E.V., Matveeva M.Yu., Arshinova T.V., Ponomarenko P.M., Arkova O.V., Bondar N.P., Savinkova L.K., Kolchanov N.A. Candidate SNP markers of aggressiveness-related complications and comorbidities of genetic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. BMC Genomics. 2016;17(Suppl. 14):995. DOI 10.1186/s12864-016-3353-3.

Chadaeva I., Ponomarenko P., Rasskazov D., Sharypova E., Kashina E., Kleshchev M., Ponomarenko M., Naumenko V., Savinkova L., Kolchanov N., Osadchuk L., Osadchuk A. Natural selection equally supports the human tendencies in subordination and domination: a genome-wide study with in silico confirmation and in vivo validation in mice. Front. Genet. 2019;10:73.

Chadaeva I.V., Ponomarenko P.M., Rasskazov D.A., Sharypova E.B., Kashina E.V., Zhechev D.A., Drachkova I.A., Arkova O.V., Savinkova L.K., Ponomarenko M.P., Kolchanov N.A., Osadchuk L.V., Osadchuk A.V. Candidate SNP markers of reproductive potential are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. BMC Genomics. 2018;19(Suppl. 3). DOI 10.1186/s12864-018-4478-3.

Chang H.H., Liu G.Y., Dwivedi N., Sun B., Okamoto Y., Kinslow J.D., Deane K.D., Demoruelle M.K., Norris J.M., Thompson P.R., Sparks J.A., Rao D.A., Karlson E.W., Hung H.C., Holers V.M., Ho I.C. A molecular signature of preclinical rheumatoid arthritis triggered by dysregulated PTPN22. JCI Insight. 2016;1(17):e90045. DOI 10.1172/jci.insight.90045.

Clark I.A., Rockett K.A., Burgner D. Genes, nitric oxide and malaria in African children. Trends Parasitol. 2003;19(8):335-337. DOI 10.1016/S1471-4922(03)00147-8.

Deplancke B., Alpern D., Gardeux V. The genetics of transcription factor DNA binding variation. Cell. 2016;166(3):538-554. DOI 10.1016/j.cell.2016.07.01.

Drachkova I., Savinkova L., Arshinova T., Ponomarenko M., Peltek S., Kolchanov N. The mechanism by which TATA-box polymorphisms associated with human hereditary diseases influence interactions with the TATA-binding protein. Hum. Mutat. 2014;35(5):601-608. DOI 10.1002/humu.22535.

Drobiecki A., Pasiarski M., Hus I., Sokolowska B., Watek M. Acquired hemophilia in the patient suffering from rheumatoid arthritis: case report. Blood Coagul. Fibrinolysis. 2013;24(8):874-880. DOI 10.1097/mbc.0b013e3283646635.

Erlandsson M.C., Doria Medina R., Töyra Silfversward S., Bokarewa M.I. Smoking functions as a negative regulator of IGF1 and impairs adipokine network in patients with rheumatoid arthritis. Mediators Inflamm. 2016;2016:3082820. DOI 10.1155/2016/3082820.

Giakoumi X., Tsironi M., Floudas C., Polymeropoylos E., Papalambros E., Aessopos A. Rheumatoid arthritis in thalassemia intermedia: coincidence or association? Isr. Med. Assoc. J. 2005;7(10):667-669.

Ginsburg G.S., Willard H.F. Genomic and personalized medicine: foundations and applications. Transl. Res. 2009;154(6):277-287. DOI 10.1016/j.trsl.2009.09.005.

Gonzalez-Martinez J.A., Moddel G., Ying Z., Prayson R.A., Bingaman W.E., Najm I.M. Neuronal nitric oxide synthase expression in resected epileptic dysplastic neocortex. J Neurosurg. 2009;110(2):343-349. DOI 10.3171/2008.6.17608.

Gunbin K.V., Ponomarenko M.P., Suslov V.V., Gusev F., Fedonin G.G., Rogaev E.I. Evolution of brain active gene promoters in human lineage towards the increased plasticity of gene regulation. Mol. Neurobiol. 2018;55(3):1871-1904. DOI 10.1007/s12035-017-0427-4.

Haeussler M., Raney B.J., Hinrichs A.S., Clawson H., Zweig A.S., Karolchik D., Casper J., Speir M.L., Haussler D., Kent W.J. Navigating protected genomics data with UCSC genome browser in a box. Bioinformatics. 2015;31(5):764-766. DOI 10.1093/bioinformatics/btu712.

Hahn S., Buratowski S., Sharp P.A., Guarente L. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. Proc. Natl. Acad. Sci. USA. 1989;86(15):5718-5722.

Haldane J.B.S. The cost of natural selection. J. Genet. 1957;55:511-524.

Ho M.F., Weinshilboum R.M. Immune mediator pharmacogenomics: TCL1A SNPs and estrogen-dependent regulation of inflammation. J. Nat. Sci. 2017;3(8):e416.

Hunninghake G.M., Cho M.H., Tesfaigzi Y., Soto-Quiros M.E., Avila L., Lasky-Su J., Stidley C., Melen E., Soderhall C., Hallberg J., Kull I., Kere J., Svartengren M., Pershagen G., Wickman M., Lange C., Demeo D.L., Hersh C.P., Klanderman B.J., Raby B.A., Sparrow D., Shapiro S.D., Silverman E.K., Litonjua A.A., Weiss S.T., Celedon J.C. MMP12, lung function, and COPD in high-risk populations. N. Engl. J. Med. 2009;361(27):2599-2608. DOI 10.1056/nejmoa0904006.

Jones O.Y., Spencer C.H., Bowyer S.L., Dent P.B., Gottlieb B.S., Rabinovich C.E. A multicenter case-control study on predictive factors distinguishing childhood leukemia from juvenile rheumatoid arthritis. Pediatrics. 2006;117(5):e840-e844. DOI 10.1542/peds.2005-1515.

Karas H., Knuuppel R., Schulz W., Sklenar H., Wingender E. Combining structural analysis of DNA with search routines for the detection of transcription regulatory elements. Comput. Appl. Biosci. 1996; 12(5):441-446.

Karlson E.W., Mandl L.A., Hankinson S.E., Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. Arthritis Rheum. 2004;50(11):3458-3467. DOI 10.1002/art.20621.

Kavlie A., Hiltunen L., Rasi V., Prydz H. Two novel mutations in the human coagulation factor VII promoter. Thromb. Haemost. 2003; 90(2):194-205. DOI 10.1160/th02-09-0050.

Kimura M. Evolutionary rate at the molecular level. Nature. 1968; 217(5129):624-626. DOI 10.1038/217624a0.

Klareskog L., Padyukov L., Lorentzen J., Alfredsson L. Mechanisms of disease: genetic susceptibility and environmental triggers in the development of rheumatoid arthritis. Nat. Clin. Pract. Rheumatol. 2006;2(8):425-433. DOI 10.1038/ncprheum0249.

Kokkonen H., Stenlund H., Rantapaa-Dahlqvist S. Cardiovascular risk factors predate the onset of symptoms of rheumatoid arthritis: a nested case-control study. Arthritis Res. Ther. 2017;19(1):148. DOI 10.1186/s13075-017-1351-8.

Köller M., Nöbauer-Huhmann I. Early arthritis: action desired – treatment required. Wien Med. Wochenschr. 2009;159(3-4):66-69. DOI 10.1007/s10354-009-0653-0.

Krasselt M., Baerwald C. Sex, symptom severity, and quality of life in rheumatology. Clin. Rev. Allergy Immunol. 2017. DOI 10.1007/s12016-017-8631-6.

Landrum M.J., Lee J.M., Riley G.R., Jang W., Rubinstein W.S., Church D.M., Maglott D.R. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014;42(Database issue):D980-D985. DOI 10.1093/nar/gkt1113.

Lim D.S., Bae Y.S. Metastatic lymph node 51 and fibroblast-like synoviocyte hyperproliferation in rheumatoid arthritis pathogenesis. Rheumatol. Int. 2011;31(7):843-847. DOI 10.1007/s00296-011-1818-x.

Liu M., Sun H., Wang X., Koike T., Mishima H., Ikeda K., Watanabe T., Ochiai N., Fan J. Association of increased expression of macrophage elastase (matrix metalloproteinase 12) with rheumatoid arthritis. Arthritis Rheum. 2004;50(10):3112-3117. DOI 10.1002/art.20567.

Lu Z. PubMed and beyond: a survey of web tools for searching biomedical literature. Database (Oxford). 2011. 2011:baq036. DOI 10.1093/database/baq036. Print 2011.

Lyberg T., Prydz H., Baklien K., Hoyeraal H.M. Effect of immune complex-containing sera from patients with rheumatic diseases on thromboplastin activity of monocytes. Thromb. Res. 1982;25(3): 193-202. DOI 10.1016/0049-3848(82)90238-9.

Malm K., Bremander A., Arvidsson B., Andersson M.L., Bergman S., Larsson I. The influence of lifestyle habits on quality of life in patients with established rheumatoid arthritis-A constant balancing between ideality and reality. Int. J. Qual. Stud. Health Well-being. 2016;11:30534. DOI 10.3402/qhw.v11.30534.

Manetti M., Ibba-Manneschi L., Fatini C., Guiducci S., Cuomo G., Bonino C., Bazzichi L., Liakouli V., Giacomelli R., Abbate R., Bombardieri S., Montecucco C., Valentini G., Matucci-Cerinic M. Association of a functional polymorphism in the matrix metalloproteinase-12 promoter region with systemic sclerosis in an Italian population. J. Rheumatol. 2010;37(9):1852-1857. DOI 10.3899/jrheum.100237.

Martianov I., Viville S., Davidson I. RNA polymerase II transcription in murine cells lacking the TATA binding protein. Science. 2002;298(5595):1036-1039. DOI 10.1126/science.1076327.

Martiney J.A., Cerami A., Slater A.F. Inhibition of hemozoin formation in *Plasmodium falciparum* trophozoite extracts by heme analogs: possible implication in the resistance to malaria conferred by the beta-thalassemia trait. Mol. Med. 1996;2(2):236-246.

Matsunaga A., Sasaki J., Han H., Huang W., Kugi M., Koga T., Ichiki S., Shinkawa T., Arakawa K. Compound heterozygosity for an apolipoprotein A1 gene promoter mutation and a structural nonsense mutation with apolipoprotein A1 deficiency. Arterioscler. Thromb. Vasc. Biol. 1999;19(2):348-355. DOI 10.1161/01.atv.19.2.348.

McCarty M.F. Vegan proteins may reduce risk of cancer, obesity, and cardiovascular disease by promoting increased glucagon activity. Med. Hypotheses. 1999;53(6):459-485. DOI 10.1054/mehy.1999.0784.

Michon P., Woolley I., Wood E.M., Kastens W., Zimmerman P.A., Adams J.H. Duffy-null promoter heterozygosity reduces DARC expression and abrogates adhesion of the *P. vivax* ligand required for blood-stage infection. FEBS Lett. 2001;495(1-2):111-114. DOI 10.1016/S0014-5793(01)02370-5.

Mogno I., Vallania F., Mitra R.D., Cohen B.A. TATA is a modular component of synthetic promoters. Genome Res. 2010;20(10): 1391-1397. DOI 10.1101/gr.106732.110.

Muller F., Lakatos L., Dantonel J., Strahle U., Tora L. TBP is not universally required for zygotic RNA polymerase II transcription in zebrafish. Curr. Biol. 2001;11(4):282-287. DOI 10.1016/S0960-9822(01)00076-8.

Nair N., Wilson A.G., Barton A. DNA methylation as a marker of response in rheumatoid arthritis. Pharmacogenomics. 2017;18(14): 1323-1332. DOI 10.2217/pgs-2016-0195.

Nalls M.A., Wilson J.G., Patterson N.J., Tandon A., Zmuda J.M., Huntsman S., Garcia M., Hu D., Li R., Beamer B.A., Patel K.V., Akylbekova E.L., Files J.C., Hardy C.L., Buxbaum S.G., Taylor H.A., Reich D., Harris T.B., Ziv E. Admixture mapping of white cell count: genetic locus responsible for lower white blood cell count in the Health ABC and Jackson Heart studies. Am. J. Hum. Genet. 2008;82(1):81-87. DOI 10.1016/j.ajhg.2007.09.003.

Niemann S., Broom W.J., Brown R.H. Jr. Analysis of a genetic defect in the TATA box of the *SOD1* gene in a patient with familial amyotrophic lateral sclerosis. Muscle Nerve. 2007;36(5):704-707. DOI 10.1002/mus.20855.

Nisihara R., Skare T., Capeletto C.M., Moreira L., Goeldner I., Messias-Reason I., Utiyama S.R. Mannose binding lectin deficiency and susceptibility to infections in patients with rheumatoid arthritis. Rheumatology (Oxford). 2016;55(5):951-952. DOI 10.1093/rheumatology/kev413.

Ohtsuka M., Konno F., Honda H., Oikawa T., Ishikawa M., Iwase N., Isomae K., Ishii F., Hemmi H., Sato S. PPA250 [3-(2,4-difluorophenyl)-6-[2-[4-(1H-imidazol-1-ylmethyl) phenoxy]ethoxy]-2-phenylpyridine], a novel orally effective inhibitor of the dimerization of inducible nitric-oxide synthase, exhibits an anti-inflammatory effect in animal models of chronic arthritis. J. Pharmacol. Exp. Ther. 2002; 303(1):52-57. DOI 10.1124/jpet.102.035857.

Okada Y., Wu D., Trynka G., Raj T., Terao C., Ikari K., Kochi Y., Ohmura K., Suzuki A., Yoshida S., Graham R.R., Manoharan A., Ortmann W., Bhangale T., Denny J.C., Carroll R.J., Eyler A.E., Greenberg J.D., Kremer J.M., Pappas D.A., Jiang L., Yin J., Ye L., Su D.F., Yang J., Xie G., Keystone E., Westra H.J., Esko T., Metspalu A., Zhou X., Gupta N., Mirel D., Stahl E.A., Diogo D., Cui J., Liao K., Guo M.H., Myouzen K., Kawaguchi T., Coenen M.J., van Riel P.L., van de Laar M.A., Guchelaar H.J., Huizinga T.W., Dieude P., Mariette X., Bridges S.L. Jr., Zhernakova A., Toes R.E., Tak P.P., Miceli-

Richard C., Bang S.Y., Lee H.S., Martin J., Gonzalez-Gay M.A., Rodriguez-Rodriguez L., Rantapaa-Dahlqvist S., Arlestig L., Choi H.K., Kamatani Y., Galan P., Lathrop M., RACI consortium, GARNET consortium, Eyre S., Bowes J., Barton A., de Vries N., Moreland L.W., Criswell L.A., Karlson E.W., Taniguchi A., Yamada R., Kubo M., Liu J.S., Bae S.C., Worthington J., Padyukov L., Klareskog L., Gregersen P.K., Raychaudhuri S., Stranger B.E., De Jager P.L., Franke L., Visscher P.M., Brown M.A., Yamanaka H., Mimori T., Takahashi A., Xu H., Behrens T.W., Siminovitch K.A., Momohara S., Matsuda F., Yamamoto K., Plenge R.M. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014;506(7488):376-381. DOI 10.1038/nature12873.

Olkkonen J., Kouri V.P., Hynninen J., Konttinen Y.T., Mandelin J. Differentially expressed in chondrocytes 2 (DEC2) increases the expression of IL-1β and is abundantly present in synovial membrane in rheumatoid arthritis. PLoS One. 2015;10(12):e0145279. DOI 10.1371/journal.pone.0145279.

Ponomarenko M., Arkova O., Rasskazov D., Ponomarenko P., Savinkova L., Kolchanov N. Candidate SNP markers of gender-biased autoimmune complications of monogenic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. Front. Immunol. 2016;7:130. DOI 10.3389/fimmu.2016.00130.

Ponomarenko M., Mironova V., Gunbin K., Savinkova L. Hogness box. In: Maloy S., Hughes K., (Eds.) Brenner's Encyclopedia of Genetics. Vol. 3. San Diego: Acad. Press; Elsevier Inc., 2013;491-494. DOI 10.1016/B978-0-12- 374984-0.00720-8.

Ponomarenko M., Rasskazov D., Arkova O., Ponomarenko P., Suslov V., Savinkova L., Kolchanov N. How to use SNP_TATA_Comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter. BioMed Res. Int. 2015;2015:359835. DOI 10.1155/2015/359835.

Ponomarenko P., Chadaeva I., Rasskazov D.A., Sharypova E., Kashina E.V., Drachkova I., Zhechev D., Ponomarenko M.P., Savinkova L.K., Kolchanov N. Candidate SNP markers of familial and sporadic Alzheimer's diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. Front. Aging Neurosci. 2017;9:231. DOI 10.3389/fnagi.2017.00231.

Ponomarenko P.M., Ponomarenko M.P., Drachkova I.A., Lysova M.V., Arshinova T.V., Savinkova L.K., Kolchanov N.A. Prediction of the affinity of the TATA-binding protein to TATA boxes with single nucleotide polymorphisms. Mol. Biol. (Mosk.). 2009;43(3):472-479. DOI 10.1134/S0026893309030157.

Ponomarenko P., Rasskazov D., Suslov V., Sharypova E., Savinkova L., Podkolodnaya O., Podkolodny N., Tverdokhleb N., Chadaeva I., Ponomarenko M., Kolchanov N. Candidate SNP markers of chronopathologies are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. BioMed Res. Int. 2016;2016:8642703. DOI 10.1155/2016/8642703.

Ponomarenko P.M., Savinkova L.K., Drachkova I.A., Lysova M.V., Arshinova T.V., Ponomarenko M.P., Kolchanov N.A. A step-by-step model of TBP/TATA box binding allows predicting human hereditary diseases by single nucleotide polymorphism. Dokl. Biochem. Biophys. 2008;419:88-92. DOI 10.1134/S1607672908020117.

Ponomarenko P.M., Suslov V.V., Savinkova L.K., Ponomarenko M.P., Kolchanov N.A. A precise equilibrium equation for four steps of binding between TBP and TATA-box allows for the prediction of phenotypical expression upon mutation. Biophysics. 2010;55(3): 358-369. DOI 10.1134/S0006350910030036.

Richardson M.L., Helms C.A., Vogler J.B. 3rd, Genant H.K. Skeletal changes in neuromuscular disorders mimicking juvenile rheumatoid arthritis and hemophilia. AJR. Am. J. Roentgenol. 1984;143(4):893-897. DOI 10.2214/ajr.143.4.893.

Rutten M.J., Dijk F., Savci-Heijink C.D., Buist M.R., Kenter G.G., van de Vijver M.J., Jordanova E.S. HLA-G expression is an independent predictor for improved survival in high grade ovarian carcinomas. J. Immunol. Res. 2014;2014:274584. DOI 10.1155/2014/274584.

Sakkas L.I., Daoussis D., Liossis S.N., Bogdanos D.P. The infectious basis of ACPA-positive rheumatoid arthritis. Front. Microbiol. 2017; 8:1853. DOI 10.3389/fmicb.2017.01853.

Sarkander J., Hojyo S., Tokoyoda K. Vaccination to gain humoral immune memory. Clin. Transl. Immunol. 2016;5(12):e120. DOI 10.1038/cti.2016.81.

Sato K., Takahashi N., Kato T., Matsuda Y., Yokoji M., Yamada M., Nakajima T., Kondo N., Endo N., Yamamoto R., Noiri Y., Ohno H., Yamazaki K. Aggravation of collagen-induced arthritis by orally administered Porphyromonas gingivalis through modulation of the gut microbiota and gut immune system. Sci. Rep. 2017;317(1):6955. DOI 10.1038/s41598-017-07196-7.

Savinkova L., Drachkova I., Arshinova T., Ponomarenko P., Ponomarenko M., Kolchanov N. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. PLoS One. 2013;8(2):e54626. DOI 10.1371/journal.pone.0054626.

Scott D.L., Wolfe F., Huizinga T.W. Rheumatoid arthritis. Lancet. 2010; 376(9746):1094-1108. DOI 10.1016/S0140-6736(10)60826-4.

Sherry S.T., Ward M.H., Kholodov M., Baker J., Phan L., Smigielski E.M., Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29(1):308-311. DOI 10.1093/nar/29.1.308.

Smolen J.S., Aletaha D. Rheumatoid arthritis therapy reappraisal: strategies, opportunities and challenges. Nat. Rev. Rheumatol. 2015; 11(5):276-289. DOI 10.1038/nrrheum.2015.8.

Smolen J.S., Aletaha D., McInnes I.B. Rheumatoid arthritis. Lancet. 2016;388(10055):2023-2038. DOI 10.1016/S0140-6736(16)30173-8.

Sood S., Brownlie R.J., Garcia C., Cowan G., Salmond R.J., Sakaguchi S., Zamoyska R. Loss of the protein tyrosine phosphatase PTPN22 reduces mannan-induced autoimmune arthritis in SKG mice. J. Immunol. 2016;197(2):429-440. DOI 10.4049/jimmunol.1502656.

Starodubtseva N.L., Sobolev V.V., Soboleva A.G., Nikolaev A.A., Bruskin S.A. Genes expression of metalloproteinases (MMP-1, MMP-2, MMP-9, and MMP-12) associated with psoriasis. Russ. J. Genet. 2011;47(9):1117-1123. DOI 10.1134/S102279541109016X.

Staron A., Mąkosa G., Koter-Michalak M. Oxidative stress in erythrocytes from patients with rheumatoid arthritis. Rheumatol. Int. 2012; 32(2):331-334. DOI 10.1007/s00296-010-1611-2.

Suslov V.V., Ponomarenko P.M., Ponomarenko M.P., Drachkova I.A., Arshinova T.V., Savinkova L.K., Kolchanov N.A. TATA box polymorphisms in genes of commercial and laboratory animals and plants associated with selectively valuable traits. Russ. J. Genet. 2010a;46(4):394-403. DOI 10.1134/S1022795410040022.

Suslov V.V., Ponomarenko P.M., Efimov V.M., Savinkova L.K., Ponomarenko M.P., Kolchanov N.A. SNPs in the HIV-1 TATA box and the AIDS pandemic. J. Bioinform. Comput. Biol. 2010b;8(3):607-625. DOI 10.1142/S0219720010004677.

Sziller I., Babula O., Hupuczi P., Nagy B., Rigo B., Szabo G., Papp Z., Linhares I.M., Witkin S.S. Mannose-binding lectin (MBL) codon 54 gene polymorphism protects against development of pre-eclampsia, HELLP syndrome and pre-eclampsia-associated intrauterine growth restriction. Mol. Hum. Reprod. 2007;13(4):281-285. DOI 10.1093/molehr/gam003.

Telenti A., Pierce L.C., Biggs W.H., di Iulio J., Wong E.H., Fabani M.M., Kirkness E.F., Moustafa A., Shah N., Xie C., Brewerton S.C., Bulsara N., Garner C., Metzker G., Sandoval E., Perkins B.A., Och F.J., Turpaz Y., Venter J.C. Deep sequencing of 10,000 human genomes. Proc. Natl. Acad. Sci. USA. 2016;113(42):11901-11906. DOI 10.1073/pnas.1613365113.

