

ВАВИЛОВСКИЙ

ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

VAVILOV JOURNAL OF GENETICS AND BREEDING

2019 • 23 • 5

Актуальные технологии / Селекция растений на иммунитет и продуктивность / Генетика животных /
Физиологическая генетика / Популяционная генетика / Медицинская генетика /
Биоинформатика и клеточная биология



Буряя ржавчина пшеницы

vavilov.ejpub.ru/
bioenet.nsc.ru/vogjs
vavilov_journal@bioenet.nsc.ru
Индекс издания 42153



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. Афтanas – 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. Rogozin – 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. Турусбеков – канд. биол. наук (Казахстан)
М. Чен – д-р биол. наук (Китайская Народная Республика)
Ю. Шавруков – д-р биол. наук (Австралия)

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. Beshpalova, 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

507

ОТ РЕДАКТОРА

Актуальные технологии

508

ОБЗОР

Новейшие технологии высоко-производительного секвенирования транскриптома отдельных клеток.

Е.А. Водясова, Э.С. Челебиева, О.Н. Кулешова

519

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Оценка количественных характеристик клубнеобразования дикого картофеля на основе анализа изображений клубней с использованием компьютерного приложения SeedCounter. К.А. Иванова, Е.Г. Комышев, М.А. Генаев, А.А. Егорова, К.А. Колошина, Н.А. Чалая, Д.А. Афонников, А.В. Кочетов, Е.В. Рогозина, С.В. Герасимова

527

ОБЗОР

Биотехнологические основы получения клонированных эмбрионов свиней.

А.В. Лопухов, Г.Н. Сингина, Н.А. Зиновьева

Селекция растений на иммунитет и продуктивность

534

ОБЗОР

Перспективы маркер-ориентированной селекции томата *Solanum lycopersicum* L. А.Б. Щербань

542

ОБЗОР

Разнообразие механизмов устойчивости, вовлеченных в многоуровневый иммунитет пшеницы к ржавчинным заболеваниям. Е.С. Сколотнева, Е.А. Салина

551

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Характеристика генофонда яровой мягкой пшеницы (*Triticum aestivum* L.) по устойчивости к пыльной головне в условиях лесостепи Западной Сибири. Е.А. Орлова, Н.П. Бехтольд

Генетика животных

559

ОБЗОР

Происхождение, селекция и адаптация российских пород крупного рогатого скота по данным полногеномных исследований. Н.С. Юдин, Д.М. Ларкин

569

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Оценка дифференциации линий в чистокровной верховой породе лошадей с использованием микросателлитов ДНК. Л.А. Храброва, Н.В. Блохина, О.И. Сулейманов, Г.А. Рождественская, В.Ф. Пустовой (на англ. языке)

Физиологическая генетика

575

ОБЗОР

Неканонические эффекты вазопрессина в ангиогенезе. И.И. Хегай

582

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Уровень тревожности и содержание нейрометаболитов в гиппокампе и амигдале крыс после завершения хронического предаторного стресса. О.Б. Шевелев, В.Э. Цейликман, Н.В. Хоцкин, А.С. Хоцкина, Г.В. Концевая, М.С. Лапшин, М.П. Мошкин, М.В. Комелькова, И.В. Фекличева, О.Б. Цейликман, Е.Б. Манухина, Г.Ф. Дауни, Е.Л. Завьялов

Популяционная генетика

588

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Истоки митохондриального генофонда русских по результатам анализа современных и палеогеномных данных. Б.А. Мальярчук

594

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Ассоциация полиморфизма *TP53 Arg72Pro* с риском развития радон-индуцированного рака легкого в казахской популяции. О. Булгакова, А. Кусаинова, А. Какабаев, А. Каусбекова, Р. Берсимбаев (на англ. языке)

Медицинская генетика

600

ОБЗОР

Эпигенетика суицидального поведения. Р.Н. Мустафин, А.В. Казанцева, Р.Ф. Еникеева, Ю.Д. Давыдова, С.Б. Малых, В.В. Викторов, Э.К. Хуснутдинова

608

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Связь генетических вариантов *rs439401* и *rs4420638* в хромосоме 19 с ожирением и артериальной гипертензией у жителей Алжира. Х. Бульнуар, С. Медин Бенчекор, Х. Оухаби Джеллоули, С. Ларжам Хетраф, Л. Хоути, И. Хаммани-Меджау (на англ. языке)

Биоинформатика и клеточная биология

615

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Предсказание функций некоторых пероксидаз *Arabidopsis thaliana* L. на основе биоинформатического поиска. А.С. Тугбаева, А.А. Ермошин, И.С. Киселева

624

ОБЗОР

Тернистый путь макрофаг-активирующего фактора (GcMAF): от открытия к клинической практике. А.А. Останин, С.С. Кирикович, Е.В. Долгова, А.С. Проскурина, Е.Р. Черных, С.С. Богачев

507 FROM THE EDITOR

Mainstream technologies

- 508 REVIEW
The new technologies of high-throughput single-cell RNA sequencing. *E.A. Vodiasova, E.S. Chelebieva, O.N. Kuleshova*

- 519 ORIGINAL ARTICLE
Image-based analysis of quantitative morphological characteristics of wild potato tubers using the desktop application SeedCounter. *K.A. Ivanova, E.G. Komyshev, M.A. Genaev, A.A. Egorova, K.A. Koloshina, N.A. Chalaya, D.A. Afonnikov, A.V. Kochetov, E.V. Rogozina, S.V. Gerasimova*

- 527 REVIEW
Biotechnological bases of the development of cloned pig embryos. *A.V. Lopukhov, G.N. Singina, N.A. Zinovieva*

Plant breeding for immunity and performance

- 534 REVIEW
Prospects for marker-associated selection in tomato *Solanum lycopersicum* L. *A.B. Shcherban*

- 542 REVIEW
Resistance mechanisms involved in complex immunity of wheat against rust diseases. *E.S. Skolotneva, E.A. Salina*

- 551 ORIGINAL ARTICLE
Characteristics of the gene pool of spring wheat (*Triticum aestivum* L.) for resistance to loose smut in the forest-steppe of Western Siberia. *E.A. Orlova, N.P. Baechtold*

Animal genetics

- 559 REVIEW
Whole genome studies of origin, selection and adaptation of the Russian cattle breeds. *N.S. Yudin, D.M. Larkin*

- 569 ORIGINAL ARTICLE
Assessment of line differentiation in the Thoroughbred horse breed using DNA microsatellite loci. *L.A. Khrabrova, N.V. Blohina, O.I. Suleymanov, G.A. Rozhdestvenskaya, V.F. Pustovoy*

Physiological genetics

- 575 REVIEW
Noncanonical effects of vasopressin in angiogenesis. *I.I. Khegay*

- 582 ORIGINAL ARTICLE
Anxiety and neurometabolite levels in the hippocampus and amygdala after prolonged exposure to predator-scent stress. *O.B. Shevelev, V.E. Tseilikman, N.V. Khotskin, A.S. Khotskina, G.V. Kontsevaya, M.S. Lapshin, M.P. Moshkin, M.V. Komelkova, I.V. Feklicheva, O.B. Tseilikman, E.B. Manukhina, H.F. Downey, E.L. Zavjalov*

Population genetics

- 588 ORIGINAL ARTICLE
Sources of the mitochondrial gene pool of Russians by the results of analysis of modern and paleogenomic data. *B.A. Malyarchuk*

- 594 ORIGINAL ARTICLE
Association of polymorphism *TP53* Arg72Pro with radon-induced lung cancer in the Kazakh population. *O. Bulgakova, A. Kussainova, A. Kakabayev, A. Kausbekova, R. Bersimbaev*

Medical genetics

- 600 REVIEW
Epigenetics of suicidal behavior. *R.N. Mustafin, A.V. Kazantseva, R.F. Enikeeva, Yu.D. Davydova, S.B. Malykh, V.V. Viktorov, E.K. Khusnutdinova*

- 608 ORIGINAL ARTICLE
Genetic variants on chromosome 19 (rs439401 and rs4420638) are associated with obesity and high blood pressure in the Algerian population. *H. Boulouvar, S. Mediene Benchechor, H. Ouhaibi Djellouli, S. Larjam Hetraf, L. Houti, I. Hammani-Medjaoui*

Bioinformatics and cell biology

- 615 ORIGINAL ARTICLE
Prediction of some peroxidase functions in *Arabidopsis thaliana* L. by bioinformatic search. *A.S. Tugbaeva, A.A. Ermoshin, I.S. Kiseleva*

- 624 REVIEW
A thorny pathway of macrophage activating factor (GcMAF): from bench to bedside. *A.A. Ostanin, S.S. Kirikovich, E.V. Dolgova, A.S. Proskurina, E.R. Chernykh, S.S. Bogachev*

Уважаемые читатели! Предлагаем вашему вниманию очередной выпуск журнала. Его открывает рубрика «Актуальные технологии», в которой представлены результаты использования современных методических подходов для решения задач клеточной биологии, генетики растений и животных. Обзорная статья рассказывает о технологии single-cell, позволяющей работать с единичными клетками. Описаны платформы для проведения высокопроизводительного секвенирования транскриптомов клеток, этапы их получения и анализа. В следующем обзоре рассмотрены технологические особенности современных методов, используемых для соматического клонирования эмбрионов свиней. Оригинальное исследование содержит результаты анализа параметров клубнеобразования диких видов картофеля, полученных с помощью цифрового фенотипирования – нового методического решения для сбора и анализа морфологических признаков.

Рубрика «Селекция растений на иммунитет и продуктивность» включает два обзора и оригинальную статью. Первый обзор посвящен направлениям селекции томата *Solanum lycopersicum*, в том числе для Сибирского региона. Основной акцент сделан на применении ДНК-технологий для селекции этой культуры на устойчивость к заболеваниям и вкусовые качества. Авторы второго обзора на примере злаковых обсуждают современную концепцию системы защиты растений от фитопатогенов. Оригинальное исследование представляет интерес прежде всего для специалистов-фитопатологов; в нем приведены результаты многолетнего изучения генофонда яровой пшеницы различного эколого-географического происхождения по устойчивости к пыльной головне.

В разделе «Генетика животных» рассмотрены генетические аспекты селекции крупного рогатого скота и лошадей. Литератур-

ный обзор обобщает данные по истории российских пород крупного рогатого скота с акцентом на генетические факторы, способствующие адаптации животных к климатическим условиям. Особое внимание уделено современным биотехнологическим методам, в том числе полногеномному анализу ассоциаций для выбора генетических маркеров для селекции особей по хозяйственно ценным признакам. В оригинальное исследование включены результаты изучения дифференциации российских чистокровных верховых лошадей с использованием аллель-специфичных микросателлитных локусов. Авторы выявили аллели, типичные для отечественной породы, и сделали вывод об эффективности их использования для контроля происхождения лошадей.

Обзорная статья рубрики «Физиологическая генетика» знакомит с механизмами действия и регуляцией активности вазопрессина как в норме, так и при патологических процессах. В экспериментальной работе обсуждаются вопросы проявления поведенческих реакций и накопления нейрометаболитов в различных отделах головного мозга животных в качестве ответной реакции на стрессовый фактор.

Разделы «Популяционная генетика» и «Медицинская генетика» представлены тремя экспериментальными исследованиями и обзором литературных данных по изучению эпигенетических механизмов суицидального поведения. В одной из статей приведены интересные результаты изучения формирования генофонда русских, полученные с привлечением палеогеномных данных и анализа митохондриальной ДНК. Другая работа анализирует корреляции между риском развития рака легкого в казахской популяции и полиморфизмом гена-супрессора опухоли *TP53*. Эта статья перекликается с оригинальным исследованием по изучению полиморфизмов гена *APOE* среди населения Алжира и их связи с такими метаболическими аномалиями, как ожирение и высокое артериальное давление.

Последняя рубрика «Биоинформатика и клеточная биология» знакомит читателей с оригинальной работой, в которой выполнен биоинформатический анализ структурно-функциональной организации пероксидаз у модельного объекта *Arabidopsis thaliana*. В обзорной статье этой рубрики рассмотрены механизмы противоопухолевого эффекта макрофаг-активирующего фактора GcMAF с научной и клинической точек зрения.

Академик В.К. Шумный

The new technologies of high-throughput single-cell RNA sequencing

E.A. Vodiasova , E.S. Chelebieva, O.N. Kuleshova

A.O. Kovalevsky Institute of Biology of the Southern Seas, RAS, Moscow, Russia
 e-mail: eavodiasova@gmail.com

A wealth of genome and transcriptome data obtained using new generation sequencing (NGS) technologies for whole organisms could not answer many questions in oncology, immunology, physiology, neurobiology, zoology and other fields of science and medicine. Since the cell is the basis for the living of all unicellular and multicellular organisms, it is necessary to study the biological processes at its level. This understanding gave impetus to the development of a new direction – the creation of technologies that allow working with individual cells (single-cell technology). The rapid development of not only instruments, but also various advanced protocols for working with single cells is due to the relevance of these studies in many fields of science and medicine. Studying the features of various stages of ontogenesis, identifying patterns of cell differentiation and subsequent tissue development, conducting genomic and transcriptome analyses in various areas of medicine (especially in demand in immunology and oncology), identifying cell types and states, patterns of biochemical and physiological processes using single cell technologies, allows the comprehensive research to be conducted at a new level. The first RNA-sequencing technologies of individual cell transcriptomes (scRNA-seq) captured no more than one hundred cells at a time, which was insufficient due to the detection of high cell heterogeneity, existence of the minor cell types (which were not detected by morphology) and complex regulatory pathways. The unique techniques for isolating, capturing and sequencing transcripts of tens of thousands of cells at a time are evolving now. However, new technologies have certain differences both at the sample preparation stage and during the bioinformatics analysis. In the paper we consider the most effective methods of multiple parallel scRNA-seq using the example of 10XGenomics, as well as the specifics of such an experiment, further bioinformatics analysis of the data, future outlook and applications of new high-performance technologies.

Key words: scRNA-seq; transcriptomics; Chromium 10XGenomics; sequencing; single cell.

For citation: Vodiasova E.A., Chelebieva E.S., Kuleshova O.N. The new technologies of high-throughput single-cell RNA sequencing. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5):508-518. DOI 10.18699/VJ19.520

Новейшие технологии высокопроизводительного секвенирования транскриптома отдельных клеток

Е.А. Водясова , Э.С. Челебиева, О.Н. Кулешова

Институт биологии южных морей имени А.О. Ковалевского Российской академии наук, Москва, Россия
 e-mail: eavodiasova@gmail.com

Огромное количество полногеномных и транскриптомных данных, полученных с помощью современных технологий секвенирования нового поколения для целых организмов, не смогло дать ответы на многие вопросы в онкологии, иммунологии, физиологии, нейробиологии, зоологии и других областях науки и медицины. Так как основой всех одноклеточных и многоклеточных организмов является клетка, то необходимо изучение биологических процессов на ее уровне. Это понимание дало толчок развитию нового направления и появлению технологий, позволяющих работать с единичными клетками (технологии single-cell). Быстрое развитие не только приборной базы, но и различных усовершенствованных протоколов для работы с единичными клетками обусловлено актуальностью этих исследований во многих областях науки и медицины. Изучение особенностей различных этапов онтогенеза, определение закономерностей дифференциации клеток и последующего развития тканей, проведение геномного и транскриптомного анализов в различных областях медицины (особенно востребовано в иммунологии, онкологии), классификация типов и состояний клеток, закономерностей биохимических и физиологических процессов с применением технологий single-cell позволяют проводить комплексные исследования на новом уровне. Разработанные первые платформы для осуществления секвенирования транскриптомов отдельных клеток (scRNA-seq) проводили изоляцию не более ста клеток одновременно, что оказалось недостаточным в связи с выявленной высокой гетерогенностью клеток, обнаруженными минорными типами клеток, которые не детектировались по морфологическим признакам, и сложными регуляторными путями в организме. В настоящее время появились методики изоляции, захвата и секвенирования транс-

криптомов (scRNA-seq) десятков тысяч клеток одновременно. Однако новые технологии имеют определенные отличия как на этапе пробоподготовки, так и во время проведения биоинформатического анализа. В работе рассмотрены наиболее эффективные методы множественного параллельного scRNA-seq на примере современной платформы для изоляции и баркодирования клеток 10XGenomics, а также особенности проведения такого эксперимента, дальнейший биоинформатический анализ полученных данных, перспективы использования и области применения новых высокопроизводительных технологий.

Ключевые слова: scRNA-seq; транскриптомика; Chromium 10XGenomics; секвенирование; единичные клетки.