Troelsen L.N., Garred P., Christiansen B., Torp-Pedersen C., Christensen I.J., Narvestad E., Jacobsen S. Double role of mannose-binding lectin in relation to carotid intima-media thickness in patients with rheumatoid arthritis. Mol. Immunol. 2010;47(4):713-718. DOI 10.1016/j.molimm.2009.10.021.

Trovato G.M. Sustainable medical research by effective and comprehensive medical skills: overcoming the frontiers by predictive,

preventive and personalized medicine. EPMA J. 2014;5(1):14. DOI 10.1186/1878-5085-5-14.

Varzari A., Deyneko I.V., Tudor E., Turcan S. Polymorphisms of glutathione S-transferase and methylenetetrahydrofolate reductase genes in Moldavian patients with ulcerative colitis: Genotype-phenotype correlation. Meta Gene. 2016;7:76-82. DOI 10.1016/j.mgene.2015.12.002.

Vives-Corrons J.L., Rubinson-Skala H., Mateo M., Estella J., Feliu E., Dreyfus J.C. Triosephosphate isomerase deficiency with hemolytic anemia and severe neuromuscular disease: familial and biochemical studies of a case found in Spain. Hum. Genet. 1978;42(2):171-180.

Waardenberg A.J., Basset S.D., Bouveret R., Harvey R.P. CompGO: an R package for comparing and visualizing Gene Ontology enrichment differences between DNA binding experiments. BMC Bioinformatics. 2015;16:275. DOI 10.1186/s12859-015-0701-2.

Watanabe M., Zingg B.C., Mohrenweiser H.W. Molecular analysis of a series of alleles in humans with reduced activity at the triosephosphate isomerase locus. Am. J. Hum. Genet. 1996;58(2):308-316.

Wu J., Wu M., Li L., Liu Z., Zeng W., Jiang R. dbWGFP: a database and web server of human whole-genome single nucleotide variants and their functional predictions. Database (Oxford). 2016;2016:baw024. DOI 10.1093/database/baw024.

Zerbino D.R., Wilder S.P., Johnson N., Juettemann T., Flicek P.R. The ensembl regulatory build. Genome Biol. 2015;16:56. DOI 10.1186/s13059-015-0621-5.

Zhang J.Q., Xia M., Shen Y.Q., Xu L.H., Yang J., Miao F.Q., Xie W. Research on the mechanisms and the function of abnormal HLA class I expression in hepatocellular carcinoma cell lines. Tissue Antigens. 2007;69(5):415. DOI 10.1111/j.1399-0039.2007.00836.x.

Zhang Y., Proenca R., Maffei M., Barone M., Leopold L., Friedman J.M. Positional cloning of the mouse obese gene and its human homologue. Nature. 1994;372(6505):425-432. DOI 10.1038/372425a0.

Zuber-Jerger I., Kullmann F., Schneidewind A., Scholmerich J. Diagnosis and treatment of a patient with gallstone ileus. Nat. Clin. Pract. Gastroenterol. Hepatol. 2005;2(7):331-335. DOI 10.1038/ncpgasthep0211.

ORCID ID

M.P. Ponomarenko orcid.org/0003-1663-318X

Благодарности. Текст статьи написан МПП при поддержке Минобрнауки РФ по Программе повышения конкурентоспособности российских университетов среди мировых научно-образовательных центров (Проект 5-100). Модификация Web-сервиса (ПМП, ДАР, ИВЧ и ЕАО) и анализ данных (ЕБШ, ИАД и ЛКС) проведены при поддержке бюджетных проектов № 0324-2019-0040 и № 0324-2019-0042 соответственно. Идея и общая координация работ (НАК и ВАК) – при поддержке Комплексной программы фундаментальных исследований СО РАН (проект № 0324-2018-0021).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 17.09.2017. После доработки 14.05.2019. Принята к публикации 29.08.2019.

ЭКОЛОГИЧЕСКАЯ ГЕНЕТИКА
Оригинальное исследование / Original article
УДК 577.214.337:632.72

Вавиловский журнал генетики и селекции. 2019;23(8):1059-1066
DOI 10.18699/VJ19.581

Diversity of *mariner*-like elements in Orthoptera

K. Ustyantsev[1] ✉, M. Biryukov[1], I. Sukhikh[1], N.V. Shatskaya[1], V. Fet[2], A. Blinov[1, 3], I. Konopatskaia[1]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Marshall University, Department of Biological Sciences, Huntington, USA
[3] Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia
✉ e-mail: ustyantsev@bionet.nsc.ru

Mariner-like elements (MLEs) are among the most widespread DNA transposable elements in eukaryotes. Insects were the first organisms in which MLEs were identified, however the diversity of MLEs in the insect order Orthoptera has not yet been addressed. In the present study, we explore the diversity of MLEs elements in 16 species of Orthoptera belonging to three infraorders, Acridoidea (Caelifera), Grylloidea (Ensifera), and Tettigoniidea (Ensifera) by combining data mined from computational analysis of sequenced degenerative PCR MLE amplicons and available Orthoptera genomic scaffolds. In total, 75 MLE lineages (Ortmar) were identified in all the studied genomes. Automatic phylogeny-based classification suggested that the current known variability of MLEs can be assigned to seven statistically well-supported phylogenetic clusters (I–VII), and the identified Orthoptera lineages were distributed among all of them. The majority of the lineages (36 out of 75) belong to cluster I; 20 belong to cluster VI; and seven, six, four, one and one lineages belong to clusters II, IV, VII, III, and V, respectively. Two of the clusters (II and IV) were composed of a single Orthoptera MLE lineage each (Ortmar37 and Ortmar45, respectively) which were distributed in the vast majority of the studied Orthoptera genomes. Finally, for 16 Orthoptera MLE lineages, horizontal transfer from the distantly related taxa belonging to other insect orders may have occurred. We believe that our study can serve as a basis for future researches on the diversity, distribution, and evolution of MLEs in species of other taxa that are still lacking the sequenced genomes.
Key words: *mariner*-like elements; transposable elements; Orthoptera; Insects; horizontal transfer.

For citation: Ustyantsev K., Biryukov M., Sukhikh I., Shatskaya N.V., Fet V., Blinov A., Konopatskaia I. Diversity of *mariner*-like elements in Orthoptera. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8): 1059-1066. DOI 10.18699/VJ19.581

Разнообразие *mariner*-подобных элементов Orthoptera

К.В. Устьянцев[1] ✉, М.Ю. Бирюков[1], И.С. Сухих[1], Н.В. Шацкая[1], В. Фет[2], А.Г. Блинов[1, 3], И.Д. Конопацкая[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Университет Маршалл, Отделение биологических наук, Хантингтон, США
[3] Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: ustyantsev@bionet.nsc.ru

Mariner-подобные элементы (МПЭ) – одни из самых распространенных мобильных элементов эукариот. Насекомые – первые организмы, в которых были обнаружены МПЭ, однако до сих пор еще не изучено разнообразие МПЭ насекомых отряда Orthoptera. В настоящей работе, совмещая результаты биоинформационного анализа секвенированных вырожденных ПЦР-ампликонов МПЭ и доступных геномных последовательностей Orthoptera, мы исследуем разнообразие МПЭ у 16 видов Orthoptera, принадлежащих к трем инфраотрядам: Acridoidea (Caelifera), Grylloidea (Ensifera) и Tettigoniidea (Ensifera). Всего среди всех изученных геномов выявлено 75 генетических линий МПЭ (Ortmar). Автоматизированная филогенетически опосредованная классификация подразделила все известное в настоящее время разнообразие МПЭ на семь кластеров со значимой статистической поддержкой (I–VII), в каждом из которых была найдена одна из генетических линий МПЭ Orthoptera. Большинство генетических линий (36 из 75) принадлежит к кластеру I; 20 линий относятся к кластеру VI; 6, 7, 4 и по 1 линии принадлежат к кластерам II, IV, VII, III и V соответственно. Два кластера, II и IV, были представлены единственными генетическими линиями МПЭ Orthoptera, Ortmar37 и Ortmar45, соответственно, распространенными среди большинства исследованных геномов Orthoptera. Мы также показали, что 16 генетических линий МПЭ Orthoptera могли участвовать в горизонтальном переносе с эволюционно удаленными таксонами насекомых из других отрядов.
Ключевые слова: *mariner*-подобные элементы; мобильные элементы; Orthoptera; насекомые; горизонтальный перенос.

Introduction

DNA transposons were the first mobile genetic elements discovered in eukaryotes (McClintock, 1950). Since that time, different groups of these "selfish" elements were identified throughout the whole eukaryotic tree of life (Robertson, 2002; Feschotte, Pritham, 2007; Wicker et al., 2007). The most abundant and widespread group of DNA transposons is a *Tc1/mariner* superfamily (Plasterk et al., 1999; Tellier et al., 2015) named after its two best-studied elements, *Tc1* and *mariner* from a nematode *Caenorhabditis elegans* (Emmons

© Ustyantsev K., Biryukov M., Sukhikh I., Shatskaya N.V., Fet V., Blinov A., Konopatskaia I., 2019

et al., 1983) and a fruit fly *Drosophila mauritiana* (Jacobson et al., 1986), respectively.

Mariner-like elements (MLEs) form the *mariner* family within the *Tc1/mariner* superfamily, MLEs are relatively small in size (about 1300 bp) characterized by the presence of 28–30 bp terminal inverted repeats (TIRs), an intronless gene encoding transposase, and a typical TA dinucleotide pattern of target site duplication (TSD) (Robertson, 1993; Robertson, MacLeod, 1993; Leroy et al., 2003). MLE transposases contain a common to *Tc1/mariner* N-terminal DNA-binding domain of helix-turn-helix type and a C-terminal catalytic domain with characteristic DD(34)D motif (Brillet et al., 2007). Former studies have described nine well-defined canonical subfamilies based on the similarities of the transposase catalytic core amino acid sequences (Robertson, 1997, 2002; Robertson, MacLeod, 1993; Green, Frommer, 2001): *vertumnana*, *cecropia*, *mauritiana*, *irritans*, *mellifera*, *capitata*, *elegans*, *briggsae* and *lineata*. This list was later supplemented by the more recently identified *marmoratus* and *drosophila* subfamilies (Bui et al., 2007; Wallau et al., 2014).

Probably, the most exciting feature of *Tc1/mariner* elements and MLEs in particular is their ubiquity. There are numerous reports suggesting horizontal transfers of MLEs between highly diverged taxa (Yoshiyama et al., 2001; Lampe et al., 2003; Silva et al., 2004, 2005; Oliveira et al., 2012; Dupeyron et al., 2014). Such a tendency to horizontal transfer is explained by (1) an evolutionary necessity for a long-term persistence of a DNA transposon, which otherwise will be eventually silenced by spontaneous mutations and cell's internal defense mechanisms (Lohe et al., 1995; Schaack et al., 2010), and (2) the simplicity of MLEs structure and life cycle (Plasterk, van Luenen, 2002). Presence of cis- or trans-encoded transposase and intact TIRs are the only features required for the transposition of MLEs (Plasterk et al., 1999). In addition, the MLE transposase was demonstrated to be fully active *in vitro* (Lampe et al., 1996; Tosi, Beverley, 2000). Such a low dependence on host-encoded and other biological factors most probably underlies success of MLEs as inter-taxa colonizers and their use as transgenic vectors (Delaurière et al., 2009).

Insects were the first organisms in which MLEs were identified and for which the horizontal transfer of MLEs was first proposed (Jacobson et al., 1986; Maruyama, Hartl, 1991; Robertson, 1993; Robertson, MacLeod, 1993). The insect order Orthoptera comprises two suborders: Caelifera (grasshoppers and locusts) and Ensifera (katydids and crickets). A computational analysis of the diversity of *Tc1/mariner* elements was carried out only for the genome of the migratory locust *Locusta migratoria* (Caelifera, Acrididae, Oedipodinae), one of the largest sequenced animal genome up to date (5.7 Gb) (Wang et al., 2014). However, this study provided no classification of the identified elements below the superfamily level (Bao et al., 2015). Only a few partial sequences belonging to MLEs were experimentally identified in two other grasshopper species (both in Caelifera, Acrididae): *Eyprepocnemis plorans* (Eyprepocnemedinae) (Montiel et al., 2012) and *Trimerotropis pallidipennis* (Oedipodinae) (Tevy et al., 2007). In addition, a recently published draft genome of the Hawaiian cricket *Laupala kohalensis* (Ensifera, Grylloidea, Trigonidiidae) is now also available for the analysis (Blankers et al., 2017). In this work, we explore the diversity of MLEs in Orthoptera from three infraorders, Acridoidea (Caelifera), Grylloidea (Ensifera), and Tettigoniidea (Ensifera). We identify new lineages of MLEs and take a first look into their horizontal transfer dynamics among insects.

Materials and methods

Orthoptera specimens. Specimens of 14 species of Orthoptera were provided by Dr. A.G. Bugrov, Institute of Systematics and Ecology of Animals, the Siberian Branch of Russian Academy of Sciences, Novosibirsk, Russia. We used 10 specimens of grasshoppers (Acridoidea), two of crickets (Grylloidea), and two of katydids (Tettigonioidea). The taxonomic position of each of the studied specimens is indicated in Table 1.

DNA extraction, PCR amplification and sequencing. Total DNA was extracted from muscle or brain tissue of the ethanol-preserved insect specimens using DNAeasy Blood and Tissue kit (Qiagen, Germany). We used canonical degenerate PCR primers specific for the conserved part of the MLE transposase gene catalytic domain to amplify the corresponding sequences for each of the DNA specimen. The primer sequences (MAR 124F and MAR 276R) and the cycling conditions are from the previous study (Robertson, MacLeod, 1993). PCR products were visualized and separated from unincorporated primers on 1.2 % agarose gel stained with EtBr. The resulting bands were cut out and then isolated from the gel for each of the studied Orthoptera specimens using QIAquick Gel Extraction Kit (Qiagen, Germany).

Library preparation, sequencing and assembly. Approximately 100 ng of each of the isolated 14 PCR amplicon mixes were used for barcode library preparation with Ion Xpress Plus Fragment Library Kit (Thermo Fisher Scientific) following the manufacturer's protocol. Fragmentation time was 6 min and 8 PCR amplification cycles were done. Purification of the libraries was done with AMPure XP kit (Beckman Coulter). A DNA fraction of 280–350 bp was extracted from the libraries using LabChip XT (Caliper Life Sciences) on a DNA500 chip with subsequent purification with AMPure XP kit. The quality and molarity of the obtained libraries was evaluated with DNA High Sensivity kit on the Bionanalyzer BA2100 chip (Agilent). The final libraries were amplified in emulsion PCR with Ion PGM Hi-Q View OT2 Kit (Thermo Scientific) and then sequenced on the Ion Torrent PGM system with 316v2 chip. Raw sequence reads produced by the Ion Torrent sequencing were *de novo* assembled into contigs using MIRA v. 4.0.2 (Chevreux et al., 2004), specifying "iontor" as technology parameter and "est, denovo, accurate" as job parameters in MIRA manifest configuration files.

Availability of supporting data. The genomic sequences used in this work are available at NCBI GenBank (http://www.ncbi.nlm.nih.gov/genbank). The genomes of *Locusta migratoria* and *Laupala kohalensis* include 5,760 Mb covering 1,397,492 scaffolds and 1,595 Mb covering 148,784 scaffolds, respectively (Wang et al., 2014; Blankers et al., 2017). The genome sequences from other 259 insect species (actual on 11.07.2018) were retrieved from NCBI GenBank using the following Entrez searching query on the Assembly database: "Insecta"[Organism] NOT "Orthoptera"[Organism] AND (latest[filter] AND "scaffold level"[filter] AND "representative genome"[filter] AND all[filter] NOT anomalous[filter] filtering option.

Table 1. Studied Orthoptera species and the identified diversity of MLE lineages

Family	Subfamily	Species	MLE cluster**							Number of Ortmar lineages per species
			I (36)	II (1)	III (7)	IV (1)	V (6)	VI (20)	VII (4)	
Acrididae	Acridinae	*Acrida* sp.	20	1	1	1	3	5	0	31
	Eyprepocnemidinae	*Eyprepocnemis plorans*	16	1	2	1	1	6	0	27
	Gomphocerinae	*Notostaurus albicornis*	14	1	2	1	1	4	0	23
	Melanoplinae	*Sinopodisma punctata*	18	1	3	1	1	8	0	32
		Bryodemella tuberculata	18	1	2	1	3	6	0	31
	Oedipodinae	*Locusta migratoria**	19	1	1	1	3	11	0	36
		Oedipoda fedtschenki	21	1	2	1	3	7	0	35
	Oxyinae	*Oxya japonica*	13	1	3	0	3	4	0	24
	Pezotettiginae	*Pezotettix anatolica*	11	1	2	1	3	3	0	21
Pyrgomorphidae	Pyrgomorphinae	*Chrotogonus* sp.	6	0	1	1	0	1	0	9
Pamphagidae	Pamphaginae	*Paranocaracris karadagi*	19	1	2	0	1	6	0	29
Tettigoniidae	Tettigoniinae	*Gampsocleis sedakovii*	9	1	4	1	2	2	1	20
		Tettigonia viridissima	10	1	1	1	1	1	1	16
Gryllidae	Gryllinae	*Cardiodactylus novaeguineae*	11	0	6	1	4	2	3	27
	Trogonidiinae	*Gryllus campestris*	14	1	3	1	3	1	1	24
	Eneopterinae	*Laupala kohalensis**	13	0	2	0	1	12	3	31

Notes. Clusters correspond to those in Figure. Taxonomy is according to NCBI Taxonomy. See Supplementary 4 for more details.
* Species with available genomic scaffolds.
** For each cluster, (N) designate total number of different Ortmar MLE lineages found.

Detection of MLE sequences. We collected a set of 19 known full-length transposase protein sequences belonging to eight MLE subfamilies (Supplementary 1)[1] and used them as queries in a batch tBLASTn search performed on the genomic sequences of *L. migratoria* and *L. kohalensis*. The recovered BLAST hits (High Scoring Pairs) were reassembled into copies if they were following the previously defined criteria (Wallau et al., 2014). Two hits from the same scaffold are jointed together if the distance between their centers is less than 1000 bp and if they have the same orientation. Only copies longer than 400 bp were retained. In order to extract the full-length sequence of each copy, the region of each assembled copy was expanded to 1500 bp. The copies were next clustered together into lineages by iterative BLASTn, setting the similarity threshold to 75 % which was previously defined in (Bouallègue et al., 2017), retrieving a majority rule consensus sequence for each lineage during the clustering process. Singlet sequences and consensuses derived from less than 10 sequences were removed from further analysis. Finally, we attempted to identify TIRs at the ends of the lineage consensuses using EMBOSS EINVERTED tool (Rice et al., 2000). Each newly identified complete sequence of Orthoptera MLE lineage was added to the query set and the whole search process was repeated to achieve the higher sensitivity.

The sequences of the query set obtained after the MLE screening of *L. migratoria* and *L. kohalensis* genomes were trimmed to the size of the conserved region of the MLE transposase catalytic domain which is between the canonical MLE degenerative primers (Robertson, MacLeod, 1993). Using this set, we performed a batch tBLASTn search against the assembled PCR amplicon sequences of 14 other Orthoptera species. Contigs of expected length (between 450 and 550 bp) which gave positive BLAST hits with more than 90 % coverage to the query sequence were kept for further analysis. The retrieved contigs were clustered into lineages as previously described. To remove redundancy and to keep (when possible) only a full-length representative of a lineage for further analyses, we did a BLASTn search using the consensus sequences of obtained lineages as queries against the database of full-length Orthoptera MLE lineages identified in this work.

Finally, consensus sequences of the obtained MLE lineages, both full-length and partial, were used as queries in a batch BLASTn search against the sequences originally retrieved by tBLASTn search from all the studied Orthoptera species to estimate the prevalence and distribution of each lineage.

Classification and phylogenetic analysis. Corresponding full-length and partial protein sequences of MLE transposases were identified in lineage consensuses using BLASTx search against the original transposase tBLASTn query set. The obtained protein sequences were aligned together with full-length transposase proteins of known MLEs from the previous work (Bouallègue et al., 2017) using MAFFT v7.312 (Katoh, Standley, 2013). The resulted alignment was used to build a Bayesian phylogeny of MLEs using the MrBayes v3.2.5 (Ronquist et al., 2012). The best-fitting model LG+I+G with four gamma rate categories was selected using SMS tool (Lefort et al., 2017) based on Akaike information criterion (Akaike, 1998). Bayesian posterior probabilities were estimated by running ten Markov chain Monte Carlo chains for 2,500,000 generations in two parallel runs, sampling each

[1] Supplementary Materials 1–6 are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx23.pdf

250 generations, and discarding first 25 % of samples in the end. Automatic phylogeny-based classification of the sequences on the resulted tree was performed using Cluster Picker tool (Ragonnet-Cronin et al., 2013). The minimal statistical support was set to 90 % and various genetic distance thresholds from 40 to 60 % were tested.

Horizontal gene transfer detection. The full-length nucleotide consensus sequences of the identified Orthoptera lineages were used in MEGABLAST search against the genomic scaffolds of 259 insect species from different orders available in the NCBI. Cases of potential horizontal transfer between the studied Orthoptera genomes and other insect taxa were considered when MEGABLAST hit showed more than 90 % of identity covering more than 90 % of the query sequences as it was proposed by several authors (Filée et al., 2015; Wallau et al., 2016; Bouallègue et al., 2017).

Results and discussion

Screening for MLE sequences was based on a homology approach (tBLASTn) with a set of 19 known full-length amino acid transposase sequences used as queries (see Supplementary 1). First, we initialized the search in the available genomic scaffolds of *L. migratoria* and *L. kohalensis* in order to identify Orthoptera full-length MLE sequences. A total of 4290 MLE copies from *L. migratoria* clustered in 27 lineages and 992 MLE copies from *L. kohalensis* clustered in 24 lineages were identified. Here, one lineage corresponds to a group of at least 10 sequences which have more than 75 % identity to each other detected by BLASTn comparisons.

Partial sequences of the conserved region of the MLE transposase catalytic domain were identified among the *de novo* sequenced PCR amplicons obtained from 14 other Orthoptera species. The sequences were then clustered into lineages following the same rules as described above (see Table 1). For each lineage, obtained both from the full-length or partial sequences, a majority-rule consensus sequence was derived (Supplementary 2). A total of 75 lineages were identified in all the studied Orthoptera genomes. For 49 of the lineages there was a full-length representative either from *L. migratoria* or *L. kohalensis* genome (Supplementary 3). There were no lineages distributed in all the genomes studied. Sixteen lineages were found only in the Caelifera, and 26, only in the Ensifera genomes. Thirty-three lineages were present both in Caelifera and Ensifera. Sixteen lineages were identified in a single genome each (Supplementary 4).