Introduction

The improvements in new generation sequencing technology in conjunction with the technology of multiple parallel capture and analysis of single cells raised to a new level of research in many fields of biology, biotechnology and medicine (Junker, Oudenaarden, 2014). A brief history of the development of single-cell biology is presented in the Fig. 1. The main areas of the application of innovative methods are neurobiology, immunology, embryology and oncology. The first experiments on the analysis of the transcriptome of single neurons were carried out using microarrays in 2003. Isolation of neurons were captured using laser capture microdissection instrument and data analysis revealed the unexpectedly high cellular heterogeneity (Kamme et al., 2003). The sequencing of the neuron transcriptome began to actively develop (Moroz et al., 2006; Tang et al., 2009) and at the same time microfluidic technology that allow the analysis of up to 100 cells simultaneously appeared (Marcus et al., 2006). In parallel convergence of qPCR and single-cell methods are implemented (Subkhankulova et al.,

2008). The first whole genomic sequencing of tumor cells identified unexpectedly abundant subpopulation of ‘pseudo-diploid’ cells which do not metastasize (Navin et al., 2011). In the past five years, a new trend of “epigenomics” has been developed, connected with the study of various RNA and DNA modifications using sequencing. Modern single-cell technology allows to characterize the methylation, chromosome organization, intra- and interdomain contacts, chromatin mobility, specific sites of histone modifications in single-cell resolution (Goldberg et al., 2007; Nagano et al., 2013; Rotem et al., 2015).

The need for an analysis of a huge number of cells in parallel has led to the development of high-throughput technologies aimed at increasing the sensitivity and accuracy of sequencing by using different protocols and developing platforms for capturing single cells with high efficiency. From 2003 to 2016, the number of captured cells simultaneously increased from a few to hundreds of thousands.

Analysis of the transcriptome, proteome and epigenome at the single-cell resolution allows to understand the formation,

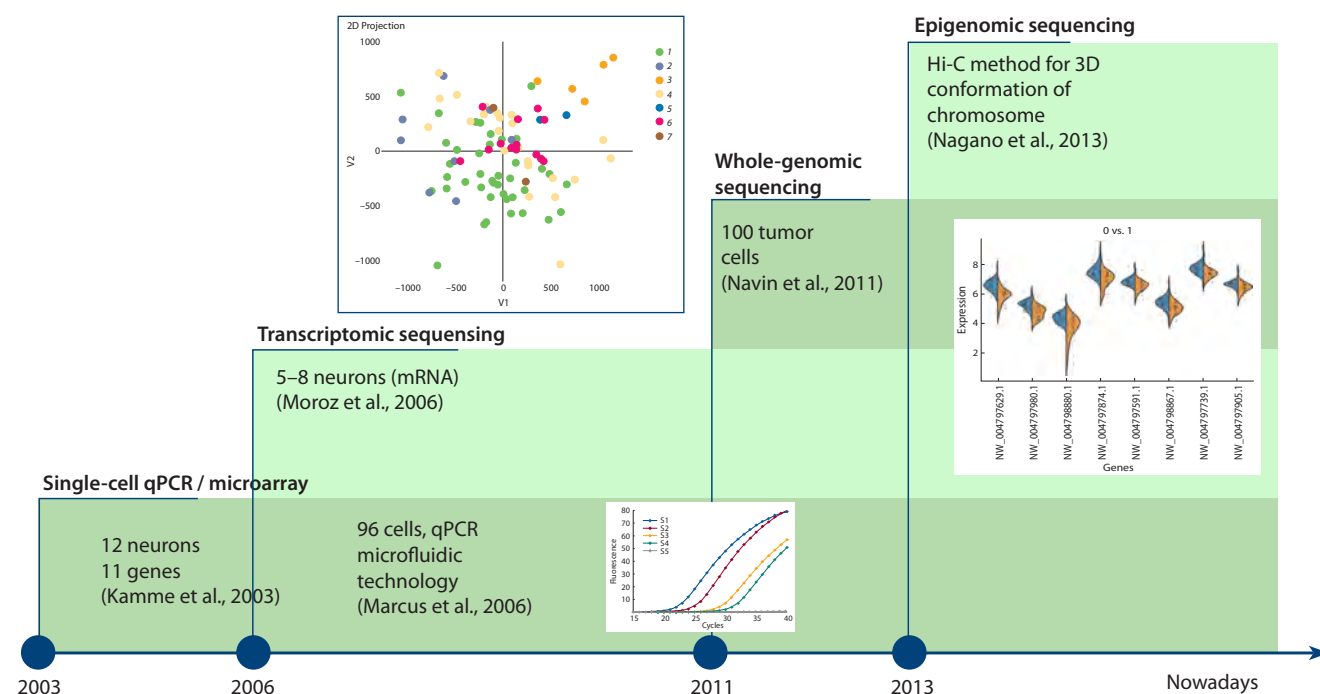


Fig. 1. The overview of the development of single-cell technologies.

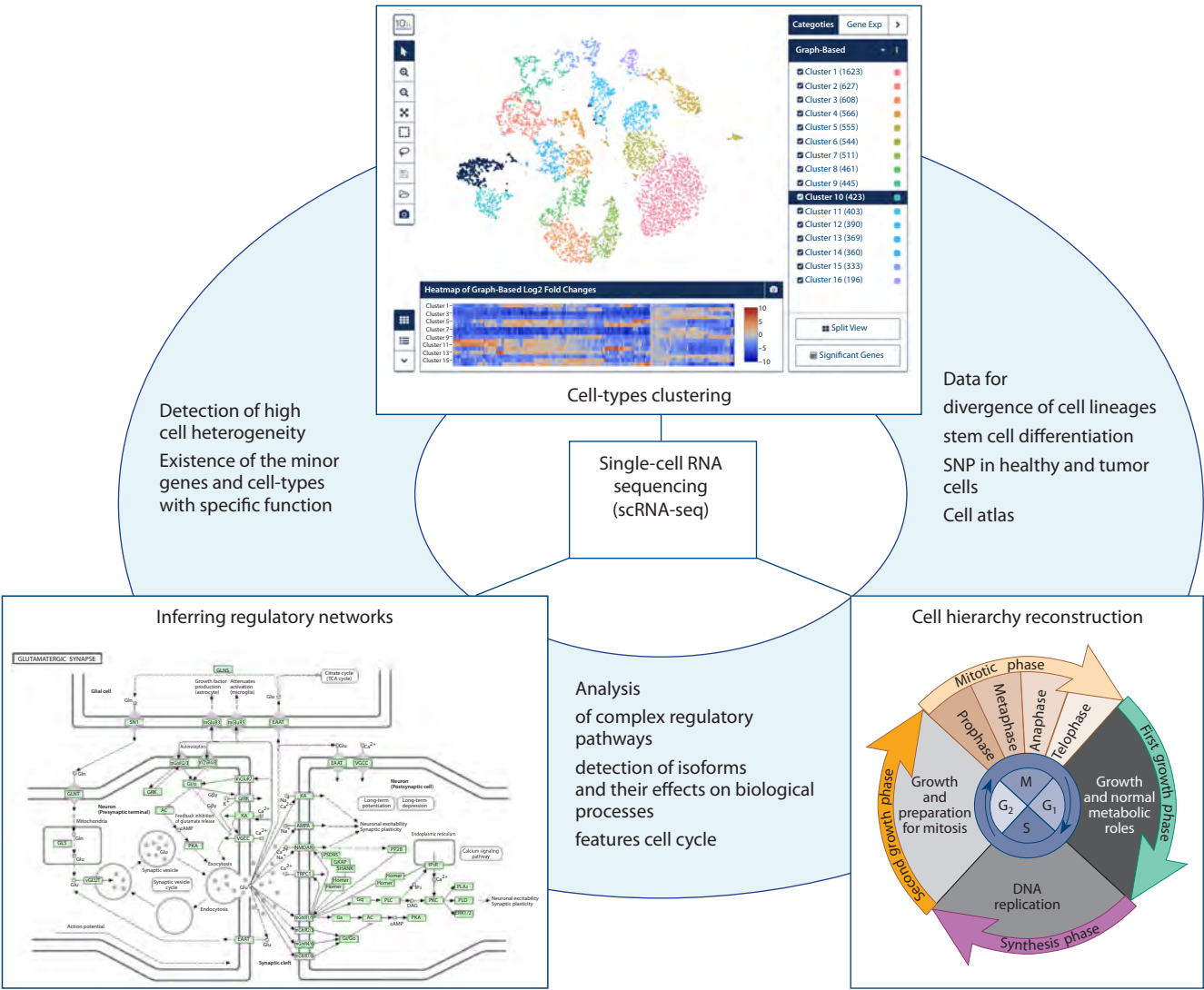


Fig. 2. The main applications of scRNA-seq data analyses.

development, organization and interaction of various cells and tissues, explore the problems of stress and adaptations (Frieda et al., 2017), identify the types and states of cells, establish patterns of their differentiation in the ontogeny (Nowogrodzki, 2017; Mi et al., 2018), study various diseases in immunology, oncology, neurobiology and other fields of medicine (Leung et al., 2017), answer questions of evolution, speciation and formation of global biodiversity (Moroz, 2018).

In this article we briefly review some of the latest and most efficient commercial available platforms for single cell collection. The capturing and specific protocols for single-cell RNA-sequencing (scRNA-seq) as well as the bioinformatics analysis and application of single-cell technologies are considered.

Applications of single-cell RNA-seq technologies

The main areas and approaches for scRNA-seq data in various fields of biology and medicine are presented in the Fig. 2. Bioinformatics pipelines for scRNA-seq data allow

to analyze different regulatory networks based on gene expression of different genes for all single cell (Janes et al., 2010; Shalek et al., 2013; Trapnel et al., 2014; Treutlein et al., 2014). Also, the new single-cell technology focuses in identification of cell-types and cell-stages (Jaitin et al., 2014), revealing rare genes and cell populations, which can be discarded from analysis of bulk transcriptome as minor fractions (Gerber et al., 2016), point mutations (Gawad et al., 2016; Ludwig et al., 2019).

Today, due to the expression profiles of thousands of genes in hundreds of thousands of individual cells, unique cell-types and cell-states have been identified, new signaling pathways have been uncovered (Saliba et al., 2014; Grun et al., 2015; Okaty et al., 2015; Zeisel et al., 2015; Poulin et al., 2016; Tirosh et al., 2016; Callaway, 2017; Lavin et al., 2017). The whole-organism tissue maps at the single-cell level are now constructed (Segal et al., 2004; Pijuan-Sala et al., 2019; Taylor et al., 2019), cell lineages reveal high heterogeneity (Mahata et al., 2014; Wang, Song, 2017).

The combination of various molecular methods with

scRNA-seq technology now are leading the experiments to the fundamentally new level. So, to study the cell-states and cells differentiation during ontogenesis, it is possible to use the CRISPR-Cas9 system for barcoding by the introduction of specific mutations into the genome or various fluorescent labels (Adamson et al., 2016; Jaitin et al., 2016; McKenna et al., 2016; Kalhor et al., 2017). It has been shown that long non-coding RNAs, even in several copies in a cell, may have important regulatory functions (Derrien et al., 2012). The capturing of thousands of cells and high-throughput sequencing with the detection of chromosome organization in the three-dimensional space (Hi-C) (Belton et al., 2012) shown, that chromosomal compartments, contact insulation, topological-associated domains (TADs) or long-range loops are governed by distinct cell-cycle dynamics (Nagano et al., 2017; Stevens et al., 2017).

Platform for single-cell collection

The essential requirement in studying the transcriptome of individual cells is the dissociation and capture of each cell of high quality and preparation for sequencing (barcoding of all transcripts for each cell). In order to pool, amplify and sequence cDNA in a single run and then assign each read to its original cell at the bioinformatics analysis step, it is necessary to incorporate a unique cellular identifier. Today, single-cell isolation platforms are represented by commercial models that differ in performance: Puncher Platform Vycap, CellRaft AIR System, PEPAArray System, Fluidigm C1, Wafergen ICELL8, BioRad Illumina ddSEQ, Dolomite Bio Nadia и RNA-Seq System, Tapestry Platform MissionBio, 1CellBio InDrop, BD Rhapsody, Chromium 10XGenomics (Kolodziejczyk et al., 2015; Valihrach et al., 2018).

Single-cell isolation approaches can be divided into low-throughput (capture from 10 to 100 cells) and high-throughput (capture from 100 to >10,000 cells) (Wang, Navin, 2015; Poulin et al., 2016). The first group includes serial dilution (Ham, 1965), mechanical micromanipulation (Brehm-Stecher et al., 2004), micropipetting, laser capture microdissection (LCM); the second group: fluorescence activated cell sorting (FACS), microfluidics, microdroplets (Navin et al., 2011; Landry et al., 2013; Mazutis et al., 2013). We consider here highly efficient cell capture platforms them in more detail as the most popular. FACS is method based on the size, granularity and fluorescent properties of cells, which allows to isolate hundreds of thousands of cells per minute, so it is as efficient as economical. However, the use of fluorescent dye can negatively affect cell viability (Lindström et al., 2010).

The methods for cell isolation based on microfluidics and microdroplets technologies are actively developing and used in the analysis of not only the transcriptome, but also the whole-genome of individual cells, as well as epigenomics sequencing (Zheng et al., 2017). These integrated systems allow to manipulate small volumes (10^{-9} – 10^{-18} l), isolate DNA and output accurate results with high resolution (Whitesides et al., 2006; Salafi et al., 2016). Today there are a lot

of high-throughput commercial platforms for cell isolation based on microfluidics technology, which not only capture cells but also provide further biochemical reactions (Poulin et al., 2016; Valihrach et al., 2018).

A droplet-based system became breakthrough in single-cell RNA sequencing and now we can label all mRNA molecules with barcode sequence shared for the same cell. It gives the ability to analyze huge number of cells in parallel and check the cell origination for each mRNA. Such approach is implemented in the next devices: Drop-Seq (Macosco et al., 2015), InDrop (Klein et al., 2015) and Chromium 10XGenomics (Kolodziejczyk et al., 2015). It is based on technology that separates high molecular weight DNA fragments or whole single cells into emulsifications containing beads with attached adapters and unique barcodes (Coombe, 2016). As a result, we obtain a suspension of gel beads (GEM – Gel bead in EMulsion), in each of which there is one cell and all the necessary reagents for lysis, barcoding, reverse transcription and cDNA synthesis. The main advantage of these platforms is the high efficiency of the capture and preparation of single cells for further sequencing. One of the leaders in this area is Chromium 10XGenomics, which allows you to simultaneously capture up to 80,000 + cells, while the number of cells analyzed from other high-performance platforms ranges from 10,000 to 48,000 cells (Valihrach et al., 2018).

Single-cell mRNA sequencing with high-throughput platforms

The study of single cell transcriptome consists of experimental and bioinformatics parts (Kumar et al., 2017; Li et al., 2017). Before starting an experiment, it is necessary to have a reference genome or bulk transcriptome for mapping of scRNA-seq data (Gawad et al., 2014). As mentioned above, the most productive platform is Chromium™ 10XGenomics. The features of the scRNA-seq experiment based on it in comparison with other methods for studying individual cells are discussed below.

Cell isolation and scRNA methods

The main challenge of sample preparation is to obtain a suspension of viable non-aggregated cells (cell dissociation). It is necessary to determine the concentration of cells, since a successful capture requires a certain number of them depending on the chosen method of cell isolation: from ten cells (micropipetting, cytoplasmic aspiration, laser microdissection) to thousands of cells (devices based on FACS technology, microfluidics and microdroplets) (Der et al., 2017). In the case of working with high-throughput devices (for example, Chromium 10XGenomics), the cell concentration should be about 1,000,000 cells/ml (Der et al., 2017). When working with animals or tissues containing the number of cells less than the required, you need to increase the number of specimen per sample. All stages of cell dissociation are carried out in a minimum volume of solution (from 50 to 1000 µl) to increase the concentration and reduce cell loss.

Single-cell RNA sequencing protocol may be divided into three stages: reverse transcription, cDNA amplification (WTA – full-transcriptome amplification) and library preparation. Despite the undesirability of cDNA amplification (due to the possibility of polymerase errors or the loss of rare transcripts), this step is necessary to construct a cDNA library, because the amount of total RNA in the cell is about 10 pg, which is not enough for single cell transcriptomic sequencing (Wang, Song, 2017). Protocols for reverse transcription and cDNA amplification depends on the task and platform for cell isolation (Haque et al., 2017; Kumar et al., 2017; Ziegenhain et al., 2017).

Today, three approaches are known. The first is oligo-dT-anchor approach (or Tang's method). It uses oligo-dT-primers conjugated to adapter sequence for reverse transcription and selective amplification of polyadenylated mRNA by PCR (Tang et al., 2009). However, this protocol leads to missing proximal splicing event due to the generation of 3'-end skew bias during reverse-transcription.

Later, the approach that allows to construct a full-length cDNA was developed – it is called template switching cDNA synthesis (Ramsköld et al., 2012). The main benefit is in ability to detect alternatively spliced exons and allele-specific expression (ASE) (Kolodziejczyk et al., 2015). This approach is used in such protocols as STRT (Islam et al., 2014), SMART-seq и SMART-seq2 (Ramsköld et al., 2012; Deng et al., 2014). The key features of these protocols is the exponential increase in the number of transcripts during the amplification, that leads to bias and loosing of minor expressed genes. *In vitro* transcription (IVT) is the alternative approach with linear amplification which incorporated to a such protocols as CEL-Seq (Hashimshony et al., 2012) and MARS-Seq (Jaitin et al., 2014).

Finally, the third approach based on using a 6–10 bp unique molecular identifier (UMI) with anchored oligo-dT- primer for barcoding each molecule. This labeling technology is presented in CEL-Seq (Hashimshony et al., 2012) and CEL-Seq2 (Hashimshony et al., 2016), Drop-Seq (Macosco et al., 2015), MARS-Seq (Jaitin et al., 2014), SCRIB-seq (Soumillon et al., 2014), STRT (Islam et al., 2014), In-Drop (Klein et al., 2015), Quartz-Seq2 (Sassagawa et al., 2018), the one of the latest protocols with UMI, allows to pool cDNA of up to 1536 cells into one mixture and the UMI conversion efficiency is 35 % in contrast with 22 % for the other scRNA-seq method. It gives the opportunity to effectively convert initial reads to UMI counts and detect more genes.

The modern tools for isolation and capturing hundreds of thousands of cells require an improvement in the labeling of mRNA molecules. The most innovative barcoding strategy in microfluidic and microdroplet technology is the use of as UMI as an additional cellular barcode placed in each drop with individual cells and other reagents. The cellular barcode (about 14 bp oligonucleotide) is an identifier of all nucleotide sequences from different cell. The advantage of such double barcoding is the high accuracy and the ability to assign each mRNA to its original cell (Islam et

al., 2014). The sequencing protocol STAMPs (Single-cell Transcriptomes Attached to Microparticles) was designed for platform Drop-Seq and protocol Cell-Seq for platform InDrop (Wang, Song, 2017).

The most high-throughput commercial platform for cell isolation Chromium 10XGenomics integrated the new barcoding technology GemCode. After isolation of cells, each long molecule is labeled with oligonucleotide, which consists of sequencing adapters and primers, a unique molecular identifier with anchored oligo-dT and a barcode to index GEMs, before library preparation (Eisenstein, 2015; Coombe et al., 2016). The double barcoding leads to reducing technical noise during the analyzing of thousands of different cells in parallel. It is especially important when working with complex tissues on a single-cell resolution and you need to get detailed information about cell-to-cell variability. This makes it possible to determine gene expression profiles, alternative splicing and cellular heterogeneity (Zheng et al., 2017).

Also, it is known various adapters for small RNA sequencing library preparation: for example, a pre-adenylated 3'-end adapter containing a 5',5'-adenyl pyrophosphoryl moiety (Hafner et al., 2008; Chen et al., 2012).

cDNA preparation is followed by sequencing. The choice of sequencing technology should be directly dependent on the goals. The different lengths of the reads (from fifty to several thousand nucleotides), accuracy, efficiency should be considered before choosing a sequencing method (Liu et al., 2012).

Computational analyses

The development of high-throughput cell isolation platforms, capturing hundreds of thousands of cells, requires new bioinformatics approaches to analyze so massive and complex single-cell transcriptomic data sets. It should be noted that in the analysis of scRNA-seq it is not always possible to use traditional computational tools for the bulk-transcriptome. In the sections below, we discuss the challenges single-cell transcriptomics and possible solutions.

A single scRNA-seq experiment is able to provide information about the profiling of all RNAs present in each of hundreds of thousands of cells, which leads to generation of massive volume of data and technical difficulties in analyzing them. The alignment and counting of reads is performed independently per cell, which requires more challenging parallelism and, accordingly, more computational power (Tokunaga et al., 2014; Yu, Lin, 2016).

scRNA-seq allows you to analyze the expression profiles for each individual cell. This led to the emergence of three main areas of application single-cell transcriptomics (see Fig. 2): identification of cell populations (cell-types and cell-states), the cell hierarchy reconstruction (cell differentiation during embryogenesis or cell responses to various stimulus) and inferring regulatory networks (based on gene expression) (Stegle et al., 2015). Each of these tasks is reduced to the following stages of bioinformatics analysis: quality control of reads, filtering, alignment and mapping to the

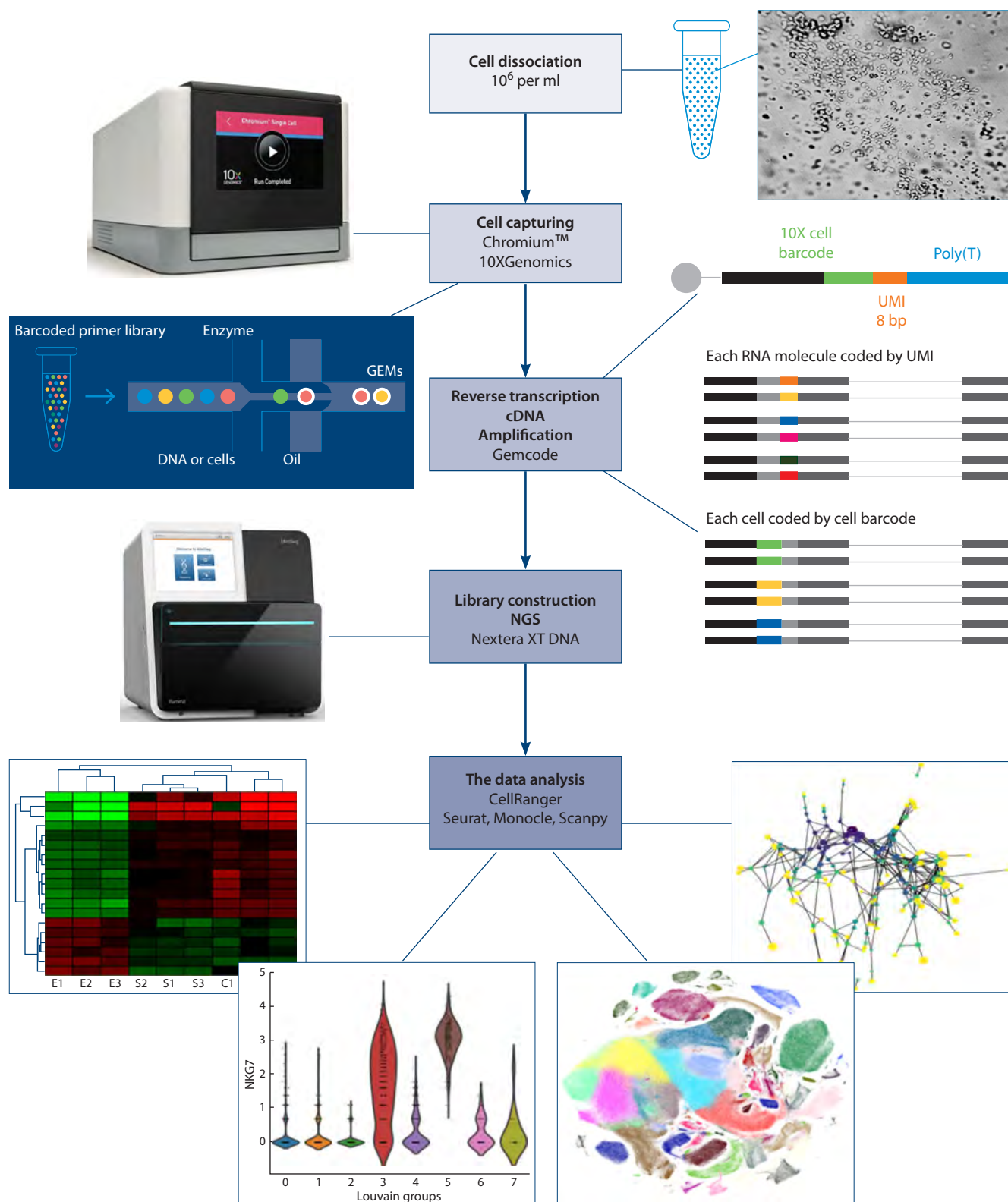


Fig. 3. The process of single-cell RNA-seq experiment using the powerful platform for capturing cell Chromium 10X Genomics.

reference genome/transcript, formation of a matrix of read counts, dimension reduction, normalization, searching for external factors introducing an error in the analysis, clustering cells, searching for marker and high variable genes, analysis of differential expression and transcript isoforms,

search for correlations in the expression of various genes (Stegle et al., 2015; Hwang et al., 2018). It is possible to use standard programs and approaches developed for bulk transcriptome for some goals.

Before interpretation of single cell expression profiles, it is necessary to incorporate technical variability estimates. The confounding factors can be divided into two categories: technical noise (for example, different cell capture efficiency, duplicated reads, dropout – missed transcript during the reverse-transcription step, batch effect – systematic differences arising as a result of sample preparation) and biological factors (stochastic gene expression, cell cycle, environment condition) (Stegle et al., 2015; Andrews, Hemberg, 2018; Hwang et al., 2018).

A specific feature of the protocols for capturing cells is the probability that two cells fall into one drop, GEM (1–10 % depending on the type of platform used), which leads to the appearance of doublets during the transcriptomic sequencing and an error during cell-types identification (Segerstolpe et al., 2016). On the other hand, the degradation of mRNA during cell/tissues dissociation or bad-quality lysis also can result in wrong classification (Brennecke et al., 2013; Hwang et al., 2018). So, such clusters of cells must be excluded from analysis during transcript quantification and quality control. The most effective is the use of protocols in conjunction with the unique molecular identifiers (UMI) (Hashimshony et al., 2012; Macosco et al., 2015) and exogenous RNA spike-ins at known concentrations and sequence which are added to each cell extract in RNA sequencing experiment (Jiang et al., 2011). It is exploited to estimate of absolute transcript counts from observed read count and normalize gene expression levels (Stegle et al., 2015). Also, you can identify confounders and technical variability in order to reduce errors during comparing expression profiles by increasing the number of biological replicates.

Using the endogenous mRNA allows to estimate the total mRNA content of each cell during normalization (Stegle et al., 2015). The most commonly used approaches for normalization include RPKM (Mortazavi et al., 2008), FPKM, TPM (Li et al., 2010). TMM and DESeq are alternative scaling factors (Robinson, Oshlack, 2010; Li et al., 2012). At the same time, normalization methods were developed for data scRNA-seq (Lun et al., 2016; Bacher et al., 2017).

Since, a huge amount of data is made during the single-cell transcriptomic sequencing for all genes in thousands of cells, creates difficulties in statistical analysis. Therefore, an important step in working with such data is the choice of a good dimension reduction strategy (Andrews, Hemberg, 2018). There are two main approaches: dimensionality reduction and removing uninformative genes. The first are: principal component analysis – PCA (Pierson, Yau, 2015); stochastic method – tSNE (Maaten, Hinton, 2008); diffusion maps – DM (Moon et al., 2018). To identify features of interest you may search highly variable genes (HGV), estimate dropout rate by Michaelis-Menten modeling of dropouts (M3Drop), analyze positive or negative correlation of gene expression cell-to-cell, use spike-in based methods (Andrews, Hemberg, 2018). Some programs combine several methods that is more effective (for example, PAGODA) (Fan et al., 2016).

It is known that the bias during clustering cell and identification novel cell-types can be caused by the stage of cell

cycle. Thus, the observed expression profile is different for G1 and G2 (Stegle et al., 2015).

The construction of regulatory networks and the reconstruction of the cell hierarchy cannot be performed for bulk transcriptomic data. The algorithm is based on the ordering of transcriptional states of various cells, which are located on a trajectory (e. g. Monocle) that characterizes a developing biological process in the organism, for example, apoptosis (Haghverdi et al., 2016).

There are some software packages available for analyzing single cell transcriptomic data step by step (Valihrach et al., 2018). So for Chromium 10XGenomics, the developer provides the CellRanger software pipeline (Zheng et al., 2017); for Drop-Seq, the developer does not propose software for primary analysis, but there are several third-party developments, for example, zUMIs (Parekh et al., 2018), scPipe (Tian et al., 2018), Dr.Seq2 (Zhao et al., 2017).

The most popular software packages for analysis of scRNA-seq data are Seurat (Butler et al., 2018) and Monocle (Qiu et al., 2017). Seurat is a software package that provides quality control, dimension reduction, analysis and estimation of genes expression. The software includes three components: uncontrolled clustering and detection of cell types and states, spatial reconstruction of single cell data and integrated analysis of scRNA-seq by conditions, technologies and types. Monocle is a comprehensive software package that provides tools for analyzing experiments with single cell genes expression. However, these software packages do not scale for accessible large sets with numbers in excess of a million cells, combining data and comparative analyses. Scanpy is an alternative approach, which overcomes this limitation and provides similar analysis capabilities (Wolf et al., 2018).

The Fig. 3 illustrated single cell transcriptomic experiment step by step using droplet-based commercial platform Chromium 10XGenomics.

Despite the rapid progress of single-cell technologies in recent years, experimental and computational tools may lead to errors during various stages of the scRNA-seq experiment (Fustin et al., 2013; Nikolenko et al., 2013; Schwartz et al., 2013; Gawad et al., 2014, 2016; Poulin et al., 2016). It must be taken into account in planning and computational analysis.

Conclusion

Thus, the single cell transcriptomic sequencing reveals the differences in genes expression profiles not only in the cell of various tissues, but also during ontogenesis as well as under changing external conditions. Due to the cell resolution, the scRNA-seq technologies are currently actively used in various fields of biology and medicine. It is innovative for investigation of convergent evolution mechanisms, different evolutionary origins and changes in cell lines of various taxonomic groups, identification of new cell-types and their functions. It is shown the need for complex researches, for example, whole organisms, tumors or tissues, which requires the development transcriptome sequencing technologies

for a huge number of cells in parallel. Nowadays the main criteria for the technology of scRNA-seq is the efficiency for capturing and barcoding cells. The most innovative and efficient for today is the commercial platform Chromium 10XGenomics with integrated GemCode technology, which allows such an analysis to be carried out in parallel to up to eighty thousand cells.