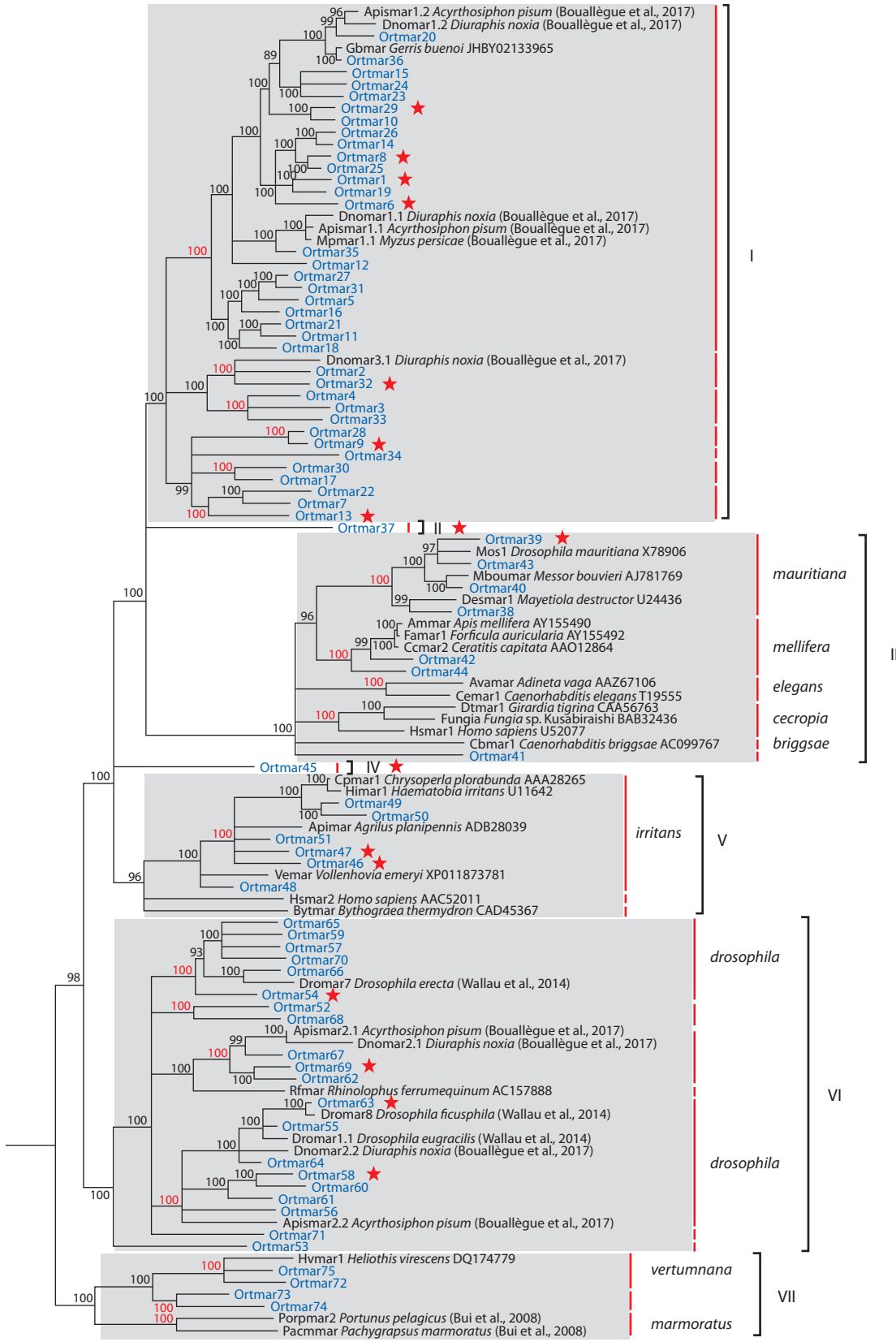
On average, 26 lineages were found for a studied orthopteran genome. The maximum (36 lineages) was identified in the genome of the migratory locust (*L. migratoria*) and the minimum (9 lineages) in the genome of an African grasshopper (*Chrotogonus* sp.) (see Table 1). It should be noted that the increase in the number of MLE lineages detected in *L. migratoria* (from 27 to 36) and *L. kohalensis* (from 24 to 31) genomes is due to the fact that these lineages are present as singlets or in less than 10 copies in the genomes and therefore were filtered out during the initial stringent mining process. Nevertheless, their minor presence is noted in Table 1 and Supplementary 4.

For the MLE lineage consensuses identified during the first stringent search in the *L. migratoria* and *L. kohalensis* genomes, TIRs as well as the characteristic to MLEs TA dinucleotide TSD have been identified. Among the 27 *L. migratoria* lineages TIRs were found in 24 lineages, and the TSD was found in 17. TIRs were found in all the 24 lineage consensuses from the *L. kohalensis* genome, 19 of which also had the TSD (see Supplementary 3).

Transposase amino acid sequences of the lineage consensuses were then used in phylogenetic analysis together with other known MLE transposases. To implement a phylogeny-based automatic classification of the identified lineages, we used the Cluster Picker tool (Ragonnet-Cronin et al., 2013) which assigns sequences into a cluster if (1) there is a substantial statistical support for their clustering on a tree and (2) a distance threshold between sequences in a cluster is satisfied. Testing different distance thresholds we found that at 43 % level all the "canonical" (historically first identified) MLE subfamilies (like *mauritiana*, *mellifera*, *cecropia*, etc.) still can be distinguished from each other, while increasing the threshold to 44 % already results in collapsing *mauritiana* and *mellifera* subfamilies into a single cluster. Using the 43 % threshold results in 28 clusters (including those represented by a single lineage), which can be supposedly of a subfamily level. However, these clusters at the same time are the parts of the higher-level ones (see Figure). For further convenience and the redundancy removal, we arbitrarily chose 55 % similarity threshold at which seven higher-level and statistically pronounced clusters (I–VII) are defined by Cluster Picker (see Figure, Table 1, and Supplementary 4). Four of these clusters (III, V, VI, and VII) include previously known subfamilies. Cluster III is comprised of *mauritiana*, *mellifera*, *cecropia*, *elegans* and *briggsae* subfamilies. Cluster V is represented solely by the *irritans* subfamily, even more supporting its status as such. Members of a recently defined *drosophila* subfamily (Wallau et al., 2014) were found within cluster VI. Cluster VII includes *vertumnana* and *marmoratus* subfamilies (Bui et al., 2007, 2008). Although no previously defined subfamilies were found grouped in cluster I, it includes elements recently identified in the insect order Hemiptera (Bouallègue et al., 2017) and assigned to *irritans* subfamily (which, however it is not supported by our phylogenetic analysis). Clusters II and IV are comprised of single Orthoptera lineage each (Ortmar37 and Ortmar45, respectively). However, these lineages were found in the vast majority (13 out of 16) of the studied genomes (see Figure, Table 1, and Supplementary 4).

The identified Orthoptera MLE lineages were distributed among all the seven clusters. The majority of the lineages (36 out of 75) belong to cluster I; 20 to cluster VI; and seven, six, four, one and one lineages, to the clusters II, IV, VII, II, and V, respectively (see Figure, and Supplementary 4). Interestingly, lineages of the cluster VII were identified only among the studied Ensifera genomes.

Horizontal transfer was suggested to be one of the main features explaining an extremely widespread occurrence of MLEs (Robertson, Lampe, 1995). To ensure the long-term survival, a transposable element has to be transferred horizontally or otherwise it would be eventually "fossilized" in the genome by vertical inactivation mechanisms (Lohe et al., 1995). Therefore, to fully understand the evolution of a transposable element lineage it is important to evaluate its horizontal transfer rate. For 16 lineages identified in the studied Orthoptera genomes, horizontal transfer between dis-



Phylogeny of MLE transposases reconstructed by a Bayesian analysis using the LG+I+G amino acid substitution model.

The tree was rooted with transposases from *rosa* and *LTIR Tc1/mariner* families (see Supplementary 6 for the names and accessions). Posterior probability is indicated for each node. Nodes with posterior probability below 90 % are collapsed and not shown. Names of the 75 Orthoptera MLE lineages identified in this study start with "Ortmar" and highlighted in blue. Clusters corresponding to original MLE subfamilies and predicted by Cluster Picker (Ragonnet-Cronin et al., 2013) with 43 % genetic distance threshold are indicated by bold red lines to the right of the tree, and their corresponding names (where applicable) are shown. Posterior probabilities of the nodes corresponding to subfamily-level clusters are highlighted in red. Higher-level clusters predicted by Cluster Picker with 55 % genetic distance threshold are indicated with black bold brackets and assigned with Roman numerals. Red stars indicate Orthoptera lineages which may have been horizontally transferred between Orthoptera species and species from other insect orders (see Table 2 for more details).

Table 2. Cases of potential horizontal transfer of MLE lineages identified in the studied Orthoptera genomes between other insects detected by MEGABLAST search

Orthoptera lineage*	Non-Orthoptera insect orders**							
	Lepidoptera (4/45)	Hymenoptera (15/67)	Hemiptera (3/23)	Diptera (3/100)	Coleoptera (2/10)	Odonata (1/2)	Phasmatodea (1/2)	Blattodea (2/4)
Ortmar1 (13)	–	–	–	–	–	–	1/320/94.8	–
Ortmar6 (10)	–	–	–	–	1/10/94.0	–	–	1/2/92.4
Ortmar8 (10)	–	–	–	–	1/43/92.0	–	–	–
Ortmar9 (10)	–	–	–	–	–	–	–	1/80/92.9
Ortmar13 (9)	–	–	–	–	–	1/1/90.2	–	–
Ortmar29 (3)	–	–	–	–	–	–	–	1/12/91.4
Ortmar32 (1)	–	1/1/90.6	–	–	–	–	–	1/78/95.3
Ortmar37 (13)	–	–	–	–	–	1/10/97.8	–	–
Ortmar39 (11)	–	9/351/93.2	–	–	–	–	–	–
Ortmar45 (13)	–	–	–	–	–	–	–	1/1/90.5
Ortmar46 (11)	–	–	1/5784/95.4	–	–	–	–	–
Ortmar47 (10)	–	–	1/1/90.6	–	–	–	–	–
Ortmar54 (9)	–	–	2/19/97.0	–	–	–	–	–
Ortmar58 (4)	–	3/8/96.5	–	–	–	–	1/51/96.8	–
Ortmar63 (2)	4/6/97.4	8/135/98.4	1/56/95.9	3/49/99.3	–	–	–	–
Ortmar69 (1)	–	–	1/50/95.6	–	–	–	–	–

Notes. See Supplementary 5 for more details.

* For each lineage, (N) – number of studied Orthoptera genomes in which the lineage was identified;

** for each non-Orthoptera order, (X/Y) – (number of genomes with detected potential horizontal transfers/number of genomes in the database); for the corresponding Orthoptera MLE lineage, values X/Y/Z designate number of genomes with MEGABLAST hit with more than 90 % identity/total number of MEGABLAST hits among the genomes/the maximum MEGABLAST identity (%) found.

tantly related species (divergence time over 250 Mya (http://www.timetree.org/home)) from other insect orders may have occurred (see Table 2, and Supplementary 5). Among these potential transfers, some draw an additional attention. For instance, high-identity (up to 99 %) MEGABLAST hits with the full-length sequence of the Ortmar63 lineage consensus from the *drosophila* subfamily (see Figure, cluster VI) were identified simultaneously in 16 genomes belonging to Diptera, Hemiptera, Lepidoptera, and Hymenoptera, while the lineage itself was identified in ten and two copies in two crickets only: *L. kohalensis* and *Cardiodactylus novaeguineae*, respectively (see Supplementary 4). A less prominent signal was found for the lineage Ortmar39 of the *mauritiana* subfamily identified in eleven Orthoptera genomes with 10 copies per genome on average (see Figure, cluster III). MEGABLAST hits to Ortmar39 (up to 92 % identity) were found for nine Hymenoptera genomes with the number of matches ranging from 1 to 257 (see Table 2, Supplementary 5). Two substantially diverged lineages from the cluster I (Ortmar8 and Ortmar9, each was found in ten studied Orthoptera genomes) gave 80 and 78 high scoring hits with a single genome of a termite *Cryptotermes secundus*. A substantial amount of high-identity hits with a genome of a stick insect *Clitarchus hookeri* (Phasmatodea) were found for the lineages Ortmar1 (320 hits up to 95 %, cluster I, distributed in 13 Orthoptera genomes) and Ortmar58 (51 hits up to 97 %, cluster VI, distributed in the 4 Orthoptera genomes). Another example is even more extraordinary: 5784 high-identity (up to 95 %) hits with the sequence of the Ortmar46 lineage of the *irritans* subfamily, widespread within the Orthoptera (see Figure, cluster V, Supplementary 4 and S5), were found in a single genome of the blood-sucking bug *Rhodnius prolixus* (Hemiptera) (see Table 2, and Supplementary 5).

Interestingly, previous studies showed that the *R. prolixus* genome is extremely saturated with members of a single lineage of the *irritans* subfamily, suggesting a relatively recent transposition burst with the estimated expansion occurred around the time of speciation of *R. prolixus*, approximately 1.4 Mya (Fernández-Medina et al., 2016). With Orthoptera species as a potential source, it can be hypothesized that the horizontal transfer of the Ortmar46 lineage could be a cause of the *irritans* subfamily expansion in the *R. prolixus* genome. Although in all the identified potential horizontal transfer cases the exact mechanism(s) and direction of the transfer remain elusive, these findings can later help in deciphering of each particular cases of the process in insects, as suggested by some authors (Silva et al., 2004; Loreto et al., 2008). Overall, the horizontal transfer cases identified here are in line with the recent report on massive horizontal transfer of transposable elements between insects which was based on more sophisticated genome to genome comparisons (Peccoud et al., 2017).

Conclusion

In this study, we used a non-standard approach to investigate the variability of MLEs. First of all, we took advantage of modern high-throughput sequencing technologies to avoid the bias of non-exhausting sampling which would occur if

the standard molecular cloning of the PCR amplicons with the subsequent Sanger sequencing was used. Here, the Ion Torrent PGM was a platform of choice, however, any of the currently available next-generation sequencing technologies can be utilized. The robustness of the obtained results was additionally controlled by the results of computational analysis of the two genome sequences from distantly related taxa of Orthoptera, *L. migratoria* (suborder Caelifera) and *L. kohalensis* (suborder Ensifera). Secondly, the automatic phylogeny-based classification of MLEs with Cluster Picker (Ragonnet-Cronin et al., 2013) allowed the definition of clusters on a rational, researcher-independent basis. In all previous studies on MLEs, it was a choice of the researchers whether to include newly identified MLE lineages to previously defined clusters, or designate new ones. Using this approach, we were able to identify and classify MLE lineages in the diverse 16 species of Orthoptera (see Table 1 and Figure), for which so far the genomic sequences were available only for the two species. We believe that our study can serve as a basis for future researches on diversity, distribution, and evolution of MLEs in species of other taxa that are still lacking the sequenced genomes.

References

Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. New York: Springer, 1998;199-213.

Bao W., Kojima K.K., Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA. 2015;6(1):11. DOI 10.1186/s13100-015-0041-9.

Blankers T., Oh K.P., Bombarely A., Shaw K.L. The genomic architecture of a rapid island radiation: mapping chromosomal rearrangements and recombination rate variation in *Laupala*. bioRxiv. 2017; 160952. DOI 10.1101/160952.

Bouallègue M., Filée J., Kharrat I., Mezghani-Khemakhem M., Rouault J.D., Makni M., Capy P. Diversity and evolution of *mariner*-like elements in aphid genomes. BMC Genomics. 2017;18(1):1-12. DOI 10.1186/s12864-017-3856-6.

Brillet B., Bigot Y., Augé-Gouillou C. Assembly of the *Tc1* and *mariner* transposition initiation complexes depends on the origins of their transposase DNA binding domains. Genetica. 2007;130(2):105-120. DOI 10.1007/s10709-006-0025-2.

Bui Q.T., Casse N., Leignel V., Nicolas V., Chénais B. Widespread occurrence of *mariner* transposons in coastal crabs. Mol. Phylogenet. Evol. 2008;47(3):1181-1189. DOI 10.1016/j.ympev.2008.03.029.

Bui Q.T., Delaurière L., Casse N., Nicolas V., Laulier M., Chénais B. Molecular characterization and phylogenetic position of a new *mariner*-like element in the coastal crab, *Pachygrapsus marmoratus*. Gene. 2007;396(2):248-256. DOI 10.1016/j.gene.2007.03.004.

Chevreux B., Pfisterer T., Drescher B., Driesel A.J., Müller W.E.G., Wetter T., Suhai S. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res. 2004;14(6):1147-1159. DOI 10.1101/gr.1917404.

Delaurière L., Chénais B., Hardivillier Y., Gauvry L., Casse N. *Mariner* transposons as genetic tools in vertebrate cells. Genetica. 2009; 137(1):9-17. DOI 10.1007/s10709-009-9370-2.

Dupeyron M., Leclercq S., Cerveau N., Bouchon D., Gilbert C. Horizontal transfer of transposons between and within crustaceans and insects. Mob. DNA. 2014;5(1):4. DOI 10.1186/1759-8753-5-4.

Emmons S.W., Yesner L., Ruan K.S., Katzenberg D. Evidence for a transposon in *Caenorhabditis elegans*. Cell. 1983;32(1):55-65.

Fernández-Medina R.D., Granzotto A., Ribeiro J.M., Carareto C.M.A. Transposition burst of *mariner*-like elements in the sequenced genome of *Rhodnius prolixus*. Insect Biochem. Mol. Biol. 2016;69: 14-24. DOI 10.1016/j.ibmb.2015.09.003.

Feschotte C., Pritham E.J. DNA transposons and the evolution of eukaryotic genomes. Annu. Rev. Genet. 2007;41(1):331-368. DOI 10.1146/annurev.genet.40.110405.090448.

Filée J., Rouault J.-D., Harry M., Hua-Van A. Mariner transposons are sailing in the genome of the blood-sucking bug *Rhodnius prolixus*. BMC Genomics. 2015;16:1061. DOI 10.1186/s12864-015-2060-9.

Green C.L., Frommer M. The genome of the Queensland fruit fly *Bactrocera tryoni* contains multiple representatives of the *mariner* family of transposable elements. Insect Mol. Biol. 2001;10(4): 371-386.

Jacobson J.W., Medhora M.M., Hartl D.L. Molecular structure of a somatically unstable transposable element in *Drosophila*. Proc. Natl. Acad. Sci. USA. 1986;83(22):8684-8688.

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

Lampe D.J., Churchill M.E., Robertson H.M. A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. EMBO J. 1996; 15(19):5470-5479.

Lampe D.J., Witherspoon D.J., Soto-Adames F.N., Robertson H.M. Recent horizontal transfer of mellifera subfamily *mariner* transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. Mol. Biol. Evol. 2003;20(4):554-562. DOI 10.1093/molbev/msg069.

Lefort V., Longueville J.-E., Gascuel O. SMS: smart model selection in PhyML. Mol. Biol. Evol. 2017;34(9):2422-2424. DOI 10.1093/molbev/msx149.

Leroy H., Castagnone-Sereno P., Renault S., Augé-Gouillou C., Bigot Y., Abad P. Characterization of Mcmar1, a *mariner*-like element with large inverted terminal repeats (ITRs) from the phytoparasitic nematode *Meloidogyne chitwoodi*. Gene. 2003;304:35-41.

Lohe A.R., Moriyama E.N., Lidholm D.A., Hartl D.L. Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. Mol. Biol. Evol. 1995;12(1):62-72. DOI 10.1093/oxfordjou rnals.molbev.a040191.

Loreto E.L.S., Carareto C.M.A., Capy P. Revisiting horizontal transfer of transposable elements in *Drosophila*. Heredity (Edinb.). 2008; 100(6):545-554. DOI 10.1038/sj.hdy.6801094.

Maruyama K., Hartl D.L. Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus*. J. Mol. Evol. 1991;33(6):514-524.

McClintock B. The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. USA. 1950;36(6):344-355. DOI 10.1073/pnas.36.6.344.

Montiel E.E., Cabrero J., Camacho J.P.M., López-León M.D. *Gypsy*, *RTE* and *Mariner* transposable elements populate *Eyprepocnemis plorans* genome. Genetica. 2012;140(7-9):365-374. DOI 10.1007/s10709-012-9686-1.

Oliveira S.G., Bao W., Martins C., Jurka J. Horizontal transfers of *Mariner* transposons between mammals and insects. Mob. DNA. 2012; 3(1):14. DOI 10.1186/1759-8753-3-14.

Peccoud J., Loiseau V., Cordaux R., Gilbert C. Massive horizontal transfer of transposable elements in insects. Proc. Natl. Acad. Sci. USA. 2017;114(18):4721-4726. DOI 10.1073/pnas.1621178114.

Plasterk R.H., Izsvák Z., Ivics Z. Resident aliens: the *Tc1/mariner* superfamily of transposable elements. Trends Genet. 1999;15(8):326-332.

Plasterk R.H.A., van Luenen H.G.A.M. The *Tc1*/*Mariner* family of transposable elements. In: Mobile DNA II. Washington, DC: American Society of Microbiology, 2002;519-532.

Ragonnet-Cronin M., Hodcroft E., Hué S., Fearnhill E., Delpech V., Brown A.J., Lycett S. Automated analysis of phylogenetic clusters. BMC Bioinformatics. 2013;14(1):317. DOI 10.1186/1471-2105-14-317.

Rice P., Longden I., Bleasby A. EMBOSS: the European molecular biology open software suite. Trends Genet. 2000;16(6):276-277.

Robertson H.M. The *mariner* transposable element is widespread in insects. Nature. 1993;362(6417):241-245. DOI 10.1038/362241a0.

Robertson H.M. Multiple *mariner* transposons in flatworms and hydras are related to those of insects. J. Hered. 1997;88(3):195-201. DOI 10.1093/oxfordjournals.jhered.a023088.

Robertson H.M. Evolution of DNA transposons in eukaryotes. In: Mobile DNA II. Washington, DC: American Society of Microbiology, 2002;1093-1110.

Robertson H.M., Lampe D.J. Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. Mol. Biol. Evol. 1995;12(5):850-862.

Robertson H.M., MacLeod E.G. Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. Insect Mol. Biol. 1993;2(3):125-139.

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

Schaack S., Gilbert C., Feschotte C. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. Trends Ecol. Evol. 2010;25(9):537-546. DOI 10.1016/j.tree.2010.06.001.

Silva J.C., Bastida F., Bidwell S.L., Johnson P.J., Carlton J.M. A potentially functional *mariner* transposable element in the protist *Trichomonas vaginalis*. Mol. Biol. Evol. 2005;22(1):126-134. DOI 10.1093/molbev/msh260.

Silva J.C., Loreto E.L., Clark J.B. Factors that affect the horizontal transfer of transposable elements. Curr. Issues Mol. Biol. 2004;6(1):57-71.

Tellier M., Bouuaert C.C., Chalmers R. *Mariner* and the *ITm* superfamily of transposons. Microbiol. Spectr. 2015;3(2):MDNA3-0033-2014. DOI 10.1128/microbiolspec.MDNA3-0033-2014.

Tevy F., Guzman N., Gonzalez G., Lia V., Poggio L., Confalonieri V.A. Mobile elements and inverted rearrangements in *Trimerotropis pallidipennis* (Orthoptera: Acrididae). Caryologia. 2007;60(3):212-221. DOI 10.1080/00087114.2007.10797939.

Tosi L.R., Beverley S.M. Cis and trans factors affecting *Mos1 mariner* evolution and transposition *in vitro*, and its potential for functional genomics. Nucleic Acids Res. 2000;28(3):784-790.

Wallau G.L., Capy P., Loreto E., Hua-Van A. Genomic landscape and evolutionary dynamics of *mariner* transposable elements within the *Drosophila* genus. BMC Genomics. 2014;15(1):1-19. DOI 10.1186/1471-2164-15-727.

Wallau G.L., Capy P., Loreto E., Le Rouzic A., Hua-Van A. VHICA, a new method to discriminate between vertical and horizontal transposon transfer: application to the *Mariner* family within *Drosophila*. Mol. Biol. Evol. 2016;33(4):1094-1109. DOI 10.1093/molbev/msv341.

Wang X., Fang X., Yang P., Jiang X., Jiang F., Zhao D., Li B., Cui F., Wei J., Ma C., Wang Y., He J., Luo Y., Wang Z., Guo X., Guo W., Wang X., Zhang Y., Yang M., Hao S., Chen B., Ma Z., Yu D., Xiong Z., Zhu Y., Fan D., Han L., Wang B., Chen Y., Wang J., Yang L., Zhao W., Feng Y., Chen G., Lian J., Li Q., Huang Z., Yao X., Lv N., Zhang G., Li Y., Wang J., Wang J., Zhu B., Kang L. The locust genome provides insight into swarm formation and long-distance flight. Nat. Commun. 2014;5:2957. DOI 10.1038/ncomms3957.

Wicker T., Sabot F., Hua-Van A., Bennetzen J.L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M., Panaud O., Paux E., SanMiguel P., Schulman A.H. A unified classification system for eukaryotic transposable elements. Nat. Rev. Genet. 2007;8(12):973-982. DOI 10.1038/nrg2165.

Yoshiyama M., Tu Z., Kainoh Y., Honda H., Shono T., Kimura K. Possible horizontal transfer of a transposable element from host to parasitoid. Mol. Biol. Evol. 2001;18(10):1952-1958. DOI 10.1093/oxfordjournals.molbev.a003735.

ORCID ID

K. Ustyantsev orcid.org/0000-0003-4346-3868
I. Sukhikh orcid.org/0000-0002-3548-4354

Acknowledgements. The work of K. Ustyantsev and A. Blinov was funded by the Russian State Budget (Project No. 0324-2019-0040). The work of M. Biryukov was funded by the Russian State Budget (Project No. 0259-2019-0010). We thank Dr. A.G. Bugrov for providing the insect specimens and the Center of Genome Investigation of the Institute of Cytology and Genetics SB RAS for the sequencing assistance.

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

Received May 28, 2019. Revised June 25, 2019. Accepted June 28, 2019.

ЭКОЛОГИЧЕСКАЯ ГЕНЕТИКА
Оригинальное исследование / Original article
УДК 582.736:581.52

Вавиловский журнал генетики и селекции. 2019;23(8):1067-1075
DOI 10.18699/VJ19.582

Внутри- и межвидовая изменчивость *Mentha arvensis* L. и *M. canadensis* L.

М.В. Семёнова ✉, О.Л. Енина, О.В. Шелепова

Главный ботанический сад им. Н.В. Цицина Российской академии наук, Москва, Россия
✉ e-mail: lab-physiol@mail.ru

При идентификации представителей рода *Mentha* L. часто возникают сложности из-за значительного внутривидового полиморфизма, активной межвидовой гибридизации, а также изменений в плоидности растений. В настоящей работе предпринята попытка комплексного подхода к изучению различных параметров растений двух трудно разделяемых видов с перекрывающимися признаками: *Mentha arvensis* L. и *M. canadensis* L. Исследованы представители восьми географически разобщенных популяций *Mentha*, собранные в разных регионах: европейская часть, Хакасия и Дальний Восток России, Западная Украина и Индокитай. Изучены диагностические для этих видов морфологические признаки, состав компонентов эфирного масла и проведен фрагментный анализ ДНК с использованием ISSR-маркеров. Полученные данные обработаны статистически с помощью кластерного анализа, метода главных компонент и метода главных координат. По результатам анализа количественных характеристик чашечки и листьев четко разделяются европейская и азиатская группы образцов, а в определении принадлежности дальневосточных растений к той или иной группе есть расхождения при использовании разных методов обработки данных и, соответственно, затруднено определение их таксономического положения с использованием только морфологических признаков. По составу эфирного масла и ISSR-фрагментов выделена группа наиболее генетически, морфологически и фитохимически близких растений, в которую вошли представители популяций Московской, Владимирской, Калужской областей, Республики Коми и Хакасии. Все они относятся к виду *M. arvensis*. Растения, собранные в природной флоре Дальнего Востока, по компонентному составу эфирного масла и составу ISSR-фрагментов имели большее сходство с *M. arvensis* из европейской части, чем с коллекционными образцами из флоры Индокитая, которые, согласно полученным данным, относятся к *M. canadensis*. Комплексное изучение морфологических признаков растений, состава эфирного масла и фрагментного ISSR-анализа ДНК позволяет уточнить видовую принадлежность, а также оценить полиморфизм и степень родства между популяциями. Выявлена корреляция между данными молекулярно-генетического анализа и составом эфирного масла для растений *M. arvensis* и *M. canadensis* и, в некоторой степени, с морфологическими признаками растений.
Ключевые слова: *Mentha arvensis*; *Mentha canadensis*; полиморфизм; ДНК; ISSR; молекулярные маркеры; компонентный состав эфирного масла.

Для цитирования: Семёнова М.В., Енина О.Л., Шелепова О.В. Внутри- и межвидовая изменчивость *Mentha arvensis* L. и *M. canadensis* L. Вавиловский журнал генетики и селекции. 2019;23(8):1067-1075. DOI 10.18699/VJ19.582

Intra- and interspecific variability of *Mentha arvensis* L. and *M. canadensis* L.

M.V. Semenova ✉, O.L. Enina, O.V. Shelepova

Tsitsin Main Botanical Garden, RAS, Moscow, Russia
✉ e-mail: lab-physiol@mail.ru

The identification of plants of the genus *Mentha* is often difficult due to significant intraspecific polymorphism, intense interspecific hybridization, and ploidy changes. An attempt was made to apply an integrated approach to the study of different parameters of two species: *Mentha arvensis* L. and *M. canadensis* L. Eight geographically dispersed populations of *Mentha* in different regions (European Russia, Khakassia and Far East, Western Ukraine, and Indochina) were studied. Diagnostic morphological characters and compositions of essential oil components were examined, and DNA was analyzed with ISSR markers. The data obtained were statistically processed by cluster, principal component, and principal coordinate analyses. The European and Asian groups of samples were clearly distinguished by the analysis of quantitative parameters of the calyx and leaves, but different methods of data processing produced different results in determining the belonging of the Far Eastern plants to a particular group. Therefore, their taxonomic positions can hardly be determined on morphological grounds. According to the composition of essential oil and ISSR fragments, a group of the genetically, morphologically, and phytochemically closest plants was identified, which included representatives of the populations of the Moscow oblast, Vladimir oblast, Kaluga oblast, the Komi Republic, and Khakassia. All these plants belonged to *M. arvensis*. Plants collected in the natural flora of the Russian Far East showed a greater resemblance in essential oil composition and ISSR markers to the European group of *M. arvensis*

© Семёнова М.В., Енина О.Л., Шелепова О.В., 2019

than to plants from Indochina, which, according to the data obtained, belonged to *M. canadensis*. It was shown that a comprehensive study of plant morphological characters, the compositions of essential oil, and ISSR fragments allows one to clarify the species identity and to assess their polymorphism and the degree of kinship between populations. A certain correlation between the data of molecular analysis and the composition of essential oil and, to a lesser extent, their correlation with morphological characters of plants was revealed.
Key words: *Mentha arvensis*; *Mentha canadensis*; polymorphism; DNA; ISSR; molecular markers; composition of essential oil.

For citation: Semenova M.V., Enina O.L., Shelepova O.V. Intra- and interspecific variability of *Mentha arvensis* L. and *M. canadensis* L. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8):1067-1075. DOI 10.18699/VJ19.582 (in Russian)

Введение

Виды *Mentha* L. обладают значительным ресурсным потенциалом благодаря содержанию эфирного масла в надземной части растений. Мята используется в медицине в качестве спазмолитического, успокоительного, желудочного, желчегонного средства. Кроме того, она находит применение в пищевой и парфюмерной промышленности. В южных регионах Азии ее культивируют для получения технического эфирного масла, при специальной обработке которого получают ментол, входящий в состав многих важных лекарственных препаратов.

По данным различных источников, род *Mentha* L. включает от 18 видов и 11 гибридов (Tucker, Naczi, 2007) до 27 видов и 15 гибридов (WCSP, 2019). При этом исследователи не могут прийти к единому мнению относительно объемов и границ морфологически сходных между собой, но в то же время полиморфных видов. Отличительной особенностью рода является значительный полиморфизм внутри видов и одновременно сходство (перекрывание) ряда морфологических признаков между видами. В связи с этим в качестве дополнительного признака для характеристики вида целесообразно использовать фитохимические особенности растений, в частности состав эфирного масла. Компонентный состав эфирного масла у всех видов *Mentha* генетически детерминирован, и при оптимальных условиях синтез компонентов протекает до полного набора терпеновых соединений, характерных для вида. В неблагоприятных условиях произрастания синтез этих веществ останавливается на более ранних стадиях, вследствие чего в качественном составе эфирного масла появляются более «простые» соединения (Gouyon et al., 1986). Это приводит к трудностям в таксономическом разделении близких видов и в полной мере относится к двум исключительно полиморфным и сложно разделяемым видам *M. arvensis* L. и *M. canadensis* L.

Mentha arvensis произрастает в Европе и Азии, на территории России распространена в европейской части, Предкавказье и в Сибири (Губанов и др., 2004). Ареал *M. canadensis* в нашей стране включает Сибирь и Дальний Восток (Доронькин, 1997), в мире – Северную Америку и Восточную Азию (Tucker, Naczi, 2007).

Mentha arvensis L. (мята полевая) – широко распространенный и очень изменчивый по морфологическим признакам вид. Габитус растений значительно меняется в зависимости от условий произрастания. Так, в тенистых сырых местах формируются высокорослые растения с приподнимающимися или лежачими стеблями и крупными зелеными или светло-зелеными листьями и бледно-лиловыми, иногда почти белыми цветками. На более сухих участках растения становятся низкорослыми, с прямостоячим стеблем, сближенными междоузлиями и мелкими листьями, на которых появляется красно-фиолетовый антоциановый оттенок, цветки приобретают ярко-лиловую окраску.

Mentha canadensis L. (мята канадская) была описана К. Линнеем (Linne, 1753), это родственный *M. arvensis* и сложно идентифицируемый вид. От *M. arvensis* отличается более высоким неветвистым, сильно опушенным стеблем, вдвое более узкими и острыми глубокопильчатыми листьями, а также формой чашечки. Для разделения этих видов по морфологическим признакам используют различные критерии. В.М. Доронькин (1997) предложил разделять их по форме чашечки и строению зубцов: у *M. arvensis* чашечки колокольчатые, зубцы короткие, широкотреугольные, тогда как у *M. canadensis* чашечки трубчатые, зубцы длинные, тонко заостренные до почти шиловидных, реже короткие, но тоже заостренные. Другие исследователи в приводимом ими ключе разделяют виды по форме листьев и аромату, однако эти признаки не могут быть надежным критерием (Tucker, Naczi, 2007).

К числу морфологических признаков, вызывающих затруднения в процессе идентификации рассматриваемых двух видов, можно отнести такие изменчивые внутривидовые и в то же время перекрывающиеся признаки, как: тип ветвления и степень ветвистости, форма и размеры листа, опушенность побегов. Все эти признаки сильно варьируют в зависимости от условий произрастания (особенно у *M. arvensis*), но очень важны в определении вида.

В разделении и определении границ двух видов исследователи часто расходятся во мнении. Так, ряд авторов рассматривали *M. canadensis* не как самостоятельный вид, а как разновидность, форму или подвид *M. arvensis*. E. Holmes (1882) описал мяту, произрастающую в Японии, как *M. arvensis* f. *piperascens* Malinv. ex Holmes. J. Briquet (1894) выделял мяту из Северной Америки как *M. arvensis* var. *canadensis*, а мяту из Азии – как *M. arvensis* var. *haplocalyx*. H. Hara (1956) описывал ее как *M. arvensis* ssp. *piperascens* (Malinv. ex Holmes) H. Hara. В настоящее время *M. canadensis* выделяют в самостоятельный вид, хотя для его обозначения и сохранились многочисленные синонимы, наиболее полно представленные в World Checklist of Selected Plant Families (WCSP, 2019).

Возможно, *M. canadensis* является амфидиплоидом, возникшим в результате межвидовой гибридизации *M. arvensis* и *M. longifolia* (Tucker, Chambers, 2002). В литературе имеются сведения об изучении хлоропластных последовательностей ДНК, которые свидетельствуют также о том, что *M. arvensis* может быть родительским видом,

выступавшим в качестве материнского растения (источника женских гамет) для образовавшегося в процессе межвидовой гибридизации вида *M. canadensis* (Bunsawat et al., 2004). На основании комплексного анализа данных проточной цитометрии и ISSR-анализа было выдвинуто предположение, что *M. canadensis* – это аллополиплоид, возникший в результате цитомиксиса между представителями родительских видов с различной плоидностью (Jedrzejczyk, Rewers, 2018).

Целью настоящей работы было изучение внутри- и межвидовой изменчивости и дифференциации *M. arvensis* и *M. canadensis* из нескольких, в разной степени удаленных друг от друга популяций. Ранее нами рассмотрен полиморфизм в популяциях *M. arvensis* на основе морфологических и молекулярно-генетических исследований (Шелепова и др., 2016б), а также предпринята попытка дифференцировать два близких вида с использованием морфологического и фитохимического анализа (Шелепова и др., 2016а), однако результаты этих исследований не дали полной картины. Для изучения межвидовых и популяционных различий в роде *Mentha* в ряде работ успешно используется комплексная оценка образцов по морфологическим признакам, составу эфирного масла и ISSR-маркеров (Hua et al., 2011; Rodrigues et al., 2013; Шелепова и др., 2017). В связи с этим для выполнения поставленной задачи нами проведено комплексное исследование ряда морфологических признаков, а также компонентного состава эфирного масла и молекулярно-генетических данных с помощью ISSR-маркеров ДНК у растений двух видов мяты.

Материалы и методы

Отбор проб в популяциях. В исследование были включены растения из восьми популяций *M. arvensis* и *M. canadensis*, собранные в четырех значительно удаленных друг от друга регионах (табл. 1).

Для изучения морфологических признаков и проведения ISSR-анализа ДНК в природных популяциях были взяты 5–8 побегов с типичных растений, собранные через каждые 10–200 м (в зависимости от размера популяции), в культивируемых на экспериментальном участке образцах – побеги с 1–3 растений. В предварительном исследовании установлено, что выборки из природных популяций отличались большим разнообразием амплифицированных фрагментов и, следовательно, значительным разнообразием генотипов в пределах одной популяции, а растения из коллекции практически не различались между собой по составу ISSR-фрагментов и представляли собой один клон. В качестве внешнего стандарта использовали вид *Mentha spicata* L. Для анализа компонентного состава эфирного масла в популяциях отбирали от 1 до 5 проб растений одновременно со сбором гербарных образцов (в фазу массового цветения) с площадки размером 0.2 м^2. Гербарные образцы хранятся в лаборатории физиологии и иммунитета растений и в гербарии ГБС РАН (MHA).

Изучение морфологических признаков. Растения были определены до уровня вида. Большая часть образцов отнесена к *M. arvensis*, а растения из природной флоры Дальнего Востока и коллекционные образцы из флоры Индокитая предварительно определены как *M. canadensis*. Исследованы следующие количественные признаки: длина чашечки цветка, длина зубцов чашечки, ширина зубцов чашечки в основании, отношение длины зубца к ширине, отношение длины чашечки к длине зубцов, отношение длины трубки к длине зубцов; длина и максимальная ширина листа, расстояние от основания до максимальной ширины листа; количество секреторных желез на нижней и верхней поверхности листа на 1 см^2, отношение количества секреторных желез на нижней и верхней поверхности листа. Строение чашечки исследовали с помощью светового стереомикроскопа при увеличении 400×. Для сравнительного анализа растений по этому признаку были сделаны фотографии с помощью видеокамеры Lumenera Infinity 2 и обработаны в программе Infinity Analyses 5.0.2. Количество секреторных желез определяли по ранее опубликованной методике (Шелепова и др., 2012).

Методы изучения состава эфирного масла. Эфирное масло получали из измельченного воздушно-сухого материала методом гидродистилляции (Гинзберг, 1932). Для этого использовали среднюю пробу из надземной массы растений (смесь соцветий и листьев). Эфирное масло растений состоит из смеси природных соединений – монотерпенов, дитерпенов, гемитерпенов, сесквитерпенов и др., а также их окислов. Вопросы классификации, биосинтеза и детальной проработки каждого класса соединений подробно рассмотрены в работе (Племенков, 2007). Качественный состав масла определяли газохроматографически в ЦКП ФИЦ «Биотехнология» РАН (RFMEFI62114X0002) на газовом хроматографе Shimadzu GS 2010 с масс-детектором GCMS-QP 2010 по описанной нами ранее методике (Шелепова и др., 2017).

Выделение ДНК и проведение ПЦР. ДНК выделяли из сухих листьев, взятых с гербарных образцов, с применением CTAB-метода (Doyle J.J., Doyle J.L., 1987). Для исследования полиморфизма ДНК использовали ISSR (inter simple sequence repeat) маркеры. Праймеры для ПЦР-анализа были синтезированы и очищены в PAAG фирмой Syntol Ltd (Москва, Россия). После предварительной проверки отобраны восемь праймеров (табл. 2). Проведены две серии ПЦР для 46 и 42 проб, причем часть образцов была включена в схему постановки ПЦР повторно. ISSR–ПЦР и разделение ПЦР-продукта выполняли по опубликованной ранее методике (Шелепова и др., 2016б, 2017).

Анализ молекулярных данных и статистическая обработка. Профили полос полученных ISSR-фрагментов в агарозном геле сравнивали с помощью визуальной оценки и программы CrossChecker. Учитывались только яркие и четкие фрагменты, неясные полосы считались отсутствующими. Каждый амплифицированный фрагмент, который был визуализирован в виде полосы в электрофоретическом геле, рассматривался как счетный признак и использовался в виде двоичного кода для составления матрицы присутствия/отсутствия фрагментов. Затем полученные результаты анализировали в программе PAST с использованием кластерного анализа и метода главных компонент и главных координат (Hammer et al., 2001). Для оценки стабильности полученных дендрограмм был проведен бутстреп-анализ с 1000 реплик.

Таблица 1. Описание исследованных образцов *Mentha* и точек сбора проб

Название образца	Кол-во образцов	Местонахождение популяции	Геогр. координаты места сбора
M-1,2	2	Москва, территория Главного ботанического сада РАН. Сухое русло ручья на левом берегу Лихоборки	55°50′35.7″ N, 37°36′52.3″ E
M-3	1	Московская область, усадьба «Знаменское Садки», берег пруда над р. Битца	55°34′44.8″ N, 37°33′19.3″ E
M-4	1	Московская область, первая речная терраса р. Битца, 3 км ниже по течению от усадьбы «Знаменское Садки»	55°34′38.1″ N, 37°33′47.1″ E
M-5	1	Московская область, 1.5 км на юго-запад от дер. Марьино, Красногорский район. У кромки воды запруды на р. Синичка	55°51′07.6″ N, 37°18′35.6″ E
M-6	1	Московская область, Валуевский лесопарк, окрестности дер. Мешково	55°35′03.5″ N, 37°18′31.6″ E
M-7	1	Москва, природный заказник «Воробьевы горы»	55°42′33.4″ N, 37°33′16.6″ E
H	5	Хакасия, Алтайский район, правый высокий берег р. Абакан, около 1.5 км к западу от пос. Изыхские Копи, на галечниковой отмели, 5 растений	53°33′4.54″ N, 91°14′38.15″ E
F	5	Владимирская область, Ковровский район, пойма р. Клязьма, 5 растений	56°24′59.1″ N, 41°23′06.0″ E
K	2	Калужская область, Жуковский район, окраина дер. Окороково, 2 растения	55°03′57.3″ N, 36°42′55.5″ E
Komi	3	Коллекционный участок лаборатории физиологии и иммунитета растений ГБС РАН, получено из Сыктывкара, природная флора Коми, 3 растения	55°50′06.5″ N, 37°35′17.4″ E
Ind1-5	5	Коллекционный участок лаборатории физиологии и иммунитета растений ГБС РАН, получено из Вьетнама, природная флора Индокитая, 3 растения	55°50′06.5″ N, 37°35′17.4″ E
Mentha spicata	1	Коллекционный участок лаборатории физиологии и иммунитета растений ГБС РАН, получено из Никитского ботанического сада, Ялта	55°50′06.5″ N, 37°35′17.4″ E
U-1	1	Украина, Львов, в черте города, парк Погулянка, заболоченная низина рядом с источником под склоном букового леса	49°49′11.0″ N, 24°02′14.3″ E
U-2	1	Украина, Львов, парк Погулянка, край болота	49°49′22.9″ N, 24°03′52.4″ E
U-3,4	2	Украина, северо-западная окраина Львова, пос. Белогорща, торфяник	49°50′28.0″ N, 23°55′02.5″ E
U-5	1	Украина, Ивано-Франковская область, Галичский район, с. Тустань, пойма р. Гнилая Липа	49°08′00.0″ N, 24°45′20.4″ E
U-6	1	Украина, Ивано-Франковская область, Галичский район, между селами Водники и Дубовцы, старица р. Днестр	49°04′33.1″ N, 24°47′53.1″ E
DV-1-5,8	6	Приморский край, Хасанский район, бухта Витязь	42°35′36″ N, 131°11′13″ E
DV-6,7	2	Приморский край, Хасанский район, бухта Астафьева	42°37′7″ N, 131°11′54″ E
DV-9,10	2	Приморский край, Лазовский район, база «Заповедное»	42°50′27″ N, 133°41′56″ E

Таблица 2. Список ISSR-праймеров, использованных для постановки ПЦР

Праймер	Последовательность	Температура отжига праймеров, °C	Кол-во полиморфных ISSR-фрагментов
M2	$(AC)_8(C/T)G$	50.0	12
M3	$(GA)_8(C/T)C$	52.7	17
M4	$(AG)_8YC$	50.0	17
M7	$(CAG)_5$	52.7	16
M8	$(GTG)_5$	52.7	13
M12	$(CA)_6(A/G)(C/T)$	50.0	13
UBC 840	$(GA)_8AYT$	50.0	12
UBC855	$(AC)_8CYT$	50.0	17

Результаты

Анализ количественных морфологических признаков с помощью метода главных компонент показал, что исследуемые образцы разделяются на две основные группы (рис. 1, *а*). В одну из них вошли все образцы из европейской части, Республики Коми, Хакассии, Дальнего Востока России, а также из Украины и один образец из Индокитая (Ind2). Дальневосточные растения расположены близко друг к другу и являются частью общего смешанного облака, сформированного остальными образцами. В отдельное облако попали растения из Индокитая, которые были определены как *M. canadensis*. При обработке тех же данных методом кластерного анализа с использованием дистанции Говера (см. рис. 1, *б*) были получены два основных кластера с несколько различающимся составом образцов внутри кластеров. В первом кластере все растения Дальнего Востока и Индокитая, во втором – европейские и растения из Хакасии.