References

- Adamson B., Norman T.M., Jost M., Cho M.Y., Nunez J.K., Chen Y., Villalta J.E., Gilbert L.A., Horlbeck M.A., Hein M.Y., Pak R.A., Gray A.N., Gross C.A., Dixit A., Parnas O., Regev A., Weissman J.S. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell*. 2016;167(7):1867-1882.e21. DOI 10.1016/j.cell.2016.11.048.
- Andrews T.S., Hemberg M. Identifying cell populations with scRNASeq. *Mol. Aspects Med.* 2018;59:114-122. DOI 10.1016/j.mam.2017.07.002.
- Bacher R., Chu L.-F., Leng N., Gasch A.P., Thomson J.A., Stewart R.M., Newton M., Kendziorski C. SCnorm: robust normalization of single-cell RNA-seq data. *Nat. Methods*. 2017;14(6):584-586. DOI 10.1038/nmeth.4263.
- Belton J.M., McCord R.P., Gibcus J.H., Naumova N., Zhan Y., Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods*. 2012;58(3):268-276. DOI 10.1016/j.jmeth.2012.05.001.
- Brehm-Stecher B.F., Johnson E.A. Single-cell microbiology: tools, technologies, and applications. *Microbiol. Mol. Biol. Rev.* 2004;68(3):538-559.
- Brennecke P., Anders S., Kim J.K., Kolodziejczyk A.A., Zhang X., Proserpio V., Baying B., Benes V., Teichmann S.A., Marioni J.C., Heisler M.G. Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods*. 2013;10:1093-1095.
- Butler A., Hoffman P., Smibert P., Papalexi E., Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 2018;36(5):411-420. DOI 10.1038/nbt.4096.
- Callaway A. The trickiest family tree in biology. *Nature*. 2017;547(7661):20-22. DOI 10.1038/547020a.
- Chen Y.R., Zheng Y., Liu B., Zhong S., Giovannoni J., Fei Z. A cost-effective method for Illumina small RNA-Seq library preparation using T4 RNA ligase 1 adenylated adapters. *Plant Methods*. 2012;8(1):41. DOI 10.1186/1746-4811-8-41.
- Coombe L., Warren R.L., Jackman S.D., Yang C., Vandervalk B.P., Moore R.A., Pleasance S., Coope R.J., Bohlmann J., Holt R.A., Jones S.J.M., Birol I. Assembly of the complete Sitka spruce chloroplast genome using 10× Genomics' GemCode sequencing data. *PLoS One*. 2016;11(9):e0163059. DOI 10.1371/journal.pone.0163059.
- Deng Q., Ramsköld D., Reinis B., Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*. 2014;343(6167):193-196. DOI 10.1126/science.1245316.
- Der E., Ranabothu S., Suryawanshi H., Akat K.M., Clancy R., Morozov P., Kustagi M., Czuppa M., Izmirly P., Belmont H.M., Wang T., Jordan N., Bornkamp N., Nwaukoni J., Martinez J., Goilav B., Buyon J.P., Tuschl T., Putterman C. Single cell RNA sequencing to dissect the molecular heterogeneity in lupus nephritis. *JCI Insight*. 2017;2(9). pii: 93009. DOI 10.1172/jci.insight.93009.
- Derrien T., Johnson R., Bussotti G., Tanzer A., Djebali S., Tilgner H., Guernec G., Martin D., Merkel A., Knowles D.G., Lagarde J., Veeravalli L., Ruan X., Ruan Y., Lassmann T., Carninci P., Brown J.B., Lipovich L., Gonzalez J.M., Thomas M., Davis C.A., Shiekhattar R., Gingeras T.R., Hubbard T.J., Notredame C., Harrow J., Guigo R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*. 2012;22(9):1775-1789. DOI 10.1101/gr.132159.111.
- Eisenstein M. Startups use short-read data to expand long-read sequencing market. *Nat. Biotechnol.* 2015;33(5):433-435. DOI 10.1038/nbt0515-433.
- Fan J., Salathia N., Liu R., Kaeser G.E., Yung Y.C., Herman J.L., Kaper F., Fan J.-B., Zhang K., Chun J., Kharchenko P.V. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat. Methods*. 2016;13:241-244. DOI 10.1038/nmeth.3734.
- Frieda K.L., Linton J.M., Hormoz S., Choi J., Chow K.-H.K., Singer Z.S., Budde M.W., Elowitz M.B., Cai L. Synthetic recording and *in situ* readout of lineage information in single cells. *Nature*. 2017;541(7635):107-111. DOI 10.1038/nature20777.
- Fustin J.M., Doi M., Yamaguchi Y., Hida H., Nishimura S., Yoshida M., Isagawa T., Morioka M.S., Kakeya H., Manabe I., Okamura H. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell*. 2013;155(4):793-806. DOI 10.1016/j.cell.2013.10.026.
- Gawad C., Koh W., Quake S.R. Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. *Proc. Natl. Acad. Sci. USA*. 2014;111(50):17947-17952. DOI 10.1073/pnas.1420822111.
- Gawad C., Koh W., Quake S.R. Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* 2016;17(3):175-188. DOI 10.1038/nrg.2015.16.
- Gerber T., Willscher E., Loeffler-Wirth H., Hopp L., Schadendorf D., Schartl M., Anderegg U., Camp G., Treutlein B., Binder H., Kunz M. Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. *Oncotarget*. 2017;8(1):846-862. DOI 10.18632/oncotarget.13666.
- Goldberg A.D., Allis C.D., Bernstein E. Epigenetics: a landscape takes shape. *Cell*. 2007;128(4):635-638.
- Grun D., Lyubimova A., Kester L., Wiebrands K., Basak O., Sasaki N., Clevers H., Oudenaarden A. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature*. 2015;525(7568):251-255. DOI 10.1038/nature14966.
- Hafner M., Landgraf P., Ludwig J., Rice A., Ojo T., Lin C., Holloch D., Lim C., Tuschl T. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods*. 2008;44(1):3-12.
- Haghverdi L., Büttner M., Wolf F.A., Büttner F., Theis F.J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods*. 2016;13:845-848.
- Ham R.G. Clonal growth of mammalian cells in a chemically defined, synthetic medium. *Proc. Natl. Acad. Sci. USA*. 1965;53:288-293.
- Haue A., Engel J., Teichmann S.A., Lonnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med*. 2017;9:75.
- Hashimshony T., Senderovich N., Avital G., Klochendler A., Leeuw Y., Anavy L., Gennert D., Li S., Livak K.J., Rozenblatt-Rosen O., Dor Y., Regev A., Yanai I. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*. 2016;17:77. DOI 10.1186/s13059-016-0938-8.
- Hashimshony T., Wagner F., Sher N., Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*. 2012;2(3):666-673. DOI 10.1016/j.celrep.2012.08.003.
- Hwang B., Lee J.H., Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med*. 2018;50:96.
- Islam S., Zeisel A., Joost S., Manno G.L., Zajac P., Kasper M., Lonnberg P., Linnarsson S. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods*. 2014;11(2):163-166. DOI 10.1038/nmeth.2772.
- Jaitin D.A., Kenigsberg E., Keren-Shaul H., Elefant N., Paul F., Zaretzky I., Mildner A., Cohen N., Jung S., Tanay A., Amit I. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*. 2014;343(6172):776-779. DOI 10.1126/science.1247651.

- Jaitin D.A., Weiner A., Yofe I., Lara-Astiaso D., Kern-Shaul H., David E., Salame T.M., Tanay A., Oudenaarden A., Amit I. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell*. 2016;167(7):1883-1896. DOI 10.1016/j.cell.2016.11.039.
- Janes K.A., Wang C.C., Holmberg K.J., Cabral K., Brugge J.S. Identifying single-cell molecular programs by stochastic profiling. *Nat. Methods*. 2010;7(4):311-317. DOI 10.1038/nmeth.1442.
- Jiang L., Schlesinger F., Davis C.A., Zhang Y., Li R., Salit M., Gingeras T.R., Oliver B. Synthetic spike-in standards for RNA-seq experiments. *Genome Res*. 2011;21:1543-1551.
- Junker J.P., Oudenaarden A. Every cell is special: genome-wide studies add a new dimension to single-cell biology. *Cell*. 2014;157(1):8-11. DOI 10.1016/j.cell.2014.02.010.
- Kalhor R., Mali P., Church G.M. Rapidly evolving homing CRISPR barcodes. *Nat. Methods*. 2017;14(2):195-200. DOI 10.1038/nmeth.4108.
- Kamme F., Salunga R., Yu J., Tran D.T., Zhu J., Luo L., Bittner A., Guo H.Q., Miller N., Wan J., Erlander M. Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity. *J. Neurosci*. 2003;23(9):3607-3615.
- Klein A.M., Mazutis L., Akartuna I., Tallapragada N., Veres A., Li V., Peshkin L., Weitz D.A., Kirschner M.W. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 2015;161(5):1187-1201. DOI 10.1016/j.cell.2015.04.044.
- Kolodziejczyk A.A., Kim J.K., Svensson V., Marioni J.C., Teichmann S.A. The technology and biology of single-cell RNA sequencing. *Mol. Cell*. 2015;58(4):610-620. DOI 10.1016/j.molcel.2015.04.005.
- Kumar P., Tan Y., Cahan P. Understanding development and stem cells using single cell-based analyses of gene expression. *Development*. 2017;144:17-32. DOI 10.1242/dev.133058.
- Landry Z.C., Giovanonni S.J., Quake S.R., Blainey P.C. Optofluidic cell selection from complex microbial communities for single-genome analysis. *Methods Enzymol*. 2013;531:61-90. DOI 10.1016/B978-0-12-407863-5.00004-6.
- Lavin Y., Kobayashi S., Leader A., Amir E.D., Elefant N., Bigenwald C., Remark R., Sweeney R., Becker C.D., Levine J.H., Meinhof K., Chow A., Kim-Shulze S., Wolf A., Medaglia C., Li H., Rytlewski J.A., Emerson R.O., Merad M. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. *Cell*. 2017;169(4):750-765.e17. DOI 10.1016/j.cell.2017.04.014.
- Leung M.L., Davis A., Gao R., Casasent A., Wang Y., Sei E., Vilar E., Maru D., Kopetz S., Navin N.E. Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res*. 2017;27(8):1287-1299. DOI 10.1101/gr.209973.116.
- Li B., Ruotti V., Stewart R.M., Thomson J.A., Dewey C.N. RNA-seq gene expression estimation with read mapping uncertainty. *Bioinformatics*. 2010;26:493-500.
- Li H., Courtois E.T., Sengupta D., Tan Y., Chen K.H., Goh J.J.L., Kong S.L., Chua C., Hon L.K., Tan W.S., Wong M., Choi P.J., Wee L.J.K., Hillmer A.M., Tan I.B., Robson P., Prabhakar S. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet*. 2017;49:708-718.
- Li J., Witten D.M., Johnstone I.M., Tibshirani R. Normalization, testing, and false discovery rate estimation for RNA-sequencing data. *Biostatistics*. 2012;13:523-538.
- Lindström S., Andersson-Svahn H. Overview of single-cell analyses: microdevices and applications. *Lab. Chip*. 2010;10(24):3363-3372. DOI 10.1039/c0lc00150c.
- Liu L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L., Law M. Comparison of next-generation sequencing systems. *J. Biomed. Biotechnol*. 2012;2012:251364. DOI 10.1155/2012/251364.
- Ludwig L.S., Lareau C.A., Ulirsch J.C., Christian E., Muus C., Li L.H., Pelka K., Ge W., Oren Y., Brack A., Law T., Rodman C., Chen J.H., Boland G.M., Hacohen N., Rozenblatt-Rosen O., Aryee M.J., Buenrostro J.D., Regev A., Sankaran V.G. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell*. 2019;176(6):1325-1339. DOI 10.1016/j.cell.2019.01.022.
- Lun A.T.L., Bach K., Marioni J.C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol*. 2016;17:75. DOI 10.1186/s13059-016-0947-7.
- Maaten L. van der, Hinton G. Visualizing data using t-SNE. *J. Mach. Learn. Res*. 2008;9:2579-2605.
- Macosko E.Z., Basu A., Satija R., Nemesh J., Shekhar K., Goldman M., Tirosh I., Bialas A.R., Kamitaki N., Martersteck E.M., Trombetta J.J., Weitz D.A., Sanes J.R., Shalek A.K., Regev A., McCarroll S.A. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202-1214. DOI 10.1016/j.cell.2015.05.002.
- Mahata B., Zhang X., Kolodziejczyk A.A., Proserpio V., Haim-Vilmsky L., Taylor A.E., Hebenstreit D., Dingler F.A., Moignard V., Gottgens B., Arlt W., McKenzie A.N.J., Teichmann S.A. Single-cell RNA sequencing reveals T helper cells synthesizing steroids *de novo* to contribute to immune homeostasis. *Cell Rep*. 2014;7(4):1130-1142. DOI 10.1016/j.celrep.2014.04.011.
- Marcus J.S., Anderson W.F., Quake S.R. Microfluidic single-cell mRNA isolation and analysis. *Anal. Chem*. 2006;78(9):3084-3089.
- Mazutis L., Gilbert J., Ung W.L., Weitz D.A., Griffiths A.D., Heyman J.A. Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc*. 2013;8(5):870-891. DOI 10.1038/nprot.2013.046.
- McKenna A., Findlay G.M., Gagnon J.A., Horwitz M.S., Schier A.F., Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science*. 2016;353(6298):aaf7907. DOI 10.1126/science.aaf7907.
- Mi D., Li Z., Lim L., Li M., Moissidis M., Yang Y., Gao T., Hu T.X., Pratt T., Price D.J., Sestan N., Marin O. Early emergence of cortical interneuron diversity in the mouse embryo. *Science*. 2018;360(6384):81-85. DOI 10.1126/science.aar6821.
- Moon K.R., Dijk D., Wang Z., Burkhardt D., Chen W.S., Yim K., Elen A., Hirn M.J., Coifman R.R., Ivanova N.B., Wolf G., Krishnaswamy S. Visualizing structure and transitions for biological data exploration. *Cell*. 2018;65. <http://dx.doi.org/10.2139/ssrn.3155891>.
- Moroz L.L. NeuroSystematics and periodic system of neurons: model vs reference species at single-cell resolution. *ACS Chem. Neurosci*. 2018;9:1884-1903. DOI 10.1021/acschemneuro.8b00100.
- Moroz L.L., Edwards J.R., Puthanveetil S.V., Kohn A.B., Ha T., Heyland A., Knudsen B., Sahni A., Yu F., Liu L., Jezzini S., Lovell P., Iannuccilli W., Chen M., Nguyen T., Sheng H., Shaw R., Kalachikov S., Panchin Y.V., Farmerie W., Russo J.J., Ju J., Kandel E.R. Neuronal transcriptome of Aplysia: neuronal compartments and circuitry. *Cell*. 2006;127(7):1453-1467.
- Mortazavi A., Williams B.A., McCue K., Schaeffer L., Wold B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods*. 2008;5:621-628.
- Nagano T., Lubling Y., Stevens T.J., Schoenfelder S., Yaffe E., Dean W., Laue E.D., Tanay A., Fraser P. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*. 2013;502(7469):59-64. DOI 10.1038/nature12593.
- Nagano T., Lubling Y., Varnai C., Dudley C., Leung W., Baran Y., Cohen N.M., Wingett S., Fraser P., Tanay A. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*. 2017;547(7661):61-67. DOI 10.1038/nature23001.
- Navin N., Kendall J., Troge J., Andrews P., Rodgers L., McIndoo J., Cook K., Stepansky A., Levy D., Esposito D., Muthuswamy L., Krasnitz A., McCombie W.R., Hicks J., Wingler M. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90-94. DOI 10.1038/nature09807.
- Nikolenko S.I., Korobeynikov A.I., Alekseyev M.A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics*. 2013;14(Suppl. 1):S7. DOI 10.1186/1471-2164-14-S1-S7.
- Nowogrodzki A. The cell seeker. *Nature*. 2017;547:24-26.
- Okaty B.W., Freret M.E., Rood B.D., Brust R.D., Hennessy M.L., Bai-

- ros D., Kim J.C., Cook M.N., Dymecki S.M. Multi-scale molecular deconstruction of the serotonin neuron system. *Neuron*. 2015; 88(4):774-791. DOI 10.1016/j.neuron.2015.10.007.
- Parekh S., Ziegenhain C., Vieth B., Enard W., Hellmann I. zUMIs – A fast and flexible pipeline to process RNA sequencing data with UMIs. *GigaScience*. 2018;7(6). DOI 10.1093/gigascience/giy059.
- Pierson E., Yau C., ZIFA: Dimensionality reduction for zero-inflated single-cell gene expression analysis. *Genome Biol.* 2015;16:241. DOI 10.1186/s13059-015-0805-z.
- Pijuan-Sala B., Griffiths J.A., Guibentif C., Hiscock T.W., Jawaid W., Calero-Nieto F.J., Mulas C., Ibarra-Soria X., Tyser R.C.V., Ho D.L.L., Reik W., Srinivas S., Simons B.D., Nihols J., Marioni J.C., Gottgens B. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*. 2019;566:490-495. DOI 10.1038/s41586-019-0933-9.
- Poulin J.F., Tasic B., Hjerling-Leffler J., Trimarchi J.M., Awatramani R. Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.* 2016;19(9):1131-1141. DOI 10.1038/nn.4366.
- Qiu X., Mao Q., Tang Y., Wang L., Chawla R., Pliner H., Trapnell C. Reversed graph embedding resolves complex single-cell developmental trajectories. *Nat. Methods*. 2017;14(10):979-982. DOI 10.1038/nmeth.4402.
- Ramsköld D., Luo S., Wang Y.C., Li R., Deng Q., Faridani O.R., Daniels G.A., Khrebtkova I., Loring J.F., Laurent L.C., Schroth G.P., Sandberg R. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* 2012; 30(8):777-782.
- Robinson M.D., Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010; 11:R25. DOI 10.1186/gb-2010-11-3-r25.
- Rotem A., Ram O., Shores N., Sperling R.A., Goren A., Weitz D.A., Bernstein B.E. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat. Biotechnol.* 2015;33(11):1165-1172. DOI 10.1038/nbt.3383.
- Salafi T., Zemling K.K., Zhang Y. Advancements in microfluidics for nanoparticle separation. *Lab. Chip*. 2016;17(1):11-33.
- Saliba A.E., Westermann A.J., Gorski S.A., Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 2014; 42(14):8845-8860. DOI 10.1093/nar/gku555.
- Sassagawa Y., Danno H., Takada H., Ebisawa M., Tanaka K., Hayashi T., Kurisaki A., Nikaido I. Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biol.* 2018;19:29. DOI 10.1186/s13059-018-1407-3.
- Schwartz S., Agarwala S.D., Mumbach M.R., Jovanovic M., Mertins P., Shishkin A., Tabach Y., Mikkelsen T.S., Satija R., Ruvkun G., Carr S.A., Lander E.S., Fink G.R., Regev A. High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell*. 2013;155(6):1409-1421. DOI 10.1016/j.cell.2013.10.047.
- Segal E., Friedman N., Koller D., Regev A. A module map showing conditional activity of expression modules in cancer. *Nat. Genet.* 2004;36(10):1090-1098.
- Segerstolpe A., Palasantza A., Eliasson P., Andersson E.M., Andreasson A.C., Sun X., Picelli S., Sabirsh A., Clausen M., Bjursell M.K., Smith D.M., Kasper M., Ammalu C., Sandberg R. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab.* 2016;24:593-607. DOI 10.1016/j.cmet.2016.08.020.
- Shalek A.K., Satija R., Adiconis X., Gertner R.S., Gaublomme J.T., Raychowdhury R., Schwartz S., Yosef N., Malboeuf C., Lu D., Trombetta J.J., Gennert D., Gnirke A., Goren A., Hacohen N., Levin J.Z., Park H., Regev A. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature*. 2013; 498(7453):236-240. DOI 10.1038/nature12172.
- Soumillon M., Cacchiarelli D., Semrau S., van Oudenaarden A., Mikkelsen T.S. Characterization of directed differentiation by high-throughput single-cell RNA-seq. *bioRxiv*. 2014. DOI 10.1101/003236.
- Stegle O., Teichmann S.A., Marioni J.C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* 2015;16(3):133-145. DOI 10.1038/nrg3833.
- Stevens T.J., Lando D., Basu S., Atkinson L.P., Cao Y., Lee S.F., Leeb M., Wohlfahrt K.J., Boucher W., O'Shaughnessy-Kirwan A., Cramard J., Faure A.J., Ralser M., Blanco E., Morey L., Sanso M., Palayret M.G.S., Lehner B., Croce L.D., Wutz A., Hendrich B., Klennerman D., Laue E.D. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*. 2017;544(7648):59-64. DOI 10.1038/nature21429.
- Subkhankulova T., Gilchrist M.J., Livesey F.J. Modelling and measuring single cell RNA expression levels find considerable transcriptional differences among phenotypically identical cells. *BMC Genomics*. 2008;9:268. DOI 10.1186/1471-2164-9-268.
- Tang F., Barbacioru C., Wang Y., Nordman E., Lee C., Xu N., Wang X., Bodeau J., Tuch B.B., Siddiqui A., Lao K., Surani M.A. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods*. 2009; 6(5):377-382. DOI 10.1038/nmeth.1315. Epub 2009 Apr 6.
- Taylor D.M., Aronow B.J. The pediatric cell atlas: defining the growth phase of human development at single-cell resolution. *Dev. Cell*. 2019;49. DOI 10.1016/j.devcel.2019.03.001.
- Tian L., Su S., Dong X., Amann-Zalcenstein D., Biben C., Seidi A., Hilton D.J., Naik S.H., Ritchie M.E. scPipe: a flexible R/Bioconductor preprocessing pipeline for single-cell RNA-sequencing data. *PLoS Comput. Biol.* 2018;14(8):e100636. DOI 10.1371/journal.pcbi.1006361.
- Tirosh I., Izar B., Prakadan S.M., Wadsworth M.H., Treacy D., Trombetta J.J., Rotem A., Rodman C., Lian C., Murphy G., Fallahi-Sichani M., Dutton-Regester K., Lin J.-R., Cohen O., Shah P., Lu D., Genshaft A.S., Hughes T.K., Ziegler C.G.K., Kazer S.W., Gaillard A., Kolb K.E., Villani A.C., Johannessen C.M., Andreev A.Y., Allen E.M.V., Bertagnolli M., Sorger P.K., Sullivan R.J., Flaherty K.T., Frederick D.T., Jane-Valbuena J., Yoon C.H., Rozenblatt-Rosen O., Shalek A.K., Regev A., Garraway L.A. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352(6282):189-196. DOI 10.1126/science.aad0501.
- Tokunaga T., Hirose O., Kawaguchi S., Toyoshima Y., Teramoto T., Ikebata H., Kuge S., Ishihara T., Iino Y., Yoshida R. Automated detection and tracking of many cells by using 4D live-cell imaging data. *Bioinformatics*. 2014;30:43-51.
- Trapnell C., Cacchiarelli D., Grimsby J., Pokharel P., Li S., Morse M., Lennon N.J., Livak K.J., Mikkelsen T.S., Rinn J.L. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 2014;32(4):381-386. DOI 10.1038/nbt.2859.
- Treutlein B., Brownfield D.G., Wu A.R., Neff N., Mantalas G.L., Espinoza F.H., Desai T.J., Krasnow M.A., Quake S.R. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature*. 2014;509(7500):371-375. DOI 10.1038/nature13173.
- Valihrach L., Androvic P., Kubista M. Platforms for single-cell collection and analysis. *Int. J. Mol. Sci.* 2018;19:807.
- Wang J., Song Y. Single cell sequencing: a distinct new field. *Clin. Transl. Med.* 2017;6(1):10. DOI 10.1186/s40169-017-0139-4.
- Wang Y., Navin N.E. Advances and applications of single-cell sequencing technologies. *Mol. Cell*. 2015;58(4):598-609. DOI 10.1016/j.molcel.2015.05.005.
- Whitesides G.M. The origins and the future of microfluidics. *Nature*. 2006;442(7101):368-373.
- Wolf F.A., Angerer P., Theis F.J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 2018;19(1):15. DOI 10.1186/s13059-017-1382-0.
- Yu P., Lin W. Single-cell transcriptome study as big data. *Genom. Proteom. Bioinf.* 2016;14:21-30.
- Zeisel A., Munoz-Manchado A.B., Codeluppi S., Lonnerberg P., Manno G., Jureus A., Margues S., Munguba H., He L., Betsholtz C., Rolny C., Castelo-Branco G., Hjerling-Leffler J., Linnarsson S. Cell