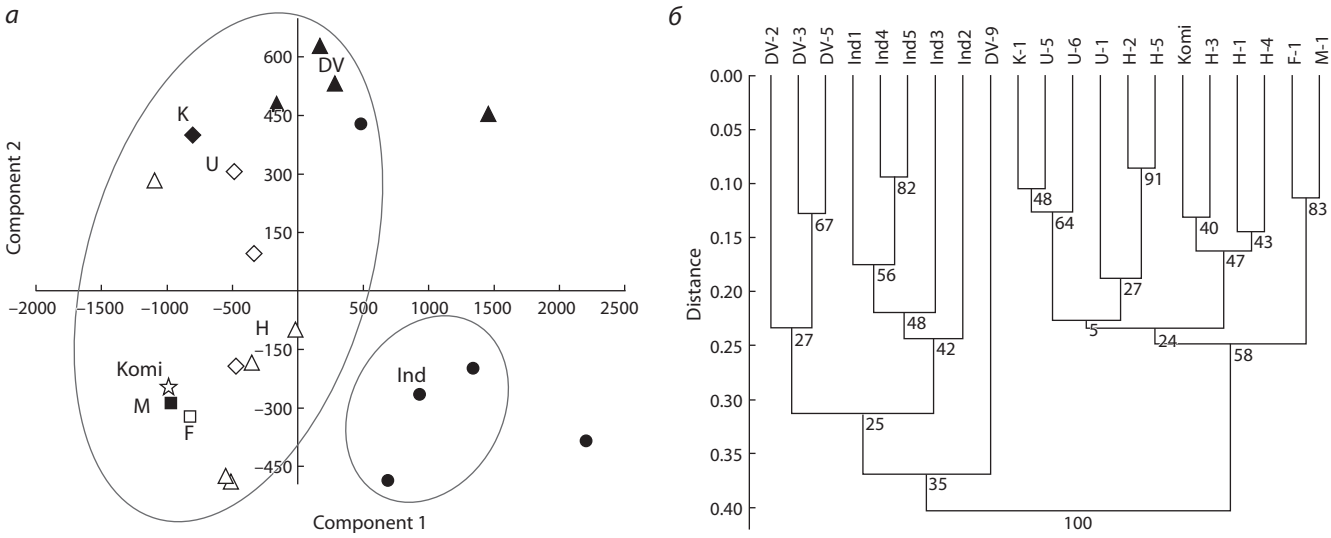


Рис. 1. Распределение образцов на основании результатов статистической обработки морфологических признаков растений *M. arvensis* и *M. canadensis*: *а* – метод главных компонент; *б* – кластерный анализ.

Условные обозначения: ■ – Московская область (M-6); ◆ – Калужская область (K-1); □ – Владимирская область (F); △ – Республика Хакасия (H); ▲ – Дальний Восток (DV); ☆ – Республика Коми; ◇ – Украина (U-2); ● – Индокитай (Ind1-5).

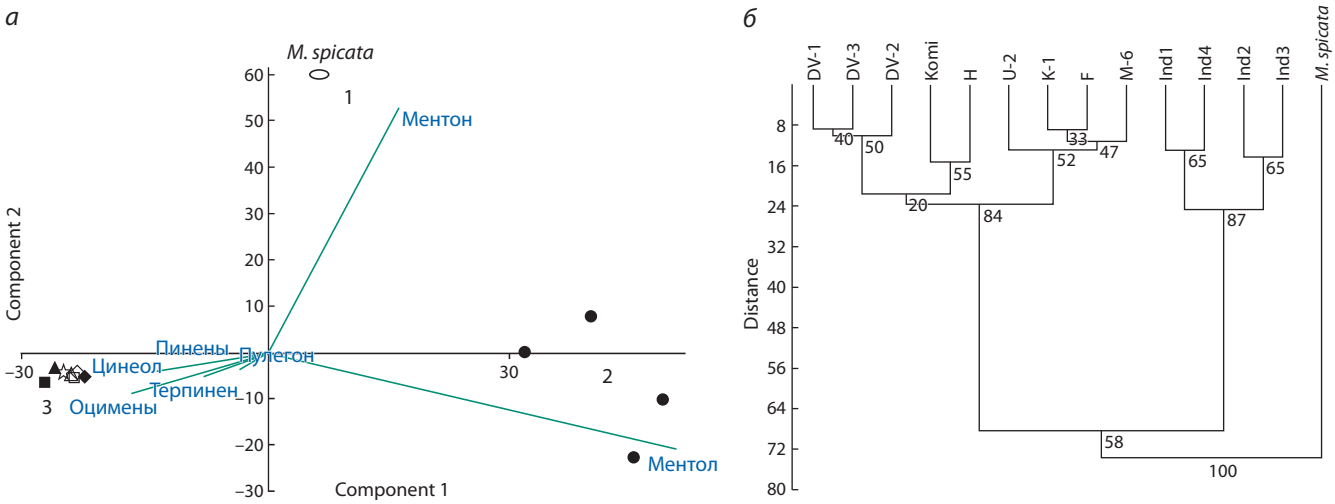


Рис. 2. Распределение образцов на основании результатов анализа состава эфирного масла в надземной части растений *M. arvensis*, *M. canadensis* и *M. spicata*: *а* – метод главных компонент; *б* – кластерный анализ.

Усл. обозн. см. рис. 1.

Самостоятельным критерием идентификации включенных в исследование растений *Mentha* может служить компонентный состав эфирного масла. В его составе обнаружено до 47 компонентов. Все компоненты с содержанием более 0.1 % от общей суммы легко идентифицировались по времени удерживания и масс-спектрам. Соотношение отдельных компонентов в составе эфирного масла значительно варьирует у растений из разных регионов.

По составу эфирного масла образцы распределились в целом сходным с морфологическими данными образом, но для растений из флоры Дальнего Востока отмечено более четкое и стабильное положение среди остальных образцов, независимо от методов статистической обработки (рис. 2). При использовании метода главных компонент (см. рис. 2, *а*) образцы разделились на три основные группы: 1 – *M. spicata* (основной компонент – ментон); 2 – индоазиатские растения (ментол); 3 – большая группа образцов, куда вошли растения Московской, Калужской и Владимирской областей, Дальнего Востока, Коми, Хакасии и Украины (для этой группы преобладающими компонентами были *транс*- и *цис*-β-оцимены, γ-терпинен, 1,8-цинеол, α- и β-пинены и пулегон).

Кластерный анализ (с использованием евклидова расстояния) компонентного состава эфирного масла разделил образцы на два кластера с высокой степенью поддержки (уровень бутстреп-поддержки 84–87 %) (см. рис. 2, *б*). *M. spicata*, взятая в качестве внешней группы, заняла базальное положение по отношению к двум кластерам, доминирующим соединением в компонентном составе ее эфирного масла был ментон (67.9 %) (табл. 3). В первый кластер вошли растения из Индокитая, для них было характерно накопление ментола (46.9–67.7 %) и его про-

Таблица 3. Компонентный состав эфирного масла растений *Mentha*

Наименование компонента	Содержание компонентов, % от цельного эфирного масла													
	M. spicata	M-6	K-1	F	H	Komi	U-2	DV-1	DV-2	DV-3	Ind1	Ind2	Ind3	Ind4
α-Пинен	0.53	2.20	1.15	1.91	2.04	2.86	0.98	5.24	3.85	4.19	0.25	0.41	0.95	0.45
β-Пинен	0.02	4.51	1.59	2.19	1.44	6.36	3.75	7.08	6.82	7.84	0.26	0.47	1.00	0.48
Сабинен	1.01	2.46	1.00	0.90	4.36	3.73	1.84	6.03	7.84	6.79	0.23	0.20	0.32	0.21
β-Мирцен	0.61	3.18	3.68	2.33	6.68	1.13	4.68	t*	6.16	2.85	0.39	0.36	0.73	0.39
Лимонен	0.69	0.65	2.09	1.02	1.88	0.82	6.05	1.24	1.30	1.19	1.99	0.93	2.04	2.13
1,8-Цинеол	3.62	***13.89*** **	3.28	***8.87***	***12.75***	***24.73***	***15.28***	***20.06***	***28.82***	***20.15***	0.20	0.31	0.41	0.20
транс-β-оцимен	0.08	***11.17***	***8.98***	2.01	7.59	5.91	7.26	***10.85***	***8.85***	6.62	0.02	t	0.02	0.15
цис-β-оцимен	0.04	***12.61***	***11.73***	2.19	9.73	6.18	***8.54***	***13.12***	***9.51***	***8.95***	t	0.28	0.01	0.06
γ-Терпинен	1.11	***21.81***	***19.32***	***13.59***	4.23	3.21	***11.79***	0.74	0.52	3.54	0.01	t	t	t
Терпинолен	0.19	1.97	13.46	5.72	3.46	0.12	3.15	0.30	0.20	0.78	0.01	t	0.01	0.01
Линалоол	0.08	0.45	1.02	2.79	2.60	1.20	1.02	0.45	0.78	0.51	0.92	0.28	0.02	0.011
Ментол	3.16	0.58	1.60	1.81	0.21	0.58	0.98	0.42	0.63	0.71	***46.92***	***67.67***	***67.32***	***52.96***
α-Терпинеол	0.32	1.58	1.65	1.79	3.19	1.72	2.74	2.81	1.84	1.21	0.17	t	0.12	t
Ментон	***67.88***	0.22	0.56	0.85	0.37	0.56	0.64	t	t	t	***19.25***	3.11	***16.36***	***30.33***
Изоментон	4.68	0.48	0.11	0.27	0.43	0.44	0.36	0.19	0.28	0.12	***22.69***	6.29	5.90	2.33
Изопулегон и пулегон	1.17	0.98	3.13	0.79	***21.14***	***15.51***	1.63	0.87	0.56	0.61	2.54	5.79	0.49	0.28
Метилацетат	0.36	0.11	0.63	0.81	0.47	0.10	0.72	0.39	0.21	0.78	8.94	7.05	0.42	1.95
Изоментол	1.03	0.21	0.22	0.25	0.18	0.18	0.22	0.38	0.42	0.73	0.38	1.03	0.04	1.51
β-Кариофиллен	1.69	2.99	1.55	1.96	2.29	7.96	6.97	5.24	2.60	2.41	0.61	0.55	0.65	0.40
Гермекрен D	1.08	4.38	5.07	3.22	3.84	5.28	3.45	0.98	0.80	1.47	0.60	0.93	0.94	0.32
Пиперитон	0.70	0.12	0.13	0.36	0.10	0.06	5.25	1.03	0.55	0.54	0.84	0.68	0.46	2.77

* Компонент отсутствует в составе эфирного масла.
** Полужирным курсивом выделены мажорные компоненты эфирного масла.

изводных – ментона (3.11–30.3 %) и изоментона (2.33–22.7 %) (см. табл. 3).

Во второй кластер вошли все остальные образцы: растения из средней полосы, Коми, Хакасии и Дальнего Востока. Второй кластер (бутстреп-поддержка <84 %) делится в свою очередь на три хорошо поддержанных субкластера. Выделяется субкластер, состоящий из растений *M. arvensis* Московской (M-6), Калужской (K-1) и Владимирской (F) областей и Украины (U-2). Два самостоятельных хорошо поддержанных сестринских подкластера образованы мятами Дальнего Востока (DV), Республики Коми (Komi) и Хакасии (H).

Как видно из табл. 3, популяции M-6, K-1, F и U-2 выделяются преимущественно по содержанию суммы *транс*- и *цис*-β-оцименов (4.2–23.8 %) и γ-терпинена (11.8–21.8 %), популяции DV, Komi и H – 1,8-цинеола (20.1–28.8 %) и суммы *транс*- и *цис*-β-оцименов (15.6–24.0 %), а в популяциях Komi и H, кроме 1,8-цинеола (24.7 и 12.8 %) и суммы *транс*- и *цис*-β-оцименов (12.1 и 17.3 %), зафиксированы пулегон и его предшественник изопулегон (15.5 и 21.1 % соответственно). Парные тесты Хотеллинга свидетельствуют о достоверном различии ($p < 0.02$ с применением коррекции Бонферрони) всех популяций. Таким образом, выявленные различия между исследованными растениями *Mentha* позволяют выделить следующие хемотипы: ментоловый (растения из Индокитая); *транс*- и *цис*-оцименовый и γ-терпиненовый (растения *M. arvensis* из Московской, Калужской и Владимирской областей и Украины), цинеоловый и *транс*- и *цис*-оцименовый (растения Дальнего Востока) и условно пулегоновый (растения из Республики Коми и Хакасии).

Для более полного изучения и выявления степени генетического сходства популяций исследовался состав ISSR-фрагментов ДНК. ПЦР проводили по двум схемам. В первую входили растения из Индокитая, Хакасии, европейской части России и Западной Украины, а также *M. spicata* в качестве внешнего стандарта. Во вторую схему опыта были добавлены образцы из флоры Дальнего Востока.

В первой схеме ПЦР была получена матрица, состоящая из 117 полиморфных фрагментов. При анализе данных методом главных координат с использованием расстояния Дайса исследуемые образцы распределились на две основные группы (рис. 3, *а*). Одна из этих групп включает растения из флоры Индокитая, во вторую вошли все образцы из европейской части России и Хакасии. Обособленное положение заняли *M. spicata* и расположенный рядом образец Ind1 флоры Индокитая (показан серым цветом). Образцы из Украины сформировали группу, почти не обособленную от других европейских и хакасских образцов.

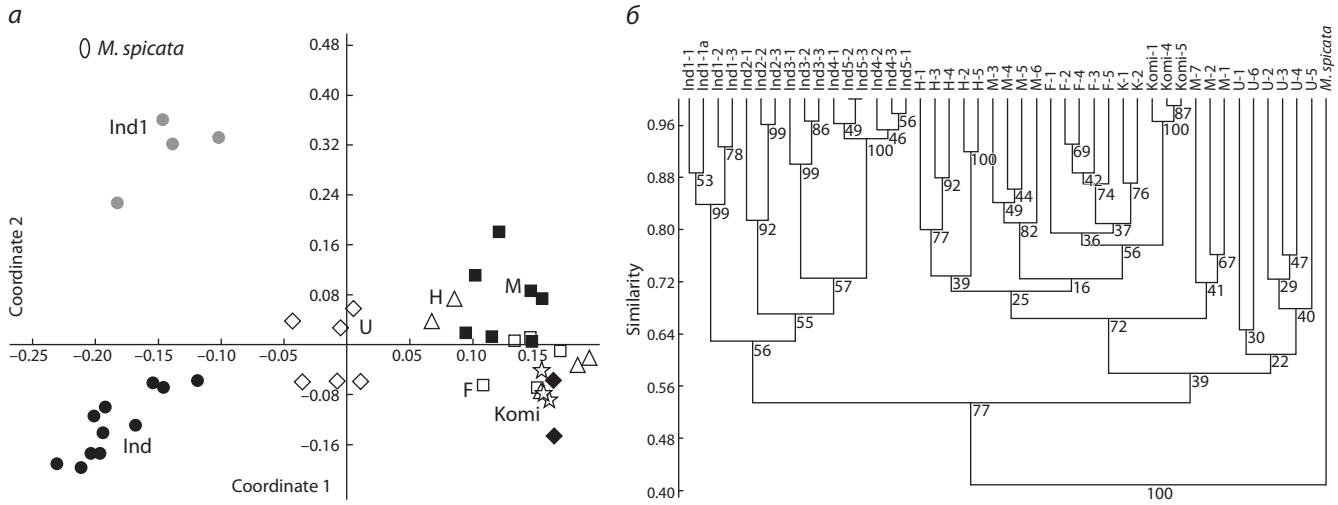


Рис. 3. Распределение образцов на основании результатов анализа состава ISSR-фрагментов ДНК в растениях *M. arvensis*, *M. canadensis* и *M. spicata*: *а* – метод главных координат; *б* – кластерный анализ.

Усл. обозн. см. рис. 1.

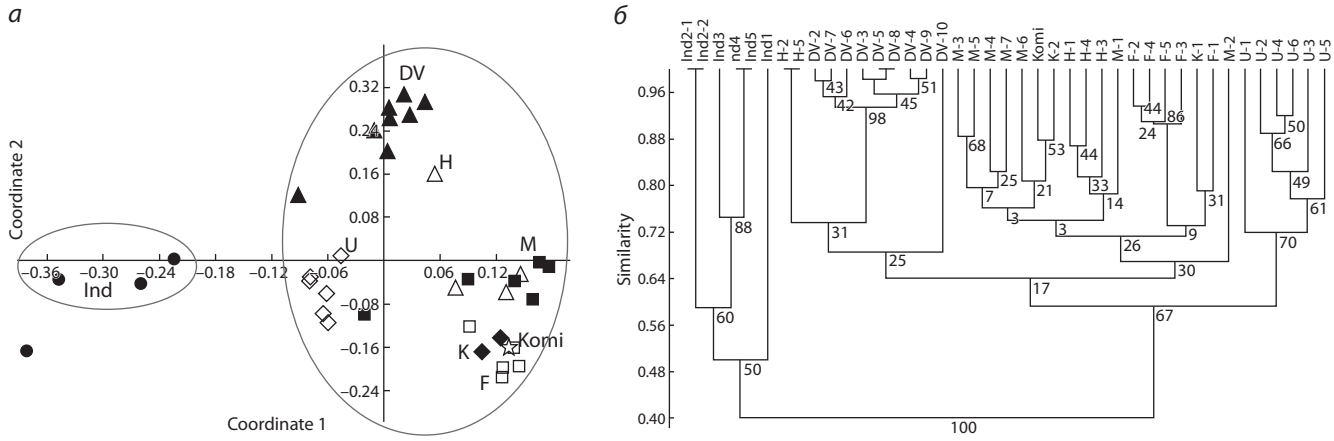


Рис. 4. Распределение образцов на основании результатов анализа состава ISSR-фрагментов ДНК в растениях *M. arvensis* и *M. canadensis*: *а* – метод главных координат; *б* – кластерный анализ.

Усл. обозн. см. рис. 1.

Кластерный анализ (см. рис. 3, *б*) также позволил выделить две большие группы растений со средним уровнем бутстреп-поддержки (39–56 %) и достаточно высокой мерой подобия (0.6): в первой находятся индоазиатские растения, во второй – европейские и хакасские.

По результатам второй схемы ПЦР была составлена матрица на основе 80 полиморфных фрагментов. После обработки полученных данных методом главных координат и кластерного анализа с использованием критерия Дайса растения тоже разделились на две группы (рис. 4). Первую составили растения из Индокитая, во вторую вошли все остальные растения, причем часть из них образовала общее облако/кластер, а другие сформировали отдельные группы, примыкающие к общей. Для образцов из Индокитая характерно наличие специфических для всей группы ампликонов, которые отсутствовали у остальных растений, в том числе у образцов с Дальнего Востока.

Внутри сборной второй группы популяции разделились на несколько подгрупп. Растения из Московской, Владимирской и Калужской областей, а также Республики Коми и Хакасии образовали общее облако, в котором образцы из разных популяций практически неразделимы, что свидетельствует о значительном генетическом сходстве между ними. К этой смешанной группе примыкают две относительно обособленные группы образцов, одна из которых образована популяцией Дальнего Востока вместе с двумя растениями из Хакасии, а вторая – образцами из Западной Украины, которые формируют облако, соприкасающееся с растениями из европейской части России. Все растения, собранные на Дальнем Востоке, имели очень высокую меру подобия (0.9) и бутстреп-поддержку кластера (98 %), что говорит о близком родстве изученных локальных популяций Дальнего Востока.

Обсуждение

В результате анализа количественных диагностических характеристик чашечки и листьев четко определяются европейская и индоазиатская группы образцов. В случае

же установления принадлежности дальневосточных растений к той или иной группе есть расхождения при использовании разных методов обработки данных, и, соответственно, затруднено определение их таксономического положения только по морфологическим признакам. Дополнительным критерием для разделения *M. arvensis* и *M. canadensis* может служить такой качественный показатель, как ультраструктура поверхности семян (Шелепова и др., 2016а), однако у растений из Индокитая семена отсутствовали, поскольку при выращивании этих образцов в условиях средней полосы России завязывание и созревание семян происходит редко.

По составу эфирного масла в надземной части растений все образцы распределились на разные хемотипы, что обусловлено генетическими особенностями растений. В литературе как для *M. arvensis*, так и для *M. canadensis* упоминается очень много хемотипов, выделенных на основе компонентов, присутствующих в количестве более 10 % (Tucker, Chambers, 2002). Для *M. arvensis* среди девяти хемотипов упоминаются такие, как лимоненовый, пулегоновый и 1.8-цинеоловый и/или β-пиненовый. В составе эфирного масла проанализированных нами растений *M. arvensis* доминировали транс- и цис-β-оцимены и γ-терпинен. Это ациклические предшественники лимонена и циклические промежуточные продукты терпеноидов. Известно, что лимонен через ряд биохимических реакций с участием гидро-, дегидроксилаз и редуктазы трансформируется в изопулегон, который, в свою очередь, является предшественником пулегона (Бугаенко, 2011), что и зафиксировано у растений *M. arvensis* из Коми и Хакасии (условно пулегоновый хемотип).

Для *M. canadensis* из Северной Америки выделяют среди прочих 1.8-цинеоловый и β-оцименовый, а для *M. canadensis* из Азии характерны многочисленные хемотипы, преимущественно с преобладанием ментола и ментона (Tucker, Chambers, 2002). Аналогичные результаты получены и в наших исследованиях: ментоловый хемотип у растений *M. canadensis* из Индокитая и цинеоловый и *транс*- и *цис*-оцименовый – у растений Дальнего Востока.

Анализируя состав ISSR-фрагментов, можно заключить, что растения *M. arvensis* из различных популяций европейской части России обладают значительным генетическим сходством и друг от друга почти не дифференцированы. Украинская и дальневосточная популяции, напротив, хорошо обособлены от остальных. Растения, произрастающие на Дальнем Востоке, согласно данным ISSR-анализа, хотя и образуют четко обособленную от европейских популяций группу, но все же находятся с ними в одном кластере/облаке. Растения из флоры Индокитая по составу ISSR-фрагментов значительно отличаются от остальных образцов и относятся к другому виду.

Между результатами молекулярно-генетического анализа, составом эфирного масла для растений *M. arvensis* и *M. canadensis* и, в некоторой степени, морфологическими признаками растений прослеживается сходный характер распределения образцов, т. е. наблюдается корреляция между молекулярными, фитохимическими и морфологическими параметрами. Таким образом, можно выделить группу наиболее генетически, морфологически и фитохимически близких растений, в которую вошли представители популяций Московской, Владимирской, Калужской областей, Республики Коми и Хакасии. Все они относятся к *M. arvensis*.

Заключение

Согласно полученным данным, как по компонентному составу эфирного масла, так и по составу ISSR-фрагментов растения, собранные в природной флоре Дальнего Востока, имели большее сходство с *M. arvensis* из европейской части, чем с коллекционными образцами из флоры Индокитая. В то же время для растений дальневосточной популяции отмечено определенное сходство с *M. canadensis* по некоторым морфологическим признакам (длина зубцов чашечки была промежуточной между *M. arvensis* из европейской части России и *M. canadensis* из Индокитая), по отдельным компонентам эфирного масла, которые характерны для *M. canadensis* из Северной Америки, кроме того, выявлена генетическая обособленность от остальных растений *M. arvensis*. Возможно, дальневосточная популяция представлена растениями гибридного происхождения и имеет более тесное родство с *M. arvensis*, чем с растениями из Индокитая, которые, по всей вероятности, относятся к *M. canadensis*. Исходя из вышесказанного, можно заключить, что комплексный анализ морфологических, фитохимических и молекулярных данных позволяет наиболее полно исследовать полиморфизм видов и уточнить видовую принадлежность растений.

Список литературы / References

Бугаенко Л.А. Генетические закономерности биосинтеза терпеноидов у мяты. Симферополь: Бизнес-Информ, 2011.
[Bugaenko L.A. Genetic Patterns of Terpenoid Biosynthesis in Mint. Simferopol: Business Inform Publ., 2011. (in Russian)]

Гинзберг А.С. Упрощенный способ определения эфирного масла в эфироносах. Хим.-фармацевт. промышленность. 1932;8(9): 326-329.
[Ginsberg A.S. A simple method for essential oil assay in essential-oil-bearing plants. Khimiko-Farmatsevticheskaya Promyshlennost = Chemical-Farm. Industry. 1932;8(9):326-329. (in Russian)]

Губанов И.А., Киселева К.В., Новиков В.С., Тихомиров В.Н. *Mentha arvensis* L. s. l. – Мята полевая. Иллюстрированный определитель растений Средней России. М.: Т-во науч. изданий КМК; Ин-т технол. исследований, 2004;3:133.
[Gubanov I.A., Kiseleva K.V., Novikov V.S., Tikhomirov V.N. *Mentha arvensis* L. s. l. – Field Mint. Illustrated Guide to Plants of Central Russia. Moscow: KMK Publ.; Institute of Technological Research, 2004;3:133. (in Russian)]

Доронькин В.М. *Mentha* L. В: Флора Сибири. Под ред. Л.И. Малышева. Новосибирск: Наука, 1997;11:222-225.
[Doron'kin V.M. *Mentha* L. In: Malyshev L.I. (Ed.) Flora of Siberia. Novosibirsk: Nauka Publ., 1997;11:222-225. (in Russian)]

Племенков В.В. Химия изопреноидов. Барнаул: Изд-во Алт. ун-та, 2007.
[Plemenkov V.V. Chemistry of Terpenoids. Barnaul: Altay University Publ. 2007. (in Russian)]

Шелепова О.В., Воронкова Т.В., Смирнова И.М., Енина О.Л. Особенности железистого аппарата растений *Mentha arvensis* L. разного географического происхождения. Изв. Самар. НЦ РАН. 2012;14(1):1884-1887.
[Shelepova O.V., Voronkova T.V., Smirnova I.M., Enina O.L. Features of the glandular apparatus of *Mentha arvensis* L. plants from different geographical regions. Izvestiya Samarskogo Nauchnogo Tsentra Rossiyskoy Academii Nauk = Proceedings of the

Samara Research Center of the Russian Academy of Sciences. 2012;14(1):1884-1887. (in Russian)]

Шелепова О.В., Семенова М.В., Воронкова Т.В., Олехнович Л.С., Бидюкова Г.Ф. Анатомо-морфологические и фитохимические отличия *Mentha canadensis* L. и *Mentha arvensis* L. Бюл. Ботан. сада-института ДВО РАН. 2016а;15:89-92.
[Shelepova O.V., Semenova M.V., Voronkova T.V., Olekhnovich L.S., Bidyukova G.F. Anatomical, morphological and phytochemical differences between *Mentha canadensis* L. and *Mentha arvensis* L. Bulleten Botanicheskogo Sada-Instituta DVO RAN = Bulletin of the Botanical Garden-Institute, Far-Eastern Branch of the RAS. 2016a;15:89-92. (in Russian)]

Шелепова О.В., Семенова М.В., Шанцер И.А., Степанова Н.Ю. Генетические и морфологические особенности южно-сибирских популяций *Mentha arvensis* L. В: Материалы XV Междунар. науч.-практ. конф. «Проблемы ботаники Южной Сибири и Монголии». Барнаул. 2016б;15:446-450.
[Shelepova O.V., Semenova M.V., Schantser I.A., Stepanova N.Yu. Genetic and morphological features of South Siberian *Mentha arvensis* L. populations. In: Proc. of the XV Int. practical and sci. conf. “Problems of Botany of South Siberia and Mongolia”. Barnaul, 2016b;15:446-450. (in Russian)]

Шелепова О.В., Семенова М.В., Енина О.Л., Шанцер И.А. Генетический, фенотипический и фитохимический полиморфизм восточноевропейских популяций *Mentha arvensis* L. Генетика. 2017;53(1):54-62. DOI 10.7868/S0016675816120134.
[Shelepova O.V., Semenova M.V., Enina O.L., Schantser I.A. Genetic, phenotypic and phytochemical polymorphism of Eastern European populations *Mentha arvensis* L. Russ. J. Genet. 2017; 53(1):59-66. DOI 10.1134/S1022795416120139.]

Briquet J. Fragmenta monographiae Labiatarum. Fascicule troisieme. Decades Mentharum novarum. Bull. Herb. Boiss. 1894;2:691-709.

Bunsawat J., Elliott N., Hertweck K., Lawrence A. Phylogenetics of *Mentha* (Lamiaceae): evidence from chloroplast DNA sequences. Syst. Bot. 2004;29(4):959-964. DOI 10.1600/0363644042450973.

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

Gouyon P.H., Vernet Ph., Guillerm J.L., Valdeyron G. Polymorphisms and environment: the adaptive value of the oil polymorphisms in *Thymus vulgaris* L. Heredity. 1986;57:59-66.

Hammer O., Harper D.A.T., Ryan P.D. PAST: Palaeontological Statistics software package for education and data analysis. Palaeontologia Electronica. 2001;4(1):4. Available at: https://palaeo-electronica.org/2001_1/past/issue1_01.htm

Hara H. Contributions to the study of variations in the Japanese plants closely related to those of Europe or North America. Part 2. J. Fac. Sci. Univ. Tokyo Sect. 3, Bot. 1956;6:343-391.

Holmes E.M. The Japanese peppermint plant. Pharm. J. Trans. 1882; 13:381-382.

Hua C.X., Wang G.R., Lei Y. Evaluation of essential oil composition and DNA diversity of mint resources from China. Afr. J. Biotechnol. 2011;10(74):16740-16745. DOI 10.5897/AJB11.1428.

Jedrzejczyk I., Rewers M. Genome size and ISSR markers for *Mentha* L. (Lamiaceae) genetic diversity assessment and species identification. Ind. Crops Prod. 2018;120:171-179. DOI 10.1016/j.indcrop.2018.04.062.

Linne C. Mentha: Species Plantarum. Holmaiae, 1753.

Rodrigues L., Póvoa O., Van den Berg C., Figueiredo A.C., Moldão M., Monteiro A. Genetic diversity in *Mentha cervina* based on morphological traits, essential oils profile and ISSRs markers. Biochem. Syst. Ecol. 2013;51:50-59. DOI 10.1016/j.bse.2013.08.014.

Tucker A.O., Chambers H.L. *Mentha canadensis* L. (Lamiaceae): a relict amphidiploid from the Lower Tertiary. Taxon. 2002;51(4): 703-718.

Tucker A.O., Naczi R.F.C. Mentha: an overview of its classification and relationships. In: Lawrence B.M. (Ed.). Mint. The Genus *Mentha*. CRC Press, 2007;3-39.

WCSP. 2019. World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. Available at: http://apps.kew.org/wcsp/Retrieved

ORCID ID
M.V. Semenova orcid.org/0000-0002-4627-0802
O.L. Enina orcid.org/0000-0002-8971-6384
O.V. Shelepova orcid.org/0000-0003-2011-6054

Благодарности. Работа выполнена в рамках ГЗ ГБС РАН № 118021490111-5.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 19.02.2019. После доработки 17.07.2019. Принята к публикации 22.07.2019.

ЭКОЛОГИЧЕСКАЯ ГЕНЕТИКА
Оригинальное исследование / Original article
УДК 582.546.6:543.544

Вавиловский журнал генетики и селекции. 2019;23(8):1076-1081
DOI 10.18699/VJ19.583

Метод газовой хроматографии-масс-спектрометрии для таксономии мискантуса

Н.М. Слынько[1] ✉, Н.В. Бурмакина[1], О.М. Поцелуев[1], С.Ю. Капустянчик[1], Г.Ю. Галицын[1], Т.Н. Горячковская[1], Л.В. Куйбида[2], С.В. Шеховцов[1], С.Е. Пельтек[1], В.К. Шумный[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт химической кинетики и горения им. В.В. Воеводского Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: nslynko@bionet.nsc.ru

Выделение новых видоопределяющих признаков актуально для систематики мискантуса и маркирования генотипов гибридов. Культивируемый в мире гибрид мискантуса *Miscanthus × giganteus* является аллополиплоидом, содержащим геномы *Miscanthus sacchariflorus* (в качестве материнского) и *M. sinensis*. В работе проведен хемотаксономический анализ образцов *M. sinensis* и *M. sacchariflorus*, собранных на Дальнем Востоке, и гибридных образцов как природного происхождения, так и полученных искусственно. В экстрактах 11 образцов мискантуса методом газовой хроматографии-масс-спектрометрии (ГХ-МС) обнаружено 153 соединения, из которых 143 идентифицировано. Из этих соединений в соответствии с их химическим строением были выделены группы алканов (20 соединений), жирных кислот (34), фенолов (13), стеролов (18), токоферолов (8), нортерпеноидов (12) и фитолов (13) вместе с их производными. Основные компоненты экстрактов образцов мискантуса – жирные кислоты и их производные (общее содержание 19.94–41.02 %), с преобладанием пальмитиновой и линоленовых кислот, а также стеролы (особенно β-ситостерол, стигмастерол и α-амирин), на долю которых приходится 17.15–31.73 %. Значения «индекса нечетности» CPI для алкановых компонентов экстракционных смесей были в диапазоне 1.55–7.18, причем экстракты из листьев дальневосточных образцов характеризовались значениями нижней половины этого диапазона (1.55–2.74), тогда как экстракты из листьев гибридов – значениями верхней половины (5.78–7.18). Анализ экстракционных профилей методом главных компонент позволил выявить три отчетливо разделенных кластера, объединяющих образцы *M. sinensis*, *M. sachariflorus* и их гибридов, и уточнить таксономическое отнесение одного из гибридов. Хемотаксономическое отнесение в целом согласуется с результатами молекулярно-генетического анализа последовательности фрагмента пластидного генома мискантуса, который позволил также подтвердить видовую природу материнских растений, использованных для получения этих гибридов. Хемотаксономический анализ с использованием метода ГХ-МС может оказаться эффективным дополнительным инструментом в таксономическом отнесении различных морфологических форм мискантуса к *M. sinensis* или *M. sachariflorus*, а также для хемотаксономической характеристики гибридов.
Ключевые слова: мискантус; листья; таксономия; ГХ-МС; метод главных компонент; пластидный геном.

Для цитирования: Слынько Н.М., Бурмакина Н.В., Поцелуев О.М., Капустянчик С.Ю., Галицын Г.Ю., Горячковская Т.Н., Куйбида Л.В., Шеховцов С.В., Пельтек С.Е., Шумный В.К. Метод газовой хроматографии-масс-спектрометрии для таксономии мискантуса. Вавиловский журнал генетики и селекции. 2019;23(8):1076-1081. DOI 10.18699/VJ19.583

Gas chromatography-mass spectrometry in the taxonomy of Miscanthus

N.M. Slynko[1] ✉, N.V. Burmakina[1], O.M. Potseluyev[1], S.Yu. Kapustyanchik[1], G.Yu. Galitsin[1], T.N. Goryachkovskaya[1], L.V. Kuybida[2], S.V. Shekhovtsov[1], S.E. Peltek[1], V.K. Shumny[1]

[1] Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
[2] Voevodsky Institute of Chemical Kinetics and Combustion, SB RAS, Novosibirsk, Russia
✉ e-mail: nslynko@bionet.nsc.ru

Chemotaxonomy as a system approach deals with intra- and interspecific polymorphism of a group of taxa in order to clarify their taxonomic positions or to select material for selection or introduction. In this study we performed chemotaxonomic analysis of specimens of *Miscanthus sinensis* and *M. sacchariflorus* collected in the Russian Far East and of hybrid plants of both natural and artificial origin. We found 153 substances and identified 143 of them in extracts of eleven Miscanthus plants by gas chromatography-mass spectrometry (GC-MS). These substances can be grouped into alkanes (20 compounds), fatty acids (34), phenols (13), sterols (18) tocopherols (8), norterpenoids (12), and phytols (13), as well as their derivatives. The main components of the extracts of miscanthus samples are fatty acids and their derivatives (total content 19.94–41.02 %), dominated by palmitic and linolenic acids, and sterols (mainly β-sitosterol, stigmasterol, and α-amyrin), which constitute 17.15–31.73 %. The values of the CPI "oddness index" for the alkane components of the extracts were within 1.55–7.18, with extracts from leaves of the Far Eastern samples characterized by the lower half of this range (1.55–2.74), while extracts from leaves of hybrids fell to the upper half (5.78–7.18). Principal component analysis of extraction profiles allowed us to separate three distinct clusters: *M. sinensis*, *M. sachariflorus*, and their

© Слынько Н.М., Бурмакина Н.В., Поцелуев О.М., Капустянчик С.Ю., Галицын Г.Ю., Горячковская Т.Н., Куйбида Л.В., Шеховцов С.В., Пельтек С.Е., Шумный В.К., 2019

hybrids, as well as to verify the origin of one of the natural hybrids. The results of chemotaxonomic analysis mostly matched those of DNA sequencing of a fragment of the plastid genome, which, moreover, allowed us to identify the species nature of the maternal plants used to obtain these hybrids. Chemotaxonomic analysis using GC-MS was found to be an efficient additional technique to delimit various morphological forms of *M. sinensis*, *M. sachariflorus*, and their hybrids.
Key words: *Miscanthus*; leaves; taxonomy; GC-MS; PCA; plastid genome.

For citation: Slynko N.M., Burmakina N.V., Potseluyev O.M., Kapustyanchik S.Yu., Galitsin G.Yu., Goryachkovskaya T.N., Kuybida L.V., Shekhovtsov S.V., Peltek S.E., Shumny V.K. Gas chromatography-mass spectrometry in the taxonomy of Miscanthus. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(8): 1076-1081. DOI 10.18699/VJ19.583 (in Russian)

Введение

В хемотаксономии основной критерий разделения организмов – биохимическая характеристика вторичных метаболитов, пути синтеза которых складываются в ходе эволюционного развития. В качестве маркеров используют такие соединения, как сесквитерпеноиды (для тысячелистника) (Покровская и др., 2009), фитоэкдистероиды (для сем. гвоздичные (Caryophyllaceae Juss.)), (Дармограй и др., 2016), летучие терпеновые соединения при анализе подвидов облепихи *Hippophae rhamnoides* (Socaci et al., 2013), алкалоиды (для сем. Annonaceae) (Silva et al., 2016). Хемотаксономические маркеры представляют интерес для систематики близкородственных видов. Например, сесквитерпеновые соединения из эфирных масел оказались дополнительными хемотаксономическими критериями установления видового статуса различных полыней (Namzalov et al., 2019). Выделение новых видоопределяющих признаков актуально для родов и семейств растений, объединяющих большое количество видов.

Одним из таких растений является мискантус (веерник), широко используемый в ландшафтном дизайне. Все виды мискантуса – многолетние корневищные злаки. Высота стебля у некоторых видов достигает 7 м (Chen, Renvoize, 2006). В настоящее время в мире происходит увеличение площадей культивирования мискантуса и разрабатываются новые технологии на основе его биомассы. Промышленное значение имеют три таксона мискантуса: *Miscanthus sachariflorus*, *M. sinensis* и *M.* × *giganteus*. *M.* × *giganteus* представляет собой стерильный гибрид *M. sinensis* × *M. sacchariflorus* (Hodkinson, Renvoize, 2001). На территории России в дикой природе распространены мискантус сахароцветный (*M. sacchariflorus*) и мискантус китайский (*M. sinensis*).

Культивируемый в Европе и Новой Зеландии гибрид *M.* × *giganteus* является триплоидным (Linde-Laursen, 1993) и морфологически по соцветию более похож на *M. sacchariflorus*, а по структуре ризом и корневища – на *M. sinensis*. У мискантуса сахароцветного нижняя цветковая чешуя с короткой прямой остью, волоски хохолка белые, иногда красноватые, корневища тонкие, ползучие. Мискантус китайский имеет укороченные толстые корневища. Колосковые чешуи несут длинные волоски. Волоски хохолка негустые, сильно оттопыренные, грязнобелые, но нередко бывают красноватые (Hodkinson et al., 2016).

В настоящей работе проведен хемотаксономический анализ образцов *M. sinensis* и *M. sacchariflorus*, собранных на Дальнем Востоке, и гибридных образцов как природного происхождения, так и полученных искусственно.

Материалы и методы

Образцы растений. Из коллекции, собранной во время экспедиции в Приморском крае, взяты образцы мискантуса *M. sinensis* и *M. sacchariflorus* и, предположительно, дикие гибриды (Hodkinson et al., 2016). Коллективом, реализующим проект GrassMargins, на основе *M. sinensis* и *M. sacchariflorus* получен межвидовой гибрид *M.* × *giganteus*.

Корни образцов, выкопанные во время экспедиционной поездки в Приморский край в сентябре 2014 г., в октябре того же года были высажены в гидропонную теплицу Института цитологии и генетики Сибирского отделения Российской академии наук. Листья мискантуса были срезаны 9 сентября 2018 г. и высушены на воздухе в течение трех дней до постоянного веса.

Данные по происхождению образцов мискантуса сведены в табл. 1. Образцы 07–11 получены в проекте European Framework Programme 7, grant number 289461 (GrassMargins). Природные образцы 01–06 были классифицированы по морфологическим признакам (см. табл. 1) (Hodkinson, et al., 2016).

Методы анализа. Определение экстрактивных веществ: 5.00 г воздушно-сухого материала экстрагировали 100 мл смеси этилацетат + хлороформ в объемном соотношении 1:2 в аппарате Сокслета, экстракт фильтровали через плотный стеклянный фильтр, промывали фильтр этилацетатом до общего объема 50 мл, добавляли 0.25 мл раствора стандартного вещества (5,6,7,8-тетрафторбензобаррелена) в этилацетате (1.00 мг/мл), отбирали 1 мл, отгоняли основную часть растворителя, остаток наносили на окись алюминия III ст. акт., смоченную этилацетатом, и элюировали этилацетатом до объема 5.0 мл. Элюент анализировали методом газовой хроматографии-масс-спектрометрии (ГХ-МС).

Условия проведения анализа: хроматограф Agilent Technologies 6890 с масс-спектрофотометрическим детектором 5973; хроматографическая колонка – капиллярная DB-1 J&W с внутренним диаметром 0.25 мм и длиной 30 м; скорость газа-носителя (гелий) – 1.0 мл/мин; температура нагревателя ввода пробы 250 °C; температура термостата программировалась от 50 до 250 °C со скоростью 25 °C/мин; скорость ввода пробы 1.0 мл/мин. Ввод пробы (1 мкл) в хроматографическую колонку осуществляли без деления потока. Специфичность (селективность) определяли путем введения растворителя с целью подтвердить, что не было ложных пиков в течение целевого времени удерживания.

Хроматографический масс-спектрометрический анализ образцов для измерения полного ионного тока выпол-

Таблица 1. Характеристика анализируемых объектов

Образец	Вид*	Место сбора, координаты	Происхождение
01	*Miscanthus sinensis*	42.702491 с. ш., 130.814686 в. д. Хасанский район, восточная окраина п. г. т. Краскино, 50 м севернее а. т. 189	
02	Природный гибрид	42.612412 с. ш., 130.722623 в. д. Хасанский район, южнее с. Маячное, 250 м восточнее а. т. 189	
03	*M. sacchariflorus*	43.761063 с. ш., 131.454047 в. д. Уссурийский г. о., 2.0 км севернее с. Монакино	
04	*M. sacchariflorus*	44.473862 с. ш., 135.950393 в. д. Дальнегорский г. о., западнее кемпинга Зе Лангоу, 100 м севернее а. т. 181	
05	*M. sacchariflorus*	44.401283 с. ш., 131.774052 в. д. Пограничный район, южнее с. Нестеровка, 300 м восточнее а. т. 183	
06	*M. sacchariflorus*	44.669587 с. ш., 131.667229 в. д. Хасанский район, севернее с. Маячное, 250 м западнее а. т.183	
07	*M. sinensis* × *M. sacchariflorus*		Ru-Msi (*M. sacchariflorus*) × D-M7GoFal (*M. sinensis*) ♀
08	*M. sinensis* × *M. sacchariflorus*		(*M. sacchariflorus*) × Ru-Msi 18 (*M. sinensis*) ♀
09	*M. sinensis* × *M. sacchariflorus*		Ru-Msi (*M. sacchariflorus*) ♀ × 18 (*M. sinensis*)
10	*M. sinensis* × *M. sacchariflorus*		Ru-Msi (*M. sacchariflorus*) × D-82 (*M. sinensis*) ♀
11	*M. sinensis* × *M. sacchariflorus*		Ru-Msi (*M. sacchariflorus*) ♀ × D-82 (*M. sinensis*)

Примечание. а. т. – автотрасса; г. о. – городской округ.
* Видовое отнесение образцов 01–06 выполнено на основании морфологических признаков и описано в работе (Hodkinson et al., 2016).