- types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*. 2015;347(6226):1138-1142. DOI 10.1126/science.aaa1934.
- Zhao C., Hu S., Huo X., Zhang Y. Dr.seq2: a quality control and analysis pipeline for parallel single cell transcriptome and epigenome data. *PLoS One*. 2017;12(7):e0180583. DOI 10.1371/journal.pone.0180583.
- Zheng G.X., Terry J.M., Belgrader P., Ryvkin P., Bent Z.W., Wilson R., Ziraldo S.B., Wheeler T.D., McDermott G.P., Zhu J., Gregory M.T., Shuga J., Montesclaros L., Underwood J.G., Masquelier D.A., Nishimura S.Y., Schnall-Levin M., Wyatt P.W., Hindson C.M., Bhargava R., Wong A., Ness K.D., Beppu L.W., Deeg H.J., McFarland C., Loeb K.R., Valente W.J., Ericson N.G., Stevens E.A., Radich J.P., Mikkelsen T.S., Hindson B.J., Bielas J.H. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 2017;8:14049. DOI 10.1038/ncomms14049.
- Ziegenhain C., Vieth B., Parekh S., Reinius B., Guillaumet-Adkins A., Smets M., Leonhardt H., Heyn H., Hellmann I., Enard W. Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell*. 2017;65:631-643. DOI 10.1015/j.molcel.2017.01.023.

ORCID ID

E.A. Vodiasova orcid.org/0000-0003-3886-2880
E.S. Chelebieva orcid.org/0000-0002-7662-2573
O.N. Kuleshova orcid.org/0000-0003-3745-7066

Acknowledgements. The research were supported by the Government of the Russian Federation, grant No. 14.W03.31.0015..

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

Received September 14, 2018. Revised May 30, 2019. Accepted May 30, 2019.

Image-based analysis of quantitative morphological characteristics of wild potato tubers using the desktop application SeedCounter

K.A. Ivanova¹, E.G. Komyshev¹, M.A. Genaev^{1, 2}, A.A. Egorova^{1, 2}, K.A. Koloshina¹, N.A. Chalaya³,
D.A. Afonnikov^{1, 2}, A.V. Kochetov^{1, 2}, E.V. Rogozina³, S.V. Gerasimova^{1, 2} ✉

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

³ Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia

✉ e-mail: gerson@bionet.nsc.ru

The development of quantitative digital phenotyping methods for evaluation of wild potato (section *Petota* Dumort., genus *Solanum* L.) tuberization is required for annotation of genebank collections and selection of the suitable donor material for potato breeding. There are no available methods specifically designed for the quantitative analysis of wild potato tuber morphology. The current study is devoted to evaluation of wild potato tubers' morphological characteristics using a digital image processing technique. For this purpose, the mobile application SeedSounter developed previously for grain analysis was specifically adapted for tuber phenotyping. The application estimates the number and shape of objects scattered on a standard sheet of white paper (i.e. A3 or A4). Twelve accessions from the VIR genebank collection belonging to nine *Petota* species were grown in pots protected with garden fabric during the growing season of cultivated potato (Novosibirsk region). Tubers were collected from plants of nine genotypes. Three genotypes did not produce tubers. The weight of tubers collected from each plant was measured. The tuber yield from each plant was analyzed using SeedCounter (<http://wheatdb.org/seedcounter>). The number of tubers per plant was counted; the following characteristics were extracted from the images of individual tubers: length, width, projected area, length to width ratio, circularity, roundness, rugosity and solidity. One-way ANOVA showed a significant effect of genotype on all measured characteristics. A pairwise comparison of nine *Petota* accessions using all measured parameters revealed statistically significant differences between 86 % of pairs. The overall tuber yield volume for each plant was calculated as a sum of volumes of individual tubers; tuber volume was calculated from its length to width ratio and projected area. A strong correlation between the evaluated tuber yield volume and yield weight was shown. We propose tuber yield volume as a characteristic for a general evaluation of tuberization for wild potato, implementing the four-step scale from 0 to 3. According to this characteristic, the twelve wild potato accessions studied could be divided into four groups with different tuberization abilities. The evaluated tuberization ability is partially in accordance with previously obtained VIR data. The results presented demonstrate the possibility to use SeedCounter for wild potato collections phenotyping.

Key words: *Solanum*; *Petota*; digital phenotyping of plants; wild potato species; SeedCounter; tuberization.

For citation: Ivanova K.A., Komyshev E.G., Genaev M.A., Egorova A.A., Koloshina K.A., Chalaya N.A., Afonnikov D.A., Kochetov A.V., Rogozina E.V., Gerasimova S.V. Image-based analysis of quantitative morphological characteristics of wild potato tubers using the desktop application SeedCounter. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5):519-526. DOI 10.18699/VJ19.35-o

Оценка количественных характеристик клубнеобразования дикого картофеля на основе анализа изображений клубней с использованием компьютерного приложения SeedCounter

К.А. Иванова¹, Е.Г. Комышев¹, М.А. Генаев^{1, 2}, А.А. Егорова^{1, 2}, К.А. Колошина¹, Н.А. Чалая³,
Д.А. Афонников^{1, 2}, А.В. Кочетов^{1, 2}, Е.В. Рогозина³, С.В. Герасимова^{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 Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия

✉ e-mail: gerson@bionet.nsc.ru

Применение методов количественного цифрового фенотипирования для анализа параметров клубнеобразования диких клубненосных видов картофеля актуально для аннотации коллекций и для отбора оптимального донорного материала в работах по селекции культурного картофеля. На сегодняшний

день в литературе не описано методов, позволяющих объективно количественно оценить морфологические параметры клубней дикого картофеля. В настоящей работе был проведен эксперимент по оценке морфологических характеристик клубней диких видов секции *Petota* Dumort. рода *Solanum* L. при помощи мобильного приложения SeedCounter, анализирующего изображения объектов, размещенных на стандартном листе бумаги. Двенадцать образцов генотипов из коллекции ВИР, относящихся к девяти диким видам, были выращены в вегетационных сосудах, расположенных в парнике, обтянутом укрывным материалом, в сроки вегетации культурного картофеля (Новосибирская область). Урожай клубней получен с представителей девяти генотипов. Измерена масса клубней с одного растения, урожай каждого растения проанализирован при помощи настольной версии приложения SeedCounter (<http://wheatdb.org/seedcounter>). Произведен подсчет клубней, для каждого клубня получены следующие характеристики: длина, ширина, площадь проекции на лист бумаги, отношение длины к ширине, округлость, закругленность, шероховатость и компактность. При попарном сравнении девяти видов по перечисленным параметрам были выявлены достоверные различия в 86 % исследуемых пар. На основании показателей отношения длины к ширине и площади проекции клубня предложена формула для вычисления объема урожая клубней с одного куста. Вычисленные значения объема урожая имеют высокую корреляцию со значениями массы урожая с одного куста. Показатель объема урожая предлагается использовать для общей характеристики клубнеобразования дикого картофеля по четырехбалльной шкале (от 0 до 3). В исследуемой группе образцов выявлены генотипы, обладающие нулевым, слабым, средним и высоким клубнеобразованием. Полученные данные частично соответствуют оценке клубнеобразования, проведенной по критериям ВИР. Различие в результатах оценки, вероятно, связано с условиями выращивания материала. Результаты применения метода позволяют рассматривать его как перспективный способ стандартного фенотипирования образцов коллекций диких видов картофеля.

Ключевые слова: *Solanum*; *Petota*; цифровое фенотипирование растений; дикие виды картофеля; SeedCounter; клубнеобразование.

Introduction

The use of wild potato species as a source of resistance to adverse environmental factors forms the background of the current potato breeding (Kiru, Rogozina, 2017). The need in phenotyping of large populations is a key point in studying the diversity of the cultivated potato and related species. The cultivated potato during its domestication and the breeding focused on the customer's needs has acquired considerable differences from its wild relatives in both its phenotype and genotype (Aversano et al., 2015). One of the major distinctions is the level of tuber formation. This trait is associated with the most important characteristic of the cultivated potato cultivars, namely, a high yield. A high-throughput morphometry of the cultivated potato tubers is a rather well-developed research area that has certain tools based on several approaches, including processing of digital tuber images (Rady, Guyer, 2015). The tuberization of cultivated potato is assessed with the help of the images captured with color, multispectral, and hyperspectral cameras. For analysis of these images, the corresponding methods have been elaborated that determine the main characteristics of the tuber, such as, the length to width ratio; predict the tuber length, width, and weight, as well as the defects inside and outside the specimen; and construct a tuber 3D model (Si et al., 2017, 2018; Su et al., 2017). There is also a mobile application, PotatoSize (<https://www.hutton.ac.uk/research/groups/information-and-computational-sciences/potatosize>), for measuring the proportions of cultivated potato tubers.

Unlike the cultivated *Solanum tuberosum* with its high adaptive ability to growth conditions, the majority of wild potato species form tubers only in the case of a short photoperiod and have considerable differences in the optimal environmental characteristics for a high yield. When studying the specimens of potato and closely related species from the collection of the Institute of Plant Industry (VIR), the morphological traits

of their tubers are estimated, including its shape, tuber skin surface, depth of the eyes and stolon trail, and coloration of the skin (main and secondary) and flesh (Kiru et al., 2010). The morphological characters of tubers of the cultivated potato species are rather diverse (Huaman et al., 1977) versus the wild species of the section *Petota*, the tubers of which are less odd and diverse in their shape, color and eye depth and are sometimes missing in the botanical descriptions; however, some monographs give tuber images (Correll, 1962; Ochoa, 2004). The morphology of wild potato tubers has never been analyzed in terms of taxonomy (Spooner et al., 2004).

Any well-established criteria for tuberization and specific tuber features for the wild potato species are absent. As a rule, the wild potato plants are maintained in collections by growing in a climatic chamber or a greenhouse with differently set parameters. The phenotypic data obtained in such artificial environment are hard for processing and characterization since the phenotype of a plant grown in a climatic chamber may considerably differ from that "in field".

Here, we describe the experiment on growing wild potato species under local vegetation conditions (Novosibirsk region, Russia) during the season typical for growing the cultivated potato. In order to accelerate and standardize the estimation of tuber formation and phenotype of the wild potato species, the produced yield was phenotyped by digital imaging and image processing with the help of the SeedCounter mobile application (Komyshev et al., 2017) adapted to analyzing the wild potato tubers.

Materials and methods

In total, 12 genotypes of accessions belonging to nine wild potato species from the section *Petota* Dumort., genus *Solanum* L. according to the classification by Bukasov (1980) from the collection of the VIR Plant Genetic Resources Gene Bank, were used in the work (Table 1).

Table 1. Genotypes of wild potato species from the collection of the Institute of Plant Industry*

No.	Species	VIR catalog no.	Genotype	Tuberization (VIR data)
1	<i>S. dolichostigma</i> Buk. (syn. <i>S. chacoense</i> Bitter)	7613	6-2-n-2002	Good
2	<i>S. chacoense</i> Bitter	19759	3-42-2	
3	<i>S. chacoense</i> Bitter	22687	16-2016	
4	<i>S. commersonii</i> Dunal	21355	19-2016	
5	<i>S. fendleri</i> A. Gray	18242	549-3-2016	
6	<i>S. ehrenbergii</i> Bitter (Rydb.)	24207	715-2011	Medium
7	<i>S. jamesii</i> Torr.	24923	649-2011	Good
8	<i>S. pinnatisectum</i> Dunal	23569	18-7	
9	<i>S. pinnatisectum</i> Dunal	24239	D-36-2011	Poor
10	<i>S. polyadenium</i> Greenm.	24957	618-2011	Moderate
11	<i>S. verrucosum</i> Schltdl.	23015	346-2011	Poor
12	<i>S. tarijense</i> Hawkes	12637	14-1-2007	

* Data of the World Collection of the Institute of Plant Industry (VIR), issue 816 (2014).

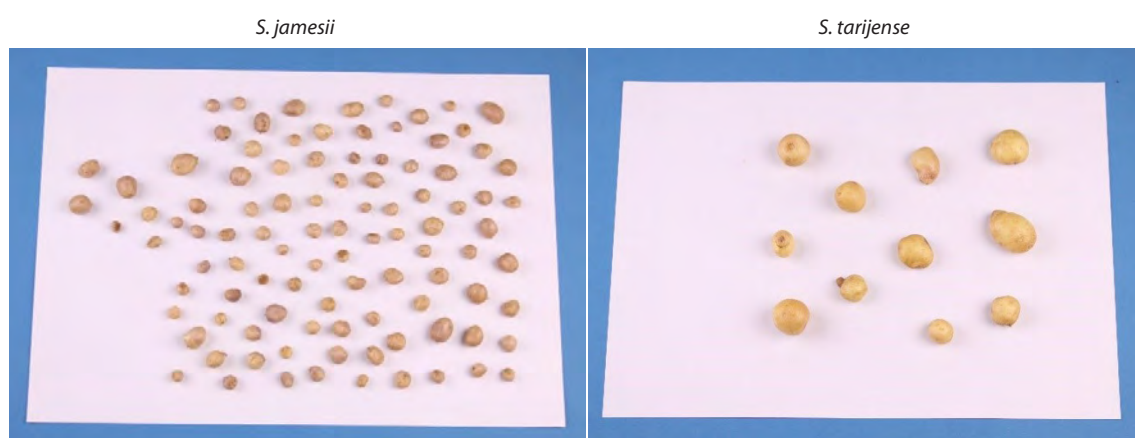


Fig. 1. Digital images of the tubers of two wild potato genotypes.

Cultivation conditions. The plants were grown in a hotbed covered with fleece, transparent for light, moisture, and air but impermeable for insects and pollen of other plants in the fields of the Siberian Research Institute of Plant Production and Breeding (Novosibirsk region, Russia). In the experiment, *in vitro* grown plantlets were planted in triplicate in 15-L buckets filled with a Terra Vita universal nutrient soil (closed company ZAO MNPP FART, St. Petersburg, Russia) mixed in equal proportions with the natural soil taken from under the turf in a birch stand and grown to the end of vegetation with regular watering. The seedlings were planted on June 20, 2018 and tubers were harvested on completion of vegetation and death of the aboveground plant part (September 20, 2018). The tubers of nine different genotypes were harvested and three accessions – *S. polyadenium*, *S. commersonii*, and *S. chacoense* (22687) – failed to tuberize.

Image capturing and processing. The tubers harvested from one plant were loosely placed on an A3 sheet of white

paper under artificial illumination (Fig. 1). The images were captured with a Canon EOS 50D camera (resolution, 4752 × 3168) according to the imaging protocol for the SeedCounter application (Komyshev et al., 2017).

Analysis of images by SeedCounter. The tuber images were analyzed by a desktop variant of the SeedCounter application (Komyshev et al., 2017; <http://wheatdb.org/seedcounter>), modified for recognition of potato tubers instead of wheat grains. In addition to the length, width, and projected area, the following indices of tuber contour shape were also computed (Cervantes et al., 2016):

- Length to width ratio, L/W , which reflects one of the main tuber shape characteristics (Si et al., 2017).
- Circularity, C , reflecting the degree to which the contour is close to a circle. The values of this index vary from 0 to 1; the latter corresponds to an ideal circle,

$$C = \frac{4\pi \times \text{area}}{\text{perimeter}^2}.$$

- Roundness, R , which is useful for the contours with numerous small bulbs on their surface. In this case, the perimeter increases and the circularity index takes on lower values, while the roundness index is independent of the irregularities of the perimeter,

$$R = \frac{4 \times \text{area}}{\pi [\text{Major axis}]^2}.$$

- Rugosity, Rg , is determined as the ratio of contour perimeter to its convex hull perimeter,

$$Rg = \frac{Ps}{Pc},$$

where Ps is the perimeter of the contour and Pc , perimeter of the contour known as a convex hull, i. e., the least convex figure that contains all points of the image; and

- Solidity, Sl , which is the ratio of contour area to its convex hull area,

$$Sl = \frac{\text{Contour Area}}{\text{Convex Hull Area}}.$$

The following assumptions were made to assess the volume of tuber yield from each plant: (1) tubers are ellipsoid with the equal height and width and the length is k -fold larger than the width ($k = L/W$ is the measured length to width ratio) and (2) the ellipsoid principle section area along the long axis is equal to the tuber projected area, S . Based on these assumptions, the following equation was obtained for assessing the tuber volume harvested from an individual plant (the total yield volume is equal to the volumes of all harvested tubers):

$$V = \sum_{i=1}^n V_i,$$

$$V_i = \frac{4}{3} \pi k_i \left(\frac{S_i}{\pi k_i} \right)^{\frac{3}{2}},$$

where n is the number of tubers; V_i is the volume of one tuber; S_i is the projected area of one tuber; and k_i is the length to width ratio.

The equation was proposed based on the following reasoning. The volume of ellipsoid is

$$V = \frac{4}{3} \pi abc.$$

The projected area of ellipsoid is an ellipse, the area of which is computed as

$$S = \pi ab.$$

The volume of ellipsoid and its projected area are related (taking into account that $b = c$) as

$$V = \frac{4}{3} Sb.$$

Expressing b via S , we get

$$a = kb,$$

$$S = \pi kb^2,$$

$$b = \sqrt{\frac{S}{\pi k}}.$$

Then,

$$V = \frac{4}{3} Sb = \frac{4}{3} S \sqrt{\frac{S}{\pi k}} = \frac{4}{3} \pi k \left(\frac{S}{\pi k} \right)^{\frac{3}{2}}.$$

Since SeedCounter computes the tuber projected area from its contour, this more precisely takes into account the specific shape features of an individual tuber. Correspondingly, the approximate estimate of the tuber volume was expressed via the projected area of the contour.

The total yield volume was also calculated from the estimate of potato tuber volume utilizing the nonlinear regression equation for the ratio of the projected area of tubers to their volume for a mixture of potato varieties proposed by Tabatabaefar (2002):

$$A = 1.1V^{0.71},$$

$$V = \left(\frac{A}{1.1} \right)^{1.408},$$

where A corresponds to the tuber projected area.

The data were statistically processed in the R environment (<http://www.r-project.org>). For each parameter, one-way ANOVA was used to analyze the effect of genotype on its value (<https://www.rdocumentation.org/packages/stats/versions/3.5.2/topics/aov>). The particular pairs of specimens that displayed statistically significant differences were determined by multiple pairwise comparison according to the mean value of each parameter using Tukey's HSD test (<https://www.rdocumentation.org/packages/stats/versions/3.5.2/topics/TukeyHSD>). The correlation between the tuber yield from one plant and the calculated values of yield volume was assessed using Pearson's product-moment correlation and the `cor.test` function in R environment (<http://www.r-project.org>).

Results

General yield characterization of tested specimens

The harvested yield was characterized according to the mean number and mean weight of the tubers from one plant (Table 2). Three genotypes failed to tuberize in all three replicates and one genotype (*S. chacoense* 3-42-2 no. 19759) formed tubers only in two replicates.

Estimation of tuber morphology

The indices that characterize tuber size and shape were computed by SeedCounter (Table 3). One-way ANOVA has shown a statistically significant effect of genotype on each parameter calculated by SeedCounter (Table 4).

Three genotypes – *S. jamesii* 649-2011, *S. tarijense* 14-1-2007, and *S. chacoense* 3-42-2 – are shown in Table 3 as examples to demonstrate statistically significant differences of one genotype from the remaining two ones. The *S. jamesii* genotype differs from the *S. tarijense* and *S. chacoense* genotypes in the tuber length, width, and projected area in a statistically significant manner (see Table 3); *S. chacoense* differs from *S. jamesii* and *S. tarijense* according to the length to width ratio in a statistically significant manner (see Table 3); and no statistically significant differences in circularity, rugosity, and solidity have been observed.

Thus, the tubers of the selected *S. tarijense* genotype can be initially characterized as large, smooth, and round; of the *S. jamesii* genotype, as small, smooth, and round; and of the *S. chacoense* genotype, as large, smooth, and elongated.

It has been shown using Tukey's test that 31 genotype pairs of the 36 (86 %) compared in a pairwise manner display sta-

Table 2. Main yield characteristics of the studied genotypes of wild potato species (mean $\pm$ SE)

No.	Species	VIR catalog no.	Genotype	Mean number of tubers per plant	Mean weight of tubers per plant, g
1	<i>S. dolichostigma</i>	7613	6-2-n-2002	27.3 $\pm$ 5.5	143.3 $\pm$ 43.6
2	<i>S. chacoense</i> *	19759	3-42-2	15.5 $\pm$ 0.7	109.0 $\pm$ 66.5
3	<i>S. chacoense</i>	22687	16-2016	–	–
4	<i>S. commersonii</i>	21355	19-2016	–	–
5	<i>S. fendleri</i>	18242	549-3-2016	13 $\pm$ 5.3	73.7 $\pm$ 3.7
6	<i>S. ehrenbergii</i>	24207	715-2011	32.7 $\pm$ 7	64.3 $\pm$ 1.5
7	<i>S. jamesii</i>	24923	649-2011	85.7 $\pm$ 13.2	100.0 $\pm$ 15.0
8	<i>S. pinnatisectum</i>	23569	18-7	33.3 $\pm$ 7	96.3 $\pm$ 29.5
9	<i>S. pinnatisectum</i>	24239	D-36-2011	8.7 $\pm$ 4.5	35.0 $\pm$ 25.2
10	<i>S. polyadenium</i>	24957	618-2011	–	–
11	<i>S. verrucosum</i>	23015	346-2011	5.3 $\pm$ 3.2	9.2 $\pm$ 11.6
12	<i>S. tarijense</i>	12637	14-1-2007	8.3 $\pm$ 2.5	95.7 $\pm$ 40.9

* Two replicates.

Table 3. Values of the tuber size and shape parameters (mean $\pm$ SE)

No.	Species	Genotype	Length, mm	Width, mm	Tuber projected area, mm ²	Length/width	Circularity	Roundness	Rugosity	Solidity
1	<i>S. dolichostigma</i>	6-2-n-2002	22.07 $\pm$ 5.55	18.57 $\pm$ 4.27	334.99 $\pm$ 154.96	1.19 $\pm$ 0.13	0.77 $\pm$ 0.07	0.84 $\pm$ 0.09	1.13 $\pm$ 0.05	0.98 $\pm$ 0.01
2	<i>S. chacoense</i>	3-42-2	29.45 $\pm$ 8.47	22.48 $\pm$ 6.07	545.79 $\pm$ 266.40	1.31 $\pm$ 0.16*	0.72 $\pm$ 0.10	0.76 $\pm$ 0.10*	1.16 $\pm$ 0.11	0.98 $\pm$ 0.01
5	<i>S. fendleri</i>	549-3-2016	22.69 $\pm$ 7.21	19.21 $\pm$ 5.76	366.09 $\pm$ 207.36	1.19 $\pm$ 0.16	0.75 $\pm$ 0.09	0.84 $\pm$ 0.11	1.14 $\pm$ 0.06	0.98 $\pm$ 0.02
6	<i>S. ehrenbergii</i>	715-2011	17.49 $\pm$ 4.68	14.70 $\pm$ 3.50	211.06 $\pm$ 102.29	1.19 $\pm$ 0.14	0.77 $\pm$ 0.08	0.84 $\pm$ 0.09	1.14 $\pm$ 0.07	0.98 $\pm$ 0.01
7	<i>S. jamesii</i>	649-2011	13.32 $\pm$ 2.82*	11.41 $\pm$ 2.19*	122.82 $\pm$ 47.83*	1.17 $\pm$ 0.12	0.82 $\pm$ 0.05	0.86 $\pm$ 0.08	1.11 $\pm$ 0.04	0.98 $\pm$ 0.01
8	<i>S. pinnatisectum</i>	18-7	18.76 $\pm$ 4.95	14.17 $\pm$ 2.75	211.72 $\pm$ 88.38	1.32 $\pm$ 0.23	0.68 $\pm$ 0.12	0.76 $\pm$ 0.13	1.19 $\pm$ 0.12	0.97 $\pm$ 0.02
9	<i>S. pinnatisectum</i>	D-36-2011	19.88 $\pm$ 3.58	17.51 $\pm$ 3.19	279.58 $\pm$ 95.11	1.14 $\pm$ 0.08	0.77 $\pm$ 0.05	0.88 $\pm$ 0.06	1.14 $\pm$ 0.04	0.98 $\pm$ 0.00
11	<i>S. verrucosum</i>	346-2011	14.46 $\pm$ 5.05	12.51 $\pm$ 4.33	156.05 $\pm$ 109.35	1.16 $\pm$ 0.09	0.79 $\pm$ 0.05	0.86 $\pm$ 0.06	1.14 $\pm$ 0.05	0.98 $\pm$ 0.01
12	<i>S. tarijense</i>	14-1-2007	28.63 $\pm$ 6.65	25.19 $\pm$ 5.42	587.38 $\pm$ 253.86	1.14 $\pm$ 0.10	0.78 $\pm$ 0.07	0.88 $\pm$ 0.07	1.14 $\pm$ 0.05	0.99 $\pm$ 0.01

Note: The genotypes mentioned in the text are in grey. * Statistically significant differences relative to the two other marked accessions (Tukey's test).

Table 4. Results of one-way ANOVA of the effect of genotype on the values of measured parameters

Significance test	Area, mm ²	Length, mm	Width, mm	Length/width	Circularity	Roundness	Rugosity	Solidity
F value	91.35	83.33	97.44	12.82	31.08	13.31	11.04	18.69
Pr (>F)	<2e–16	<2e–16	<2e–16	<2e–16	<2e–16	<2e–16	1.05e–14	<2e–16

tistically significant differences in all measured parameters. Three-thirds of the pairs of genotypes (75 %) are reliably differentiable by analyzing tuber size characteristics (tuber projected area, length, and width) and 40 % of the pairs are distinguishable by only size characteristic being the same in the remaining parameters. Not taking into account the tuber

size characteristics, the shape indices on the average allow for statistically significant differentiation of approximately 30 % of the pairs of genotypes of the studied species and three pairs (8 %) differ only in the shape characteristics and are the same in the size parameters. As has been shown, *S. pinnatisectum* genotype 18-7 no. 23569 differs from *S. jamesii* genotype 649-

Table 5. Number of the pairs of genotypes distinguishable in a statistically significant manner with the SeedCounter application according to the considered parameters

Parameter	Number of significantly distinguishable pairs	Ratio of the average number of distinguishable pairs of species to the number of pairs
Width	28	0.78
Length	27	0.75
Area	26	0.72
Length/width	14	0.39
Roundness	14	0.39
Circularity	13	0.36
Solidity	8	0.22
Rugosity	7	0.19

2011 no. 24923 and *S. dolichostigma* genotype 6-2-n-2002 no. 7613 in all eight considered parameters. Table 5 lists the data on the absolute and relative numbers of genotype pairs distinguishable in a statistically significant manner using each parameter alone.

General assessment of tuber formation

The obtained data were used to suggest the characteristic for assessment of tuberization by the tuber-forming wild potato species. The volume of the tubers harvested from one plant was selected as the main criterion since this characteristic is computable by analyzing images. According to the literature (Tabatabaefar, 2002; Su et al., 2017), the volume of potato tubers is tightly correlated with the tuber weight, which is, as a rule, used for assessing the yield. The yield volume was estimated by approximating the tuber shape with ellipsoid, using the data on the projected area and the length to width ratio of each tuber, and with the help of the earlier proposed equation for cultivated potato (Tabatabaefar, 2002).

The correlation with the measured tuber yield weight was chosen as the criterion for verification of the proposed method for estimating the tuber yield volume per plant. The total yield volume was assessed using two equations for each of the analyzed 26 plants to further calculate the Pearson correlation coefficient for the calculated and measured yield weights. The correlation coefficient for the values computed using the equation for the volume of ellipsoid expressed via the projected area of the contour was 0.87 ($p = 7.839\text{e-}09$; 95 % confidence interval: 0.7278559–0.9404020). As for the values computed according to the ratio of the volume to projected area using nonlinear regression, the correlation coefficient was 0.88 ($p = 2.359\text{e-}09$; 95 % confidence interval: 0.7533180–0.9465927). These statistical results show a high degree of correlation and suggest that the proposed equations are applicable for assessing the yield volume of wild potato species.

Figure 2 shows the characteristics of mean tuber yield weight and volume for each of the examined genotypes. The accessions fall into three arbitrary groups with poor, medium, and good tuberization according to both tuber yield weight and volume.

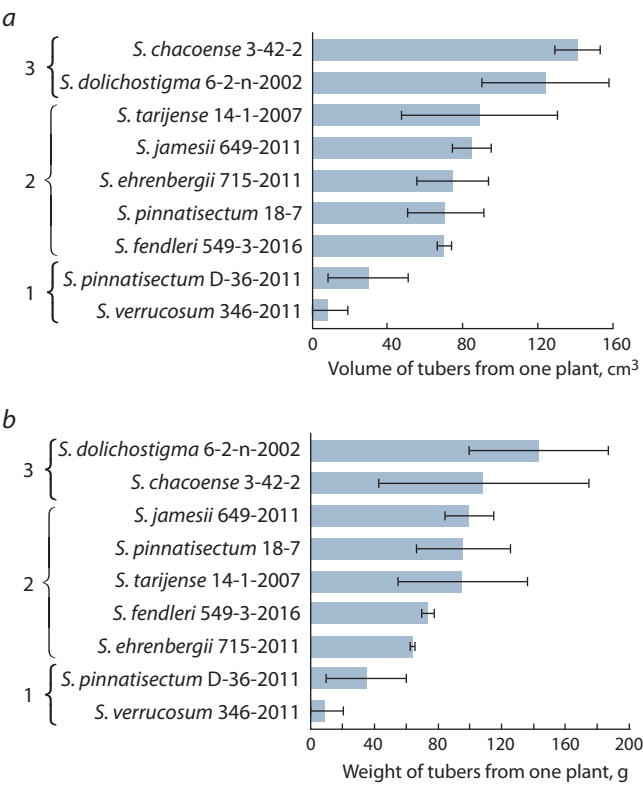


Fig. 2. Estimation of tuberization according to the (a) volume and (b) weight of the tubers from one plant: genotypes with (1) poor, (2) medium, and (3) good tuberization.