няли как в режиме сканирования в диапазоне масс 10–800 а. е. м., так и в режиме селективного сканирования ионов. Индексы удерживания RI определяли с использованием данных хроматографического анализа смеси гомологичных линейных алканов (C_8–C_{32}) и рассчитывали из времен выхода пиков по (Dool, Kratz, 1963). Площади пиков всех компонентов рассчитывали с помощью программы Xcalibur 2.0. Идентификацию компонентов проводили с использованием встроенной базы масс-спектров The NIST Mass Spectral Search Program for the NIST/EPA/NIH Mass Spectral Library Version 2.0a, а также сравнением полученных значений RI со значениями из расширенного варианта той же базы (https://webbook.nist.gov/chemistry/gc-ri/).

Массовую долю определяемого компонента смеси (мкг/г экстракта) находили через соотношение площадей пиков стандартного и определяемого веществ. Принималось, что фактор отклика для каждого отдельного соединения одинаков для всех расчетов.

В сводке данных не учитывали хроматографические пики, принадлежащие веществам с очевидным антропогенным происхождением: 2,5-ди-трет-бутил-р-бензохинон (время удерживания RT 9.398 мин); 3-трет-бутил-4-гидроксианизол (RT 9.765 мин), н-додецилакрилат (RT = 10.396 мин), дибутилфталат (RT 11.517 мин), бис(2-этилгексил)фумарат (RT 13.044 мин), ди-(2-этилгексил)фталат (15.972 мин).

Многофакторный анализ был выполнен в программе Past 3, версия 3.25 (Hammer et al., 2001). ДНК мискантуса была экстрагирована из свежих либо сухих листьев мискантуса в соответствии с методикой, описанной в статье (Shekhovtsov et al., 2012). Для амплификации пластидного гена tRNA-Leu(UAA) и фланкирующих последовательностей использовали праймеры Misc-Leu-Fw (5′-TGGAA-GCTGT-TCTAA-CGAAT-C-3′) и Misc-Leu-Rv (5′-AATGG-GACTC-TCTCT-TTATC-CTC-3′). В состав реакции входили: 60 мМ Tris-HCl (pH 8.5), 1.5 мМ $MgCl_2$, 25 мМ KCl, 10 мМ 2-меркаптоэтанол, 0.1 % Тритон X-100, 0.7 мМ праймеров, 1 ед. полимеразы TaqSE («СибЭнзим», Новосибирск). Реакцию ПЦР проводили со следующим профилем реакции: 3 мин при 94 °C; 35 циклов, включающих 15 с при 94 °C, 15 с при 54 °C и 25 с при 72 °C; финальный цикл синтеза – 5 мин при 72 °C. Секвенирование ДНК проводили в Межинститутском центре секвенирования СО РАН.

Результаты

Согласно анализу ГХ-МС (рис. 1), в экстрактах 11 образцов мискантуса было обнаружено 152 соединения, из которых идентифицировано 143, представленных в Приложении[1]. Из этих соединений в соответствии с их химическим строением были выделены группы алканов (20 соединений), жирных кислот (34), фенолов (13), стеролов (18), токоферолов (8), нортерпеноидов (12) и фитолов (13) вместе с их производными.

Содержание алканов в экстрактах из образцов листьев мискантуса было в диапазоне 2.40–18.55 %, причем максимальные значения были характерны для гибридов мискантуса. Среди алканов в экстрактах из листьев дальневосточных образцов наибольшее содержание наблюда-

[1] Приложение см. по адресу:
http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx24.pdf

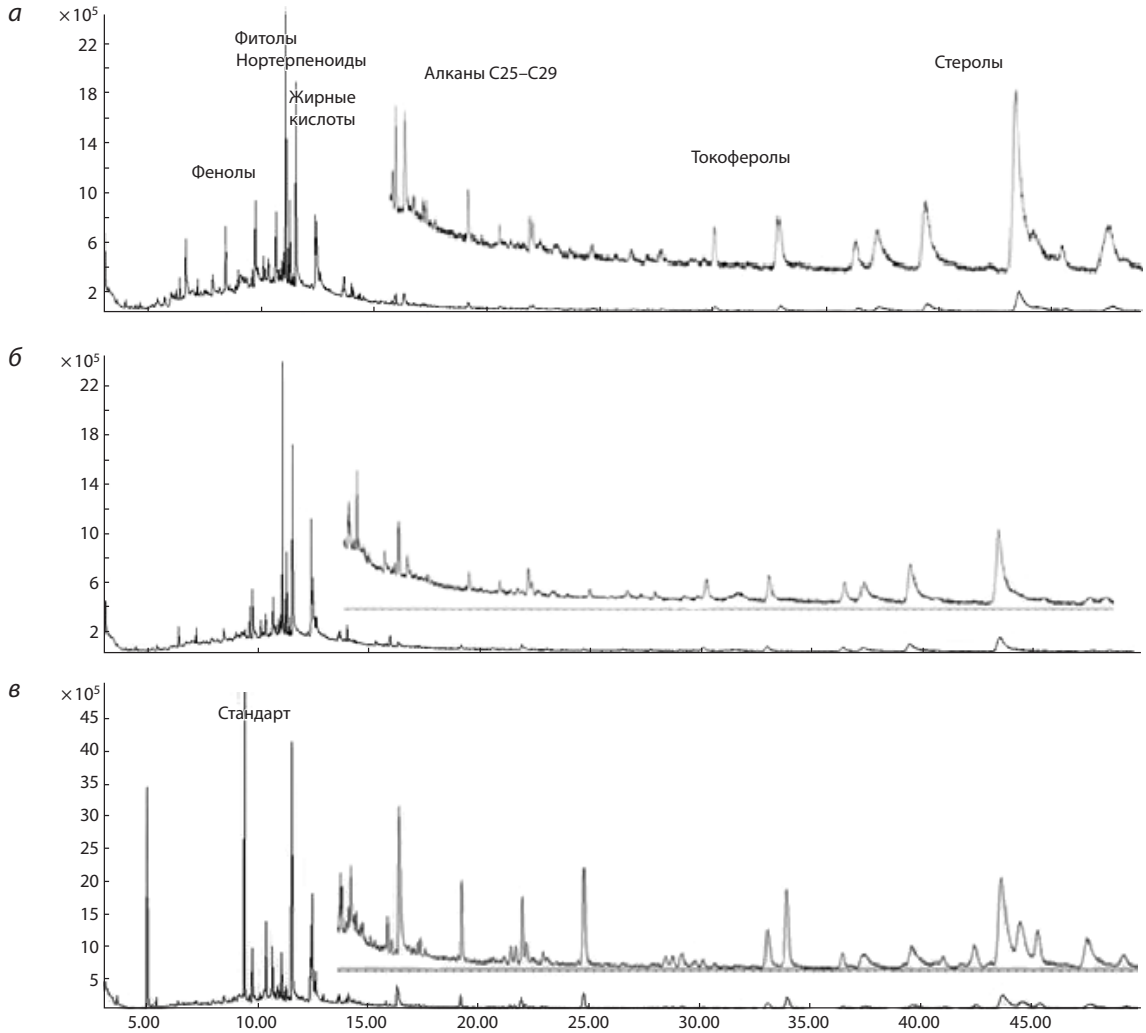


Рис. 1. Хроматограммы экстрактов мискантуса: *а* – образец 01, *M. sinensis*; *б* – образец 04, *M. sacchariflorus*; *в* – образец 09, гибрид *M. sinensis* × *M. sacchariflorus*.

лось у гептакозана (кроме образца 01), в то время как у гибридов лидером был нонакозан.

В группе жирных кислот и их производных ожидаемо доминировали кислоты с четным количеством атомов углерода и их эфиры. Пальмитиновая (C16:0) и линолевая (C16:2ω9,12) кислоты обладали наибольшим содержанием среди кислот, однако экстракты из листьев *M. sinensis* (образцы 01 и 02) содержали их в количестве 7.74–9.95 и 5.13–5.96 % соответственно, тогда как содержание этих кислот в остальных образцах было существенно выше: 11.20–15.49 %.

Фенолы же, напротив, имели максимальное содержание в образцах 01 и 02 – 11.78 и 16.20 % соответственно, тогда как их суммарное содержание в прочих образцах не превышало 7.54 %, а наименьшими концентрациями характеризовались гибриды (не более 4.19 %). Из фенолов наибольшие концентрации были зарегистрированы у 2-гидрокси-5-метилацетофенона, ранее обнаруженном только в табаке и кофе (Buckingham, 1996).

Сходное, но более сглаженное распределение характерно для фитостерольной фракции, в которой вполне ожидаемо превалировали β-ситостерол, стигмастерол и кампестерол (Winkler-Moser, 2012). Поверхностный анализ не выявляет каких-либо закономерностей в распределении этих соединений.

Представители группы нортерпеноидов, часто определяющих аромат растений, открыты в экстрактах из листьев мискантуса в относительно небольшом количестве, причем наибольшим содержанием отмечены образцы экстрактов из *M. sinensis* (3.54–5.54 %), а наименьшим – экстракты из листьев гибридов (0.75–1.24 %). Отнесение пиков изомеров мегастигматриен-3-она выполнено в соответствии с данными работы (Osorio et al., 2003).

Примерно одинаковым содержанием в экстрактах характеризуются пики токоферолов. Максимальная концентрация у витамина E: 1.99–2.59 %. Подобная картина у производных фитола, причем самые характеристические соединения в данной группе – это сам фитол (1.58–4.45 %) и продукт его дегидратации неофитадиен. Последний имеет выраженную область низких (1.09–2.30 % для экстрактов из листьев гибридов) и высоких (4.72–6.62 % для экстрактов из листьев *M. sachariflorus*) значений. Из неклассифицированных соединений (см. Приложение, «Прочие») обращает на себя внимание наличие в экстрактах листьев дальневосточных видов мискантуса азотсодержащих продуктов взаимодействия алифатических

альдегидов C3 и C4 с алифатическими аминами. Подобные соединения в экстрактах из листьев гибридов не зарегистрированы.

Обсуждение

При изучении объектов растительного мира как в живой природе, так и в палеонтологических отложениях профиль содержания алкановых соединений может снабдить исследователя важной информацией. н-Алканы способствуют гидрофобным свойствам листового воска и служат частью первого барьера защиты растения от внешней среды, спасая листья от потери воды в результате испарения (Jetter et al., 2006). Растения производят ряд н-алканов обычно с сильным превалированием нечетного числа атомов углерода над четным и одной или двумя доминирующими длинами цепей (Eglinton, Hamilton, 1967). Так, постулировано, что преобладание в смеси алканов гентриаконтана C31 является характеристичным для трав, в то время как максимумы для C27 и C29 более свойственны деревьям и кустарникам (Yi Duan, Jinxian He, 2011). Применительно к нашему исследованию это положение выполняется для гибридных образцов мискантуса 07–11, в то время как в хроматограммах экстрактов из образцов *M. sacharoflorus* 04–06 превалирует нонакозан C29, а в образцах 01 и 02 поиски гентриаконтана не увенчались успехом.

Еще один характеристичный таксономический параметр – так называемый индекс нечетности (carbon preference index, CPI), представляющий собой отношение содержания нечетных гомологов н-алканов к четным. Как правило, в растениях содержание нечетных алканов, образующихся главным образом при метаболическом декарбоксилировании более распространенных жирных кислот с четным числом атомов углерода, естественным образом должно превалировать над содержанием алканов с четным числом атомов. Действительно, рассчитанные значения CPI для алкановых компонентов экстракционных смесей лежали в диапазоне 1.55–7.18, причем экстракты из листьев дальневосточных образцов характеризовались значениями нижней половины этого диапазона (1.55–2.74), тогда как экстракты из листьев гибридов – значениями верхней половины (5.78–7.18).

Было полезным сравнить наши результаты с данными, полученными в сходных исследованиях. Газ-хроматографический анализ экстрактов листьев *M. sinensis* впервые был выполнен G.V. Weißmann и W. Lange (1991). Хотя авторы не привели численных значений содержания алканов, они отметили, что в экстрактах превалировали пентакозан, гептакозан и нанокозан. В работе (Villaverde et al., 2009) при описании состава экстрактов из стеблей и влагалищных листьев *M.* × *giganteus* среди 63 идентифицированных соединений алканы не упоминаются вовсе.

Анализ методом главных компонент (PCA) был рассчитан путем нормализации 142 переменных на основе 142 идентифицированных веществ по евклидовым расстояниям и средней связи основных компонентов первого дерева (рис. 2). Совокупная дисперсия первых двух компонентов составляет 75.77 %. Этот анализ позволяет выявить три отчетливо разделенных кластера, объединяющих образцы *M. sinensis*, *M. sachariflorus* и их гибридов,

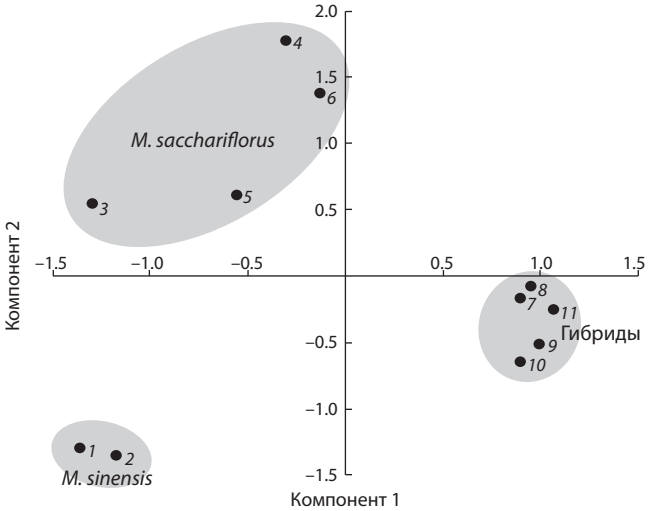


Рис. 2. PCI-график липофильных соединений, содержащихся в экстрактах 11 образцов мискантуса.

Таблица 2. Вариабельные сайты в исследуемом фрагменте хпДНК

Образец	Позиция			Видовое отнесение
	109	227	232	
01	G	A	T	*M. sacchariflorus*
03	G	A	T	»
04	G	A	T	»
05	G	A	T	»
06	G	A	T	»
07	T	A	A	*M. sinensis*
08	T	A	A	»
09	G	A	T	*M. sacchariflorus*
10	T	C	A	Промежуточный

причем бóльшая площадь кластера, принадлежащего *M. sachariflorus*, объясняется большим расстоянием между точками сбора образцов и, соответственно, принадлежностью к разным популяциям и различными условиями их произрастания. Следует отметить, что ареалы распространения *M. sinensis* и *M. sachariflorus* пересекаются и, вполне естественно, может произойти межвидовая гибридизация. Однако образец 01, отнесенный по морфологическим признакам к гибридному, по хемотаксономическим параметрам тесно кластеризуется вместе с *M. sinensis*, что дает основания предполагать их общую природу.

Представляло интерес сравнить результаты, полученные при использовании различных методов видоопределения. Молекулярно-генетические подходы считаются одними из наиболее надежных и достоверных. Мы получили последовательности фрагмента пластидного генома для девяти из одиннадцати исследованных образцов. Из табл. 2 следует, что два образца имеют гаплотип, типичный для *M. sinensis*, шесть – типичный для *M. sachariflorus* и один – редкий гаплотип, встречающийся у единичных образцов *M. sinensis*, а также *M. capensis*. При интерпретации результатов необходимо помнить, что ха-

рактер пластидных последовательностей гибридов указывает на видовую принадлежность материнского растения.

Интерпретация совокупности результатов не дает завершенной картины. Описанное в работе (Hodkinson et al., 2016) отнесение образцов 03–06 к *M. sacchariflorus* по морфологическим признакам, подтвержденное хемотаксономическим анализом; об этом же свидетельствуют результаты, полученные молекулярно-генетическим методом. Аналогично последним методом подтверждена видовая природа материнских растений, использованных для синтеза гибридов 07–09. В образце 10 видовое отнесение материнского растения к *M. sinensis* не выглядит однозначным. Отнесение образца 01 к *M. sinensis*, выполненное по хемотаксономическим признакам, результатами, полученными молекулярно-генетическим методом, не подтверждается. Для объяснения всех имеющихся противоречий, связанных с установлением видовой природы образца 01, можно предположить, что это растение является природным гибридом различных видов мискантуса, обладающим хемотаксономическим профилем, отличным от таковых у искусственных гибридов.

Заключение

Хемотаксономический анализ с использованием метода ГХ-МС в совокупности с данными, полученными другими методами, может оказаться эффективным дополнительным инструментом в таксономическом отнесении различных морфологических форм мискантуса к виду *M. sinensis* или *M. sachariflorus*. Эти данные имеют также прикладное значение, поскольку виды *M. sachariflorus* и *M. sinensis* используются для создания промышленных гибридов *Miscanthus* × *giganteus*. Изучение хемотаксономических характеристик гибридов позволит разработать «химические паспорта» технологически перспективных сортов мискантуса.

Список литературы / References

Дармограй С.В., Ерофеева Н.С., Филиппова А.С., Дармограй В.Н. Хемотаксономическое изучение некоторых видов рода ясколки (*Cerastium* L.) семейства гвоздичные (Caryophyllaceae Juss.). Усп. соврем. естествознания. 2016;7:32-36.
[Darmogray S.V., Erofeeva N.S., Filippova A.S., Darmogray V.N. Chemotaxonomical study of the *Cerastium* genus some plants of the Caryophyllaceae family. Uspekhi Sovremennogo Yestestvoznaniya = Advances in Current Natural Sciences. 2016;7:32-36. (in Russian)]

Покровская И.С., Мазова О.В., Апыхтин Н.Н., Племенков В.В. Хемотаксономия тысячелистника обыкновенного (*Achillea millefolium* L.). Химия раст. сырья. 2009;3:85-89.
[Pokrovskaya I.S., Mazova O.V., Apykhtin N.N., Plemenkov V.V. Chemotaxonomy yarrow (*Achillea millefolium* L.). Khimiya Rastitelnogo Syr'ya = Chemistry of Plant Raw Materials. 2009;3:85-89. (in Russian)]

Buckingham J. Dictionary of Organic Compounds. CRC Press, 1996; 8:3681.

Chen S.L., Renvoize S.A. *Miscanthus*. In: Wu Z.Y., Raven P.H., Hong D.Y. (Ed.). Flora of China. Beijing/St. Louis: Science Press/ Missouri Botanical Garden Press, 2006;22:581-583.

Dool H.V.D., Kratz P.D., A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. J. Chromatogr. A. 1963;11:463-471. DOI 10.1016/S0021-9673(01)80947-X.

Eglinton G., Hamilton R.J. Leaf epicuticular waxes. Science. 1967; 156:1322-1335.

Hammer Ø., Harper D.A.T., Ryan P.D. Past: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica. 2001;4(1):1-9. 178kb. http://palaeo-electronica.org/2001_1/past/issue1_01.htm.

Hodkinson T.R., Petrunenko E., Klaas M., Münnich C., Barth S., Shekhovtsov S.V., Peltek S.E. New breeding collections of *Miscanthus sinensis*, *M. sacchariflorus* and hybrids from Primorsky Krai, Far Eastern Russia. In: Perennial Biomass Crops for a Resource-Constrained World. Springer, 2016;105-118.

Hodkinson T.R., Renvoize S.A. Nomenclature of *Miscanthus* × *giganteus* (Poaceae). Kew Bull. 2001;56:759-760. DOI 10.2307/4117709.

Jetter R., Kunst L., Samuels A.L. Composition of plant cuticular waxes. In: Riederer M., Muller C. (Ed). Biology of the Plant Cuticle. Oxford: Blackwell Publ., 2006.

Linde-Laursen I.B. Cytogenetic analysis of *Miscanthus* '*Giganteus*', an interspecific hybrid. Hereditas. 1993;119:297-300. DOI 10.1111/j.1601-5223.1993.00297.x.

Namzalov B.B., Zhigzhitzhapova S.V., Dubrovsky N.G., Sakhyaeva A.B., Radnaeva L.D. Wormwoods of Buryatia: diversity analysis, ecological-geographical features, and chemotaxonomy of section Abrotanum. Acta Biologica Sibirica. 2019;5(3):178-187. DOI 10.14258/abs.v5.i3.6589.

Osorio C., Duque C., Batista-Viera F. Studies on aroma generation in lulo (*Solanum quitoense*): enzymatic hydrolysis of glycosides from leaves. Food Chem. 2003;81(3):333-340. DOI 10.1016/S0308-8146(02)00427-2.

Shekhovtsov S.V., Shekhovtsova I.N., Peltek S.E.. Phylogeny of Siberian species of *Carex* sect. *Vesicariae* based on nuclear and plastid markers. Nord. J. Bot. 2012;30(3):343-351. DOI 10.1111/j.1756-1051.2011.01405.x.

Silva F.M.A., Silva F.F.A., Lima B.R., Almeida R.A., Soares E.R., Koolen H.H.F., Souza A.D.L., Pinheiro M.L.B. Chemotaxonomy of the Amazonian *Unonopsis* species based on leaf alkaloid fingerprint direct infusion ESI-MS and chemometric analysis. J. Braz. Chem. Soc. 2016;27:599-604. DOI 10.5935/0103-5053.20150296.

Socaci S.A., Socaciu C., Tofană M., Raţi I., Pinteaa A. In-tube extraction and GC-MS analysis of volatile components from wild and cultivated sea buckthorn (*Hippophae rhamnoides* L. ssp. *Carpatica*) berry varieties and juice. Phytochem. Anal. 2013;24:319-328. DOI 10.1002/pca.2413. Epub 2013 Jan 15.

Villaverde J.J., Domingues R.M.A., Freire C.S.R., Silvestre A.J.D., Pascoal Neto C., Ligero P., Vega A. *Miscanthus* × *giganteus* extractives: a source of valuable phenolic compounds and sterols. J. Agric. Food Chem. 2009;57:3626-3631. DOI 10.1021/jf900071t.

Weißmann G.V., Lange W. Untersuchung der Extraktstoffe von *Miscanthus sinensis* Anderss. Holzforschung. 1991;45(4):285-289.

Winkler-Moser J. Gas Chromatographic Analysis of Plant Sterols. USDA, ARS, NCAUR, Functional Foods Research Unit, 2012. DOI 10.21748/lipidlibrary.40384.

Yi Duan, Jinxian He. Distribution and isotopic composition of n-alkanes from grass, reed and tree leaves along a latitudinal gradient in China. Geochem. J. 2011;45:199-207. DOI 10.2343/geochemj.1.0115.