The *S. pinnatisectum* D-36-2011 no. 24239 and *S. verrucosum* 346-2011 no. 23015 genotypes form the group with poor tuberization and the *S. dolichostigma* (synonym for *S. chacoense*) 6-2-n-2002 no. 7613 and *S. chacoense* 3-42-2 no. 19759 genotypes, members of the same species, the group with good tuberization. The remaining genotypes form an intermediate (medium) group.

The values of the degree of tuberization (1 to 3) or its absence (0) were ascribed to the groups. The results were compared with each other and the data from the World Collection of the Institute of Plant Industry (Catalog..., 2014) (Table 6). Our estimates for tuberization coincided with that of the Institute of Plant Industry for genotypes nos. 1, 2, 6, 9, and 11 (see Table 6).

Discussion

Here, we describe a new approach to harvesting and analyzing the information about morphological characteristics of wild potato tubers as well as propose a new criterion for assessing the tuberization of wild potato species.

The methods utilizing digital images of tubers and their analysis are intended for estimation of the yield of cultivated potato cultivars and provide the information about its quality and usefulness. Any specialized methods allowing for assessment of tuberization by wild potato species are unavailable in the relevant literature and the existing botanical descriptions of tubers are sparse and inapplicable to adequately compare similar characteristics for different accessions. However,

Table 6. Grouping of the examined genotypes using the proposed criteria according to tuber yield characteristics

No.	Species	Genotype	Tuberization		VIR data
			from one plant according to our estimates using		
			mean volume	mean weight	
1	<i>S. dolichostigma</i> *	6-2-n-2002	3	3	Good
2	<i>S. chacoense</i> *	3-42-2	3	3	
3	<i>S. chacoense</i>	16-2016	0	0	
4	<i>S. commersonii</i>	19-2016	0	0	
5	<i>S. fendleri</i>	549-3-2016	2	2	Medium
6	<i>S. ehrenbergii</i> *	715-2011	2	2	
7	<i>S. jamesii</i>	649-2011	2	2	
8	<i>S. pinnatisectum</i>	18-7	2	2	
9	<i>S. pinnatisectum</i> *	D-36-2011	1	1	Poor
10	<i>S. polyadenium</i>	618-2011	0	0	Moderate
11	<i>S. verrucosum</i> *	346-2011	1	1	Poor
12	<i>S. tarijense</i>	14-1-2007	2	2	

* Matching of the estimates of tuberization according to the mean tuber yield volume, mean tuber yield weight, and the data of the Institute of Plant Industry (VIR).

tuberization and tuber characteristics are basic economic features of the potato. The information about tuberization of potato accessions is of a high value when selecting the donor specimens in breeding.

Incomparable cultivation conditions and a wide reaction norm of the wild species contribute to the difficulties in estimation of tuberization. We have conducted an experiment on cultivation of wild potato species under conditions of our region, which in several parameters are close to field conditions (soil composition, light day, and temperature), which has made it possible to take into account the degree of adaptation of the examined accessions to regional climate.

Processing of the digital images of wild potato tubers with the help of the elaborated SeedCounter application allows for quantification of their phenotypic characteristics related to tuber number, size, and shape. The tuber size parameters (length, width, and projected area) give the most statistically significant differences between species, while the remaining morphological parameters are less specific; however, various combinations of these parameters detect statistically significant differences between individual accessions with a high accuracy. The proposed method distinguishes not only between individual species, but also between different accessions of the same species based on the characteristics of tuber morphology. This feature of the proposed method can be useful when selecting the optimal accession of a particular species for different breeding challenges.

The phenotyping of wild potato by the proposed method forms a high-capacity background for interpretation of the obtained data. Digital phenotyping makes it possible to elaborate new characteristics of the described phenotypes and genotypes; in particular, these characteristics may be used for estimation of tuberization. The new criterion for assessing tuber formation – the volume of tubers harvested from one plant – is proposed based on the measurement of the projected area of each tuber and the length to width ratio.

A visual assessment of the tuber shape and the fact that the shape parameters inconsiderably (as a rule, by 10 %; see Table 2) vary within the studied species suggest that the tuber shape may be approximated by an ellipsoid. As has been shown earlier, the volume and weight of the cultivated potato tubers display a stably high correlation; correspondingly, the estimation of tuber volume can be used as a tool for predicting the weight of the yield (Tabatabaefar, 2002; Su et al., 2017). A comparison of the calculated estimate of the tuber yield from one plant and the measured weight of this yield demonstrates a strong correlation of these parameters. Thus, our results comply with the literature data and suggest that the proposed criterion allows for an adequate estimation of tuberization within one experiment.

However, in the further use of this method for characterization of a wide range of genotypes, it is necessary to take into account the probability of interspecific differences in the tuber specific weight. Correspondingly, it is recommended when applying this method to a new set of accessions to construct a regression model of the dependence of tuber volume on their weight. The calculated data match the data by the Institute of Plant Industry in five cases of the overall 12. Presumably, the difference in results on tuberization is explainable by different growth conditions. The method proposed for assessing tuberization is easily adaptable to different growth conditions. This estimation in several independent experiments will make it possible to determine the degree of tuberization using a four-point scale (0–3) for each accession.

Conclusion

Estimation of quantitative characteristics of tuberization by analyzing tuber images with the help of the SeedCounter application is a promising method for phenotyping of the species belonging to the section *Petota* Dumort., genus *Solanum* L. A regular use of the SeedCounter-based phenotyping when describing collections of wild potato species will

allow a large volume of new phenotypic data to be harvested and systematized as well as new characteristics advantageous for using particular accessions in breeding to be detected.

References

- Aversano R., Contaldi F., Ercolano M.R., Grosso V., Iorizzo M., Tatino F., Delledonne M., Xumerle L., Molin A.D., Avanzato C., Ferrarini A., Sanseverino W., Cigliano R.A., Capella-Gutierrez S., Gabaldón T., Frusciante L., Bradeen J.M., Carputo D. The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. *Plant Cell*. 2015;27(4):954-968. DOI 10.1105/tpc.114.135954.
- Bukasov S.M. Taxonomy of potato species of the section *Tuberarium* (Dun.) Buk: a review. *Byulleten VIR = Bulletin of the Institute of Plant Industry (Leningrad)*. 1980;105:6. (in Russian)
- Catalog of the VIR Global Collection. Issue 816. Clone Collection of Wild Potato Species. St. Petersburg, 2014. (in Russian)
- Cervantes E., Martín J.J., Saadaoui E. Updated methods for seed shape analysis. *Scientifica*. 2016;2016:5691825. DOI 10.1155/2016/5691825.
- Correll D.S. The Potato and its Wild Relatives. Renner: Texas Res. Found., 1962.
- Huaman Z., Williams J., Salhuna W., Vincent L. Descriptors for the Cultivated Potato. Rome, Italy: Int. Board for Plant Genet. Res., 1977;77(32):47.
- Kiru S.D., Kostina L.I., Truskinov E.V., Zoteeva N.M., Rogozina E.V., Koroleva L.V., Fomina V.E., Palekha S.V., Kosareva O.S., Kirilov D.A. Guidelines for the Maintenance and Study of the World Collection of Potatoes. St. Petersburg: VIR Publ., 2010. (in Russian)
- Kiru S.D., Rogozina E.V. Mobilization, conservation and study of cultivated and wild potato genetic resources. *Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding*. 2017;21(1):7-15. DOI 10.18699/VJ17.219. (in Russian)
- Komyshev E., Genaev M., Afonnikov D. Evaluation of the SeedCounter, a mobile application for grain phenotyping. *Front. Plant Sci*. 2017;7:1990. DOI 10.3389/fpls.2016.01990.
- Ochoa C.M. The Potatoes of South America: Peru. Int. Potato Center, 2004.
- Rady A.M., Guyer D.E. Rapid and/or nondestructive quality evaluation methods for potatoes: a review. *Comput. Electron. Agr.* 2015;117:31-48. DOI 10.1016/j.compag.2015.07.002.
- Si Y., Sankaran S., Knowles N.R., Pavek M.J. Potato tuber length-width ratio assessment using image analysis. *Am. J. Potato Res.* 2017;94(1):88-93. DOI 10.1007/s12230-016-9545-1.
- Si Y., Sankaran S., Knowles N.R., Pavek M.J. Image-based automated potato tuber shape evaluation. *J. Food Meas. Charact.* 2018;12(2):702-709. DOI 10.1007/s11694-017-9683-2.
- Spooner D.M., Van den Berg R.G., Rodrigues A., Bamberg J., Hijmans R.J., Lara-Cabrera S.I. Wild Potatoes (*Solanum* section *Petota*; Solanaceae) of North and Central America. *Syst. Bot. Monogr.* 2004;68.
- Su Q., Kondo N., Li M., Sun H., Al Riza D.F. Potato feature prediction based on machine vision and 3D model rebuilding. *Comput. Electron. Agr.* 2017;137:41-51. DOI 10.1016/j.compag.2017.03.020.
- Tabatabaefar A. Size and shape of potato tubers. *Int. Agrophys.* 2002;16(4):301-306.

ORCID ID

K.A. Ivanova orcid.org/0000-0002-1868-635X
D.A. Afonnikov orcid.org/0000-0001-9738-1409
A.V. Kochetov orcid.org/0000-0003-3151-5181
E.V. Rogozina orcid.org/0000-0002-2743-068X
S.V. Gerasimova orcid.org/0000-0001-8626-1831

Acknowledgements. The work was supported by the Russian Foundation for Basic Research (grant no. 18-316-00068) and the plant material used in the work is covered by the project of the Institute of Plant Industry (no. 0662-2019-0004). The SeedCounter mobile application was adapted to analysis of tuber shape under project no. 17-74-10148 of the Russian Science Foundation. The data were analyzed using the facilities of the Bioinformatics joint access center (with support of the state project of the Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences). The work with plants was performed at the joint access center "Laboratory of Artificial Plant Growth" and experimental fields (under state project of the Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences).

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

Received February 1, 2019. Revised April 8, 2019. Accepted April 15, 2019. Published online April 29, 2019.

Biotechnological bases of the development of cloned pig embryos

A.V. Lopukhov , G.N. Singina, N.A. Zinovieva

Federal Science Center for Animal Husbandry named after Academy Member L.K. Ernst, Moscow, Russia

 e-mail: vubi_myaso@mail.ru

The term 'clone' in animal biotechnology refers to an organism derived from non-sexual reproduction, which is both a direct offspring and a genetic copy of the parent organism. To date, the pig appears to be the most interesting object in cloning research. Somatic cell nuclear transfer in pigs has a wide range of potential applications in various fields of human scientific and economic activities. However, the efficiency of producing cloned embryos in swine is still lower than that of other livestock species, in particular horses and cattle. Somatic cell nuclear transfer is a technically complex multi-stage technology, at each stage of which the pig oocytes, which are more susceptible to changes of surrounding conditions, are affected by various factors (mechanical, physical, chemical). At the stage of oocyte maturation, changes in the cell ultrastructures of the ooplasm occur, which play an important role in the subsequent nuclear reprogramming of the transferred donor cell. Before transfer to the oocyte donor somatic cells are synchronized in the G0/G1 stage of the cell cycle to ensure the normal ploidy of the cloned embryo. When removing the nucleus of pig oocytes matured *in vitro*, it is necessary to pay attention to the problem of preserving the viability of cells, which were devoid of their own nuclear material. To perform the reconstruction, a somatic cell is placed, using micro-tools, in the perivitelline space, where the first polar body was previously located, or in the cytoplasm of an enucleated oocyte. The method of manual cloning involves the removal of the oocyte nucleus with subsequent fusion with the donor cell without the use of micromanipulation techniques. The increased sensitivity of oocytes to the environmental conditions causes special requirements for the choice of the system for *in vitro* culture of cloned pig embryos. In this work, we have reviewed the modern methods used for the production of cloned embryos and identified the technological issues that prevent improving the efficiency of somatic cloning of pigs.

Key words: domestic pig; *Sus scrofa domestica*; oocytes; *in vitro*; somatic cell nuclear transfer; fusion; activation; cloned embryo.

For citation: Lopukhov A.V., Singina G.N., Zinovieva N.A. Biotechnological bases of the development of cloned pig embryos. Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding. 2019;23(5): 527-533. DOI 10.18699/VJ19.521

Биотехнологические основы получения клонированных эмбрионов свиней

А.В. Лопухов , Г.Н. Сингина, Н.А. Зиновьева

Федеральный научный центр животноводства – ВИЖ им. академика Л.К. Эрнста, Москва, Россия

 e-mail: vubi_myaso@mail.ru

Термин «клон» в биотехнологии животных обозначает организм, полученный в результате неполового размножения, который одновременно является прямым потомком и генетической копией родительского организма. На сегодняшний день домашняя свинья (*Sus scrofa domestica*) представляется наиболее интересным объектом в исследованиях по клонированию. Клонирование свиней имеет широкий спектр потенциальных возможностей использования в различных областях научной и хозяйственной деятельности человека. Тем не менее эффективность получения клонированных эмбрионов свиней все еще остается ниже, чем других видов сельскохозяйственных животных, в частности лошадей и крупного рогатого скота. Соматическое клонирование – сложная многоступенчатая технология, на каждом этапе которой более восприимчивые к изменениям окружающих условий ооциты свиней испытывают неблагоприятные воздействия различных по своей природе факторов (механические, физические, химические). На этапе созревания ооцитов происходят изменения клеточных ультраструктур ооплазмы, которые играют важную роль в последующем репрограммировании ядра пересаженной донорской клетки. Донорские соматические клетки перед переносом в ооцит синхронизируют в стадии G0/G1 клеточного цикла с целью обеспечения нормальной пloidности клонированного эмбриона. При удалении ядра у созревших *in vitro* ооцитов свиней следует обращать внимание на проблему сохранения жизнеспособности клеток после извлечения собственного ядерного материала. В ходе реконструирования соматическую клетку с помощью микроинструментов помещают в перивителлиновое пространство, где ранее находилось первое

полярное тельце, или в цитоплазму энуклеированного ооцита. Метод ручного клонирования (handmade cloning) предполагает удаление ядра ооцита с последующим слиянием с донорской клеткой без помощи микроманипуляционной техники. Повышенная чувствительность ооцитов к факторам окружающей среды обуславливает особые требования к выбору системы *in vitro* культивирования клонированных эмбрионов свиней. В рамках настоящего обзора проведен мониторинг современных методов, используемых при получении клонированных эмбрионов, выявлены технологические особенности, препятствующие повышению эффективности метода соматического клонирования свиней.

Ключевые слова: домашняя свинья; *Sus scrofa domestica*; ооциты; *in vitro*; соматическое клонирование; слияние; активация; клонированный эмбрион.

Introduction

The ability of the somatic cell nucleus, which is transferred to the enucleated oocyte, to be reprogramed is one of the most important phenomena of biological science, the discovery of which made it possible to obtain reconstructed embryos and cloned animals. In practice, this was implemented in June 1996 by a group of Scottish researchers led by Ian Wilmut, who reported the birth of the first cloned mammal (Dolly the sheep) with hereditary material identical to another adult animal, which was produced using differentiated cells (breast epithelium) (Wilmut et al., 1997). The serious interest of the scientific community, caused by the revolutionary breakthrough in this field of reproductive technologies, has led to the cloning of more than 20 mammalian species (Singina et al., 2014).

The production of cloned piglets was conducted for the first time in 2000 simultaneously by two research groups from the United States and Japan (Onishi et al., 2000; Polejaeva et al., 2000). One of the main application fields of cloning technology is the use of genetically modified pigs as models for the study of human diseases and organ donors for xenotransplantation (Bethausen et al., 2000). Currently, such animals are used in preclinical testing of preventive or therapeutic medicines (Liu et al., 2008), testing the toxicity of drugs, studies of functional genomics (Wimmers et al., 2010). Production of genetically modified pigs is a potential tool for reducing physiological and immunological barriers to obtaining and transplantation of donor organs. Another equally important area of practical application of cloning is the production of animals with desired parameters of productivity by copying boars and sows with high breeding values.

At somatic cloning, instead of its own chromosomal material, the oocyte (cytoplast) acquires the nucleus of a somatic cell (karyoplast) from the animal, a genetic copy of which is to be obtained. The main stages of the cloning technology included the preparation (*in vitro* maturation) of the recipient oocyte and donor cells, removing nuclear material from the mature oocyte, reconstruction of the cytoplast obtained (fusion with karyoplasts), activation of the reconstructed oocyte and culture of the cloned embryo (Niemann et al., 2011; Simões, Santos, Jr., 2017).