ORCID ID
S.V. Shekhovtsov orcid.org/0000-0001-5604-5601
S.E. Peltek orcid.org/0000-0002-3524-0456

Благодарности. Работа поддержана бюджетным проектом № АААА-А17-117092070032-4.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 15.11.2019. После доработки 13.12.2019. Принята к публикации 13.12.2019.

Алфавитный указатель авторов статей, опубликованных в журнале в 2019 г.

Абдуллаев К.М. 7, 941
Абдуллаев Р.А. 6, 641
Абугалиева С.И. 7, 887
Агаева Е.В. 7, 910
Агаркова Т.А. 8, 999
Агафонов А.В. 7, 817
Адонина И.Г. 6, 746; 7, 827
Азиева А.М. 2, 184
Айтназаров Р.Б. 8, 999
Акинина В.Н. 1, 86
Акопян А.А. 4, 390
Аксенович Т.И. 8, 1037
Алпатьева Н.В. 6, 641, 753
Амалова А.Ы. 7, 887
Аминева Е.Ю. 1, 15
Амстиславская Т.Г. 4, 390
Амстиславский С.Я. 8, 1006
Андреева Е.Н. 2, 190
Андреенков О.В. 2, 199, 226
Андреенкова Н.Г. 2, 199, 226
Андронов Е.Е. 8, 972
Аниськов Н.И. 6, 780
Антоненко О.В. 2, 140
Антонова О.Ю. 6, 656, 753
Апаликова О.В. 6, 753
Артемьева А.М. 6, 656, 787
Асбаганов С.В. 7, 817
Атишова М.Н. 7, 879
Афонников Д.А. 1, 38, 100; 5, 519

Бабушкина Н.П. 2, 244
Багиров В.А. 3, 355
Бадаева Е.Д. 7, 827
Бажан Н.М. 1, 62
Баженов М.С. 7, 916; 8, 964
Баженова Е.Ю. 1, 55
Байков И.К. 3, 256; 4, 398
Баклушинская И.Ю. 2, 239
Балашов А.В. 7, 941
Баранов С.Г. 4, 496
Баранова О.А. 3, 296
Барило А.А. 1, 75
Баттулин Н.Р. 1, 95
Бебякина И.В. 7, 827
Белозор О.С. 4, 473
Белоногова Н.М. 8, 1037
Белявская В.А. 8, 1026
Беляева Е.С. 2, 148
Беренсен Ф.А. 6, 656
Берсимбаев Р. 5, 594
Беспалова Е.С. 3, 281
Беспалова Л.А. 7, 910, 916
Бехтольд Н.П. 5, 551
Бикчурина Т.И. 3, 355
Бирюков М.Ю. 8, 1059
Блажко Н.В. 3, 262
Блинов А.Г. 8, 1059
Блинова Е.В. 6, 723
Блохина Н.В. 5, 569
Бобохужаев Ш.У. 7, 836
Богачев С.С. 5, 624
Богданова В.С. 8, 972
Боголюбов Д.С. 2, 129
Боголюбова И.О. 2, 129
Бойкова И.В. 3, 328
Болдаков Д.М. 7, 827, 910
Бондарь А.А. 7, 817
Бондарь Н.П. 4, 456
Бородин П.М. 3, 355
Брагин А.О. 8, 1011
Брагина М.К. 1, 38
Брач Н.Б. 6, 650, 717
Брусенцев Е.Ю. 8, 1006
Буглова Л.В. 7, 926
Будаева В.В. 7, 933
Букреева Г.И. 7, 910
Булгакова О. 5, 594
Бульнуар Х. 5, 608
Бурдеева К.В. 4, 456
Бурляева М.О. 6, 667, 708
Бурмакина Н.В. 8, 1076
Бурханова Г.Ф. 7, 856, 865, 873

Вавилин В.А. 4, 390
Васильев Г.В. 8, 972
Васильев С.А. 2, 244
Васильева О.Ю. 7, 926, 933
Вахрамеев А.Б. 3, 343; 8, 993
Веселова С.В. 7, 856, 865
Ветрова С.А. 4, 439
Викторов В.В. 5, 600
Винк М. 2, 226
Вишнякова М.А. 1, 119
Водясова Е.А. 5, 508; 6, 723
Воевода М.И. 8, 1020, 1026
Волков В.А. 6, 730
Волкова Е.И. 2, 199
Волкова Н.А. 3, 355
Волкова Н.Н. 3, 281; 4, 422
Воробьева Н.Е. 2, 212
Воронина Е.Н. 8, 1011
Вышегуров С.Х. 3, 262

Гавриленко Т.А. 3, 281; 4, 422; 6, 753
Гаврилова К.С. 4, 405
Гаджимагомедова М.Х. 6, 746
Гаитова Н.А. 3, 281
Галицын Г.Ю. 8, 1076

Гасич Е.Л. 6, 641
Гатти М. 2, 190
Гацкая С.С. 2, 135
Гвасалия М.В. 8, 958
Генаев М.А. 5, 519
Гениевская Ю.А. 7, 887
Георгиев П.Г. 2, 168, 180
Георгиева Н.А. 8, 981
Георгиева С.Г. 2, 174, 184
Герасимов С.А. 6, 683
Герасимова С.В. 1, 29; 5, 519
Гисматулина Ю.А. 7, 933
Гладких С.В. 6, 765
Гнутиков А.А. 3, 287
Головина М.А. 7, 941
Головнин А.К. 2, 168
Гончарова Е.П. 4, 398
Горшков В.М. 1, 24
Горячковская Т.Н. 8, 1076
Гриффис С. 7, 887
Губарева Н.К. 6, 738
Гудзенко В.Н. 1, 110
Гуреев А.П. 1, 15
Гусар А.С. 7, 926, 933
Гусева Е.Д. 8, 964

Давоян Р.О. 7, 827, 910
Давоян Э.Р. 7, 827, 910
Давыдова Ю.Д. 4, 465; 5, 600
Дауни Г.Ф. 5, 582
Двоеглазов Н.Г. 8, 999
Демаков С.А. 2, 148, 154, 199
Деменков П.С. 3, 312
Дементьева Н.В. 3, 343; 8, 993
Дзюбенко Е.А. 1, 49; 6, 700, 730
Дзюбенко Н.И. 6, 700
Дивашук М.Г. 7, 916
Диковская М.А. 4, 390
Дмитриев А.А. 7, 896
Добрякова К.С. 3, 287
Долгова Е.В. 5, 624
Донченко А.С. 3, 262
Дорогина О.В. 7, 926, 933
Драчкова И.А. 8, 1047
Дрейзигакер С. 7, 803
Дрозд У.С. 4, 448
Дружин А.Е. 3, 296
Дубовец Н.И. 7, 846
Дубовская А.Г. 6, 787
Дудкин Р.В. 7, 926
Дыгало Н.Н. 4, 448
Дьяченко Е.А. 7, 902
Дьячук Т.И. 1, 86

Евдокимова З.З. 6, 753
Евтушенко Е.В. 2, 135
Егорова А.А. 5, 519
Емельянова Л.А. 3, 256
Еникеева Р.Ф. 4, 465; 5, 600
Енина О.Л. 8, 1067
Ермошин А.А. 5, 615
Ефимов В.М. 8, 1032
Ефимов К.В. 8, 1032
Ефремов А.М. 8, 958

Жигалина Д.И. 2, 244
Жимулёв И.Ф. 2, 148, 226
Жмудь Е.В. 7, 926
Журавлев Ю.Н. 3, 362
Жученко А.А. 7, 896

Завьялов Е.Л. 1, 81; 4, 390; 5, 582
Зайцев Б.Н. 3, 337
Захаренко Л.П. 3, 370
Заячковский В.А. 4, 439
Зиновьева Н.А. 5, 527
Зинченко А.Н. 7, 827
Зоркольцева И.В. 8, 1037
Зубаирова У.С. 1, 100
Зубанова Ю.С. 7, 827, 910
Зуева Г.А. 7, 926, 933
Зыков И.Е. 4, 496
Зыкова Т.Ю. 2, 148
Зюте С. 6, 683

Иванисенко В.А. 3, 312
Иванисенко Т.В. 3, 312
Иванкин А.В. 2, 190, 203
Иванова К.А. 5, 519
Иванова Ю.Н. 7, 846
Иванощук Д.Е. 8, 1026
Игнатьева Е.В. 8, 999
Илларионова Н.Б. 1, 81
Ильина Ю.А. 2, 160
Ильичев А.А. 3, 337

Каверина Г.Б. 3, 256
Казанцева А.В. 4, 465; 5, 600
Казанцева О.М. 4, 473
Какабаев А. 5, 594
Камионская А.М. 4, 405
Кантемирова Е.Н. 7, 941
Капустянчик С.Ю. 8, 1076
Карлов Г.И. 7, 916
Карпенко Л.И. 3, 337
Карякин И.В. 2, 226
Каспаров С. 4, 473
Каусбекова А. 5, 594
Кашеварова А.А. 2, 244
Кашинцова А.А. 2, 239
Каштанова Е.В. 8, 1020
Кезимана П. 7, 896
Кибкало И.А. 1, 86
Кимеклис А.К. 8, 972
Кирикович С.С. 5, 624
Кириченко А.В. 8, 1037
Киселева Е.В. 4, 482
Киселева И.С. 5, 615
Клименко Н.С. 6, 753
Кобылянский В.Д. 6, 780
Ковалева В.Ю. 8, 1032
Коваленко Н.Н. 6, 765

Ковтуненко В.Я. 7, 916
Кожевникова Е.Н. 2, 190
Кожемякин В.В. 4, 412
Козарь Е.Г. 4, 439
Козаченко М.Р. 4, 430
Козлов В.А. 8, 1047
Колесникова Т.Д. 2, 140
Коломиец О.Л. 2, 239
Колошина К.А. 5, 519
Колчанов Н.А. 3, 312; 8, 1020, 1047
Комелькова М.В. 5, 582
Комина О.В. 7, 926
Комышев Е.Г. 5, 519
Конев А.Ю. 2, 160
Конопацкая И.Д. 8, 1059
Константинов Ю.М. 7, 809
Концевая Г.В. 5, 582
Конькова Н.Г. 6, 787
Копытова Д.В. 2, 174
Коробицын И.Г. 3, 362
Короленко Т.А. 4, 390
Короткова А.М. 1, 29
Косарева И.А. 6, 723
Косев В.И. 8, 981
Костерин О.Э. 8, 972
Костина Л.И. 6, 753
Костюченко М.В. 2, 168
Кохметова А.М. 7, 879
Кочетов А.В. 3, 312; 5, 519
Кочиева Е.З. 3, 270; 4, 405
Кошкин В.А. 6, 723, 738
Краснобаева И.Л. 3, 328
Краснов Г.С. 7, 896
Красноперова Е.Ю. 6, 691
Крупин П.Ю. 7, 916
Крупина А.Ю. 7, 916
Кручина-Богданов И.В. 7, 941
Крыцына Т.И. 3, 262
Крюков А.А. 3, 287
Крюков А.П. 2, 232
Кубан И.С. 7, 926
Кудрявцева А.В. 7, 896
Кудрявцева Л.П. 7, 896
Кудряшов А.А. 6, 691
Кузнецова Д.Д. 4, 496
Кузнецова К.А. 6, 691
Куйбида Л.В. 8, 1076
Кулешова О.Н. 5, 508
Куликов А.В. 1, 55
Куликова И.В. 3, 362
Кулян Р.В. 1, 24
Кумарбаева М.Т. 7, 879
Купер К.Э. 4, 482
Куркиев К.У. 6, 746
Куршакова М.М. 2, 174
Кусаинова А. 5, 594
Кутузова С.Н. 6, 650

Лапшин В.И. 6, 675
Лапшин М.С. 5, 582
Ларжам Хетраф С. 5, 608
Ларкин Д.М. 5, 559
Ларкина Т.А. 8, 993
Ларссон Я. 2, 148
Лебедев И.Н. 2, 244
Лебедев М.О. 2, 203
Лебедева В.А. 6, 753
Леонова И.Н. 7, 879
Липшин А.Г. 6, 683
Логинова Д.Б. 7, 846
Лопаткина М.Е. 2, 244
Лопухов А.В. 5, 527
Лоскутов И.Г. 6, 723
Лошкарёва С.В. 8, 958
Лутова Л.А. 6, 691

Мазина М.Ю. 2, 212
Мазур А.М. 4, 489
Макарова А.В. 3, 343
Мака рова Е.Н. 1, 62
МакКрекен К.Г. 3, 362
Максименко О.Г. 2, 180
Максимов В.Н. 8, 1026
Максимов И.В. 7, 856, 865, 873
Макунин И.В. 2, 140
Малых С.Б. 4, 465; 5, 600
Малюкова Л.С. 8, 958
Маляровская В.И. 1, 8, 24; 8, 958
Малярчук Б.А. 5, 588
Манухина Е.Б. 5, 582
Матвеев А.Л. 3, 256; 4, 398
Матвеевский С.Н. 2, 239
Матвиенко И.И. 6, 717, 723
Мацухита Ю. 3, 304
Машкина О.С. 1, 15
Медин Бенчекор С. 5, 608
Мельникова Л.С. 2, 168
Мельникова Н.В. 7, 896
Меркулова Т.И. 4, 456
Миков Д.С. 7, 827, 910
Миргалиева Р.Я. 1, 67
Мироненко Н.В. 3, 304
Мирошникова Е.Е. 7, 846
Мирошниченко Е.В. 7, 941
Митрофанова О.В. 3, 343; 8, 993
Митрофанова О.П. 6, 738
Можей О.И. 4, 473
Молодина В.В. 2, 168
Морозов И.В. 7, 817
Морозова К.Н. 4, 482
Мошкин М.П. 5, 582
Мустафин Р.Н. 4, 380, 465; 5, 600

Набирочкина Е.Н. 2, 174
Назарова Л.А. 7, 916
Наумова Л.Г. 6, 772
Нежданова А.В. 4, 405
Некрасов А.Ю. 1, 119
Немкова Г.А. 3, 362
Никитина Т.В. 2, 244
Николаев С.В. 1, 100
Николин В.П. 8, 1020

Новаковский Р.О. 7, 896
Новикова И.И. 3, 328
Новикова Л.Ю. 1, 119; 6, 753, 772
Носов Н.Н. 3, 287
Нуждина Н.С. 7, 926
Нужная Т.В. 7, 856

Овэс Е.В. 3, 281
Огиенко А.А. 2, 190
Омелина Е.С. 2, 219
Орлов П.С. 8, 1026
Орлов Ю.Л. 1, 24
Орлова Е.А. 1, 100; 5, 551
Осадчий И.С. 2, 180
Осипова Л.П. 8, 1011
Осипова Н.А. 8, 999
Останин А.А. 5, 624
Оухаби Джеллоули Х. 5, 608
Ошевский С.И. 8, 1020
Ощепкова Е.А. 8, 1047

Павлов А.В. 6, 650, 717
Павлова Г.А. 2, 190
Паламарчук Д.П. 4, 430
Панкова Т.Е. 2, 154
Пантелеев С.Л. 4, 489
Пантюх К.С. 4, 489
Панченко В.В. 7, 916
Пельтек С.Е. 8, 1076
Перчук И.Н. 1, 119
Петерс Дж.Л. 3, 362
Петровский Д.В. 1, 81; 3, 370; 4, 482
Петрушин И.С. 7, 809
Пиндюрин А.В. 2, 190, 203, 219
Повхова Л.В. 7, 896
Подорожный В.Н. 6, 675
Пожиленкова Е.А. 4, 473
Полонская Я.В. 8, 1020
Полонский В.И. 6, 683
Поминов А.В. 1, 86
Пономарев С.Н. 3, 320
Пономарева М.Л. 3, 320
Пономаренко М.П. 8, 1047
Пономаренко П.М. 8, 1047
Попкова А.М. 8, 1006
Попов В.Н. 1, 15
Попова Н.А. 8, 1020
Попова Ю.В. 2, 190
Пороховинова Е.А. 6, 650, 717
Портнов С.М. 4, 489
Потокина Е.К. 1, 49; 6, 730; 7, 941
Похолкова Г.В. 2, 148
Поцелуев О.М. 8, 1076
Поценковская Э.А. 6, 691
Проворов Н.А. 8, 972
Проскурина А.С. 5, 624
Пунина Е.О. 3, 287
Пупышев А.Б. 4, 390
Пустовой В.Ф. 5, 569
Путина О.В. 6, 641
Пушкова Е.Н. 7, 896
Пшеничникова Т.А. 6, 746

Рагаева Д.С. 8, 1006
Рагино Ю.И. 8, 1020
Радченко Е.Е. 6, 641
Разумов И.А. 1, 81
Раннева С.В. 8, 1006
Рассказов Д.А. 8, 1047
Рахмангулов Р.С. 1, 24
Решетников В.В. 4, 456
Рогозина Е.В. 3, 304; 5, 519
Родионов А.В. 3, 287
Рождественская Г.А. 5, 569
Рожмина Т.А. 7, 896
Романова А.Р. 4, 465
Ромащенко А.В. 4, 482
Ростова Н.С. 6, 708
Рубец В.С. 8, 964
Рудометова Н.Б. 3, 337
Рукин И.В. 4, 489
Румянцев С.Д. 7, 856; 865
Русинов П.Г. 7, 941
Русинова Е.В. 7, 941
Русских Г.С. 4, 390
Рыжкова А.С. 1, 95
Рындин А.В. 1, 24
Рябинина В.А. 3, 262

Савинкова Л.К. 8, 1047
Савченко Р.Р. 2, 244
Сайк О.В. 3, 312
Сайлау Ж.К. 2, 129
Сакович Г.В. 7, 933
Салина Е.А. 1, 38; 5, 542; 7, 827
Самарина Л.С. 1, 8, 24; 8, 958
Самсонова М.Г. 1, 119
Санамьян М.Ф. 7, 836
Сарбаев А.Т. 7, 887
Сарварова Е.Р. 7, 873
Сафонова И.В. 6, 780
Свищёва Г.Р. 8, 1037
Святченко С.И. 4, 430
Сегал Д. 7, 803
Семёнова М.В. 8, 1067
Середин Т.М. 7, 902
Сеферова И.В. 1, 119
Сибикеев С.Н. 3, 296
Сидоренко Д.С. 2, 148
Силкова О.Г. 7, 846
Сингина Г.Н. 5, 527
Сколотнева Е.С. 1, 100; 5, 542
Скрябин К.Г. 4, 405
Скрябин Н.А. 2, 244
Слугина М.А. 3, 270
Слынько Н.М. 8, 1076
Смирнова С.В. 1, 75
Смольникова М.В. 1, 75; 4, 473
Соловей Л.А. 7, 846
Соловьева А.Е. 6, 667

Сорокань А.В. 7, 873
Сорокин И.Е. 1, 55
Сошникова Н.В. 2, 184
Стариков И.Ю. 2, 226
Стахнева Е.М. 8, 1020
Степанов В.А. 4, 439
Степанов И.В. 1, 8
Сулейманов О.И. 5, 569
Сумина А.В. 6, 683
Супрун И.И. 1, 8
Сурин Н.А. 6, 683
Сухих И.С. 8, 1059

Табацкая Т.М. 1, 15
Табиханова Л.Э. 8, 1011
Тамбовцева В.Г. 2, 239
Таранов О.С. 3, 337
Татарский В.В. 2, 184
Творогова В.Е. 6, 691
Теплякова С.Б. 1, 49; 6, 730
Тикунова Н.В. 3, 256; 4, 398
Титова Ю.А. 3, 328
Тихонова М.А. 4, 390
Ткачёв С.Е. 3, 256
Томгорова Е.К. 3, 355
Торгашева А.А. 3, 355
Третьяков А.В. 2, 239
Тугбаева А.С. 5, 615
Туов М.Т. 8, 958
Туруспеков Е.К. 7, 887
Тюрин А.В. 1, 67
Тяпкина Д.Ю. 3, 270

Устьянцев К.В. 8, 1059
Ухатова Ю.В. 3, 281; 4, 422

Федоренко Е.Н. 7, 887
Федорова М.И. 4, 439
Федорова Т.Н. 2, 180
Федорова Ю.А. 6, 691
Фекличева И.В. 5, 582
Феофанова Н.А. 1, 62
Фет В. 8, 1059
Филипенко М.Л. 8, 1011
Филюшин М.А. 4, 405; 7, 902
Фурсенко Д.В. 1, 55

Хабарова А.А. 1, 95
Хайруллин Р.М. 7, 873
Хакимова А.Г. 6, 738
Хаммани-Меджау И. 5, 608
Хатиб А. 4, 489
Хегай И.И. 5, 575
Хлесткина Е.К. 1, 29
Хлусевич Я.А. 4, 398
Хомякова О.В. 1, 86
Хорошко В.А. 2, 148
Хоути Л. 5, 608
Хоцкин Н.В. 1, 55; 5, 582
Хоцкина А.С. 5, 582
Храброва Л.А. 5, 569
Храмцов В.В. 8, 999
Хусаинова Р.И. 1, 67
Хуснутдинова Э.К. 4, 380, 465; 5, 600

Цветова М.И. 4, 412
Цейликман В.Э. 5, 582
Цейликман О.Б. 5, 582

Чадаева И.В. 8, 1047
Чалая Н.А. 3, 304; 5, 519
Челебиева Э.С. 5, 508
Черепанова Е.А. 7, 873
Черноок А.Г. 7, 916
Черных Е.Р. 5, 624
Чуйко Э.А. 8, 1006

Шабанова (Кобозева) Е.В. 7, 817
Шабурова Е.В. 4, 448
Шаварда А.Л. 1, 49; 6, 667
Шарапова М.Б. 4, 482
Шарыпова Е.Б. 8, 1047
Шатохин К.С. 3, 262
Шаухаров Р.А. 7, 941
Шацкая Н.В. 8, 972, 1059
Шварц (Беркаева) М.Б. 2, 154
Швачко Н.А. 4, 422
Шевелев О.Б. 5, 582
Шейнов А.А. 2, 184
Шеленга Т.В. 1, 49, 119; 6, 667
Шелепова О.В. 8, 1067
Шеховцов С.В. 8, 1076
Шинтяпина А.Б. 4, 390
Шнеер В.С. 3, 287
Шуваев А.Н. 4, 473
Шуваев Андр.Н. 4, 473
Шумный В.К. 8, 1076

Щеглов С.Н. 6, 675
Щенникова А.В. 4, 405
Щербакова Н.С. 3, 337
Щербань А.Б. 5, 534
Щукина Л.В. 6, 746

Эльконин Л.А. 4, 412

Юдин Н.С. 5, 559; 8, 999, 1026
Юшкова А.А. 2, 190

Яковенко В.В. 6, 675
Яковлева Д.А. 4, 473
Яковлева Т.В. 1, 62
Ялаев Б.И. 1, 67
Янагисава Х. 3, 304
Яринич Л.А. 2, 203