Oocyte maturation

The ability of the oocyte to initiate successful development into the cloned, parthenogenetic embryo and the embryo after *in vitro* fertilization is largely determined by its maturation. Cytoplasmic maturation includes modifications of the cytoplasm, in particular, redistribution of organelles, changes in cytoskeletal dynamics, micro- and macromolecular alterations

(Ferreira et al., 2009). Nuclear maturation involves modifications of chromatin during the period from the destruction of the germinal vesicle to the metaphase of the second meiotic division (MII) (Marteil et al., 2009).

Unlike *in vitro* culture of other animal species, that of pig oocytes relies on a two-phase maturation protocol is used, which makes it possible to increase their competence to fertilization or to artificial activation. At the first stage, the proteins that are necessary for the early embryonic development are produced in the oocyte under the action of hormones. The second period of maturation takes place without external signals and includes the division of the nucleus and the structuring of cell organelles. The matured oocytes at the stage of metaphase II with the visualized first polar body are usually used as sources of cytoplasts for cloning (Hardarson et al., 2000). In pigs, the optimal duration of *in vitro* maturation of oocytes varies according to different studies in the range from 24 to 44 hours (Zhang et al., 2006; Sugimura et al., 2010).

Preparation of donor somatic cells

The efficiency of reprogramming after transfer of somatic cell nuclei depends on a number of factors, including the type, the number of passages and the stage of the cell cycle of the donor cell (Enright et al., 2003; Yang et al., 2007). Particular importance for the success of cloning has the level of differentiation of the donor cell (Jaenisch, 2012). To date, cloned swine embryos have been produced using different types of somatic cells, such as fetal fibroblasts, skin fibroblasts, neural stem cells, cumulus cells, granulosa cells and breast epithelial cells (Verma et al., 2000; Cervera et al., 2009; Zheng et al., 2009).

A necessary condition for nucleus reprogramming and for the successful development of the cloned embryo is the synchronization of the cell cycle of the cytoplast and karyoplast. With this aim, the donor cells of all types are normally subjected to artificial arrest in phase G0/G1 prior to micro-manipulation procedures (Boquest et al., 1999). The transfer of somatic cell nuclei, of which DNA has not been replicated yet, to the oocyte at metaphase II reduces the risk of chromosomal abnormalities and ploidy disorders of cloned embryos (Campbell et al., 1996). Coordination between the cell cycles of oocytes and somatic cells of pigs is ensured by limitation of serum content or by contact inhibition. Despite the fact that the first cloned piglets were obtained after transfer of donor cell nuclei after the limitation of serum content (Polejaeva et al., 2000), this method is characterized by negative effects. One, for example, is the initiation of apoptotic phenomena along with increased DNA fragmentation at the culture of swine fibroblasts in a medium with a low serum content (Kues et al.,

2002). The most commonly used method for synchronizing the cell cycle in a pig cloning procedure is contact inhibition of somatic cells in confluence (Park et al., 2010). In addition, chemical antimitotic agents (mimosine, rapamycin, roscovitine etc.) have become more broadly applied in the recent years (Vacková et al., 2003; Park et al., 2010; Hyun et al., 2016).

Removal of the nuclear material of oocytes

Complete removal of nuclear material from the mature oocyte eliminates the possibility of aneuploidy, reduces the risk of spontaneous parthenogenetic activation and abnormal development of the cloned embryo. However, due to the presence of fat inclusions in the cytoplasm, the nuclei of swine oocytes are not visualized under an inverted microscope without prior staining of DNA with fluorochrome dyes in combination with ultraviolet visualization (Tatham et al., 1995). The disadvantage of the classical technique is the damage of mitochondrial DNA and ooplasmic organelles by ultraviolet light. Therefore, an alternative method of blind enucleation is used, which is based on the assumption that metaphase chromosomes in mature oocytes are attached to the mitotic spindle, and their position is determined by an indirect sign, the localization of the first polar body (FPB). Thus, after removal of the FPB and part of the cytoplasm, the mature oocyte also loses the metaphase plate (McGrath, Solter, 1983). The disadvantage of this approach is the migration of the FPB relative to the metaphase plate (Hardarson et al., 2000; Miao et al., 2004) due to aging of mature oocytes (Miao et al., 2009).

Another method of enucleation is based on the cutting of the zona pellucida close to the FBP with subsequent compression of the oocyte by a glass needle to remove a small volume of the ooplasm. The removed ooplast remains intact and is therefore convenient for DNA staining to confirm removal of the mitotic spindle without exposing the oocyte to the harmful effects of ultraviolet light. The disadvantage of the procedure is its complexity and difficulty to control the volume of the removed cytoplasm. The compression method, which is associated with removal of the metaphase plate by releasing part of the cytoplasm of the mature swine oocyte *in vitro*, is a time-consuming procedure, which is characterized by a higher rate of degeneration compared to the classical method (Lee et al., 2008).

In 2002, a group of researchers reported on the production of cloned piglets after chemical enucleation (Yin et al., 2002). Chemical enucleation is based on the use of topoisomerase II inhibitors blocking the onset of telophase II, resulting in the expelling of mitotic spindle to the cell boundary (Fulka, Mur, 1993; Savard et al., 2004). A short exposition of pig oocytes at metaphase II in 0.4 mg/ml demecolcine – a chemical agent causing the depolymerization of microtubules – in the presence of 0.05 M solution of sucrose causes membrane protrusion containing a condensed chromosome mass, which can be easily removed through aspiration (Kawakami et al., 2003).

The reconstruction of the oocytes with the aim of producing cloned embryos

The traditional method of reconstruction involves the transplantation of a donor cell into the perivitelline space of the recipient oocyte. After fixation of the oocyte with a holding

pipette, the somatic cell is transferred to the oocyte through a hole or incision formed during enucleation using a transfer pipette (Popova et al., 2009). At intracytoplasmic injection, the karyoplast is introduced directly into the cytoplasm of the enucleated oocyte, bypassing the perivitelline space (Onishi et al., 2000; Lee et al., 2003; Kong et al., 2008). The factor limiting the use of this method of producing cloned embryos is the incomprehensible mechanism of destruction of the donor cell membrane in the cytoplasm of the recipient oocyte. In the case of maintaining the integrity of the membrane of the transferred cell, and thus, not passing the reprogramming of its nucleus during intracytoplasmic injection, embryos do not develop further (Lee et al., 2003).

Cloned embryos can be produced without the use of micro-manipulation techniques (Vajta et al., 2005). In the framework of the method of manual cloning (Hand Made Cloning, HMC), oocytes at metaphase II are released from the zona pellucida with the help of the enzyme pronase, cut by a microscalpel into two equal parts, which are stained with the vital fluorescent dye Hoechst 33342 in order to accurately detect the location of the metaphase plate. Two halves of oocytes that do not contain chromatin are selected for fusion with the somatic cell (Vajta et al., 2001). The successful use of this method to produce cloned pigs has been reported (Kragh et al., 2004). The limiting factor in ensuring the reproducibility of the HMC technology is the need to set up adequate conditions for culture of zona pellucida free embryos. The use of two mature oocytes to produce one reconstructed embryo by the HMC method results in the loss of 50 % of the initial material. The presence in the HMC embryos of three genotypes of mitochondrial DNA potentially increases the level of mitochondrial heteroplasmy (Oback et al., 2003).

The development of a cloned embryo is impossible without the fusion of the recipient oocyte with the transferred donor cell. In the practice of somatic cloning, there is a technique called 'electrofusion', which is fusion of the cytoplasm and the karyoplast with the use of the phenomenon of electric breakdown of the membranes of the contacting pair in a pulsed electric field. The procedure of electrofusion involves the use of two types of electrical signals, inhomogeneous AC pulses and rectangular DC pulses. The electrical oscillations that occur during electrofusion cause excessive heating of the medium with high conductivity. For this reason, buffer solutions with low electrical conductivity, which can cause the formation of dielectric potentials within the cells to facilitate intercellular contact, are chosen for electrofusion. Mainly, the Zimmerman medium in various modifications is used for cell fusion (Robl et al., 1987; Nickoloff, 1995). The physical factors affecting the effectiveness of the fusion of the cytoplasm and karyoplast include voltage, duration and repeatability of the electric field pulse. As a rule, the effect of dielectrophoretic forces on the cells is caused by the guidance of a high-frequency (1–3 MHz) sinusoidal AC field of low amplitude (~100–300 V/cm). Cells that come into contact with plasma membranes are fused by one or two rectangular pulses of high voltage DC (1–10 kV/cm) with a duration of 10–50 μ s (Cao et al., 2008). The literature is rich in methods of cell electrofusion, which is due to both the technical characteristics of the devices used for these purposes, and the particularities of different types of cells.

Activation of reconstructed oocytes

In the process of cloning, the enucleated oocytes after fusion with the diploid donor cell are activated for further development (Campbell, 1999). For cloning of pig embryos, the activation signals used are DC pulses (Im et al., 2004; Hölker et al., 2005), the chemical agents ionomycin, ionophore Ca^{2+} A23187 (Yin et al., 2002; Hyun et al., 2003; Garcia-Mengual et al., 2008), and thimerosal in combination with dithiothreitol (Im et al., 2006; Whitworth et al., 2009). Among the existing methods, electrostimulation is the most commonly used for the production of cloned embryos. It is reported that the magnitude of the emission of Ca^{2+} ions is proportional to the number and magnitude of pores formed during electrical stimulation, and depends on the number and duration of electrical pulses and electric field strength (Fissore, Robl, 1992). One DC pulse leads to a single mobilization of intracellular calcium reserves. In contrast, the strategy of multiple electrical pulses (2–3) stimulates the generation of a long series of Ca^{2+} oscillator peaks, increasing many times the concentration of these cations in the ooplasm (Fissore et al., 1999).

In studies on somatic cloning of pigs (Verma et al., 2000; Zhu et al., 2002), the multiplication of electrical impulses positively correlated with the high level of development of reconstructed oocytes to the stage of morula and blastocysts. The specific feature of obtaining cloned pig zygotes is the simultaneous electroactivation and electrofusion of the enucleated oocyte and somatic donor cells (Hyun et al., 2003; Lee et al., 2003; Skrzyszowska et al., 2008). Transgenic cloned pig embryos produced from the oocytes reconstructed using fetal fibroblasts, which were activated by electric pulses and by subsequent incubation in a solution of ionomycin, were inferior in terms of development to the blastocyst stage to the oocytes fused with the somatic cell and activated simultaneously (Hyun et al., 2003). Simultaneous fusion and activation of pig oocytes led to the improvement of embryonal development of oocytes reconstructed using fetal fibroblasts compared to the use of shared electrical and chemical activation (Samiec, Skrzyszowska, 2010). At the same time, it should be noted that no clear and reproducible unified protocols of electrical activation – like those developed for other mammalian species – have yet been developed for pigs, and the parameters of the electric field (the number of electrical impulses, the duration of electrical stimulation and the interval between pulses) significantly vary (Koo et al., 2005; Cervera et al., 2010; Peng et al., 2013).

One of the approaches for increasing the effectiveness of artificial activation is the by using stimuli that increase the concentration of Ca^{2+} ions in the cytoplasm, in combination with factors that inhibit the activity of the maturation promoting factor (Presicce, Yang, 1994; Cheng et al., 2007). Electrical stimulation in conjunction with postactivation culture in the presence of 6-dimethylaminopurine (6-DMAP) or cycloheximide resulted in an increase in the yield of blastocysts compared to conventional electroactivation (Kim et al., 2005; Im et al., 2006). Culture in the presence of 6-DMAP after electrical stimulation leads to an increase in the oscillations of intracellular Ca^{2+} ions, which is observed throughout the entire period of postactivation of activated pig oocytes (Im et al., 2006, 2007). The combination of the chemical agents

6-DMAP + cytochalasin B improved the parthenogenetic development of embryos to the blastocyst stage, but these blastocysts were characterized by a reduced number of cells. Postactivation of pig oocytes in 6-DMAP increased the yield of blastocysts on day 7 of culture in comparison with the incubation in the presence of cycloheximide or cytochalasin B (Gruppen et al., 2002).

Culture of cloned embryos

Activated reconstructed oocytes, which have started their embryonic development, are cultured *in vitro* in special media until transplantation to the recipient animal. As is known, when comparing *in vivo* and *in vitro* produced embryos, the latter have a reduced potential to embryonic development (Uhm et al., 2009; Gil et al., 2017). On the other hand, it has been shown that cloned embryos are more sensitive to culture conditions compared to parthenogenetic embryos (Heindryckx et al., 2001). These observations, along with the increased susceptibility of embryos to environmental factors when they are maintained outside the sow's body, including temperature fluctuations, set special requirements to the systems of *in vitro* culture of reconstructed pig oocytes. Culture media should provide overcoming a developmental arrest of 4-cell pig embryos, activation of the own genome and *in vitro* development to advanced stages of embryogenesis (morula and blastocyst).

Different media such as NCSU-23 and NCSU-37, BECM-3, PZM-3, PZM-4 and PZM-5 are used for culture of pig embryos (Dobrinisky et al., 1996; Yoshioka et al., 2002; Im et al., 2004). NCSU-23 is a traditional and initially widely used medium for the development of *in vitro* fertilized and cloned swine embryos. On the other hand, PZM-3, of which the composition is similar to the environment of the pig oviduct, allows increasing the ratio of embryos developed to the blastocyst stage at the 7th day of culture, and the number of cells of inner cell mass compared to NCSU-23 media (Im et al., 2004). Cloned embryos are known to be more susceptible to apoptotic degeneration than embryos developed after *in vitro* fertilization of intact oocytes (Ju et al., 2010). It has been shown that culture in PZM-5 developed by Japanese researchers in 2004 (Suzuki et al., 2004) reduces the apoptotic index in the pig embryos, produced by nuclei transfer of somatic cells (Yamanaka et al., 2009).

However, it should be noted that, despite the local successes in the development and application of new culture media, currently the conditions for embryo culture of this animal species are still not optimal. That is why, according to scientists from Canada (Cordova et al., 2017), in most of the experiments on nuclei transfer of somatic cells, the reconstructed embryos are transplanted to the host animals prior to or at early stages of cleavage. For example, a more efficient transplantation of embryos to the recipient sow was shown at hours 4–6 following activation of reconstructed oocytes than at hours 20–24 (1–2 cell embryos), the increase in efficiency being seen as an increase in pregnancy rate and the overall effectiveness of cloning (Shi et al., 2015). On the other hand, *in vitro* incubation to the morula and blastocyst stage allows controlling each division-cleavage with the possibility to select high-quality embryos with the greatest potential for further development (Jin et al., 2018). As confirmation of this

suggestion, the extension of the duration of *in vitro* culture of cloned embryos from 20 to 40 hours increased the number of pregnant recipients by 13 %, and from 22 to 120 hours by 61.8 % (Ju et al., 2010; Rim et al., 2013).

Conclusion

Analysis of literature data has shown that the cloning technology allows the researcher to create cloned pig embryos and to obtain viable offspring after transfer to the recipient animal. While some stages of cloning have become routine in many aspects (fusion, enucleation, reconstruction), others are still uncertain and require additional research (oocyte maturation, donor cells and embryo culture). It is also obvious that it will be necessary to pay special attention to studying the mechanisms of somatic cell reprogramming and regulation of oocyte quality – for using quality oocytes as sources of cytoplasts.

References

- Betthausen J., Forsberg E., Augenstein M., Childs L., Eilertsen K., Enos J., Forsythe T., Golueke P., Jurgella G., Koppang R., Lesmeister T., Mallon K., Mell G., Misica P., Pace M., Pfister-Genskow M., Strelchenko N., Voelker G., Watt S., Thompson S., Bishop M. Production of cloned pigs from *in vitro* systems. *Nat. Biotechnol.* 2000; 18:1055-1059. DOI 10.1038/80242.
- Boquest A.C., Day B.N., Prather R.S. Flow cytometric cell cycle analysis of cultured porcine fetal fibroblast cells. *Biol. Reprod.* 1999;60: 1013-1019. DOI 10.1095/biolreprod60.4.1013.
- Campbell K.H. Nuclear transfer in farm animal species. *Semin. Cell Dev. Biol.* 1999;10(3):245-252. DOI 10.1006/scdb.1999.0310.
- Campbell K.H., Loi P., Otaegui P.J., Wilmut I. Cell cycle co-ordination in embryo cloning by nuclear transfer. *Rev. Reprod.* 1996;1(1):40-46. DOI 10.1530/ror.0.0010040.
- Cao Y., Yang J., Yin Z.Q., Luo H.Y., Yang M., Hu N., Yang J., Huo D.Q., Hou C.J., Jiang Z.Z., Zhang R.Q., Xu R., Zheng X.L. Study of high-throughput cell electrofusion in a microelectrode-array chip. *Microfluid. Nanofluidics.* 2008;5:669-675. DOI 10.1007/s10404-008-0289-1.
- Cervera R.P., Marti-Gutierrez N., Escorihuela E., Moreno R., Stojkovic M. Trichostatin A affects histone acetylation and gene expression in porcine somatic cell nucleus transfer embryos. *Theriogenology.* 2009;72:10971110. DOI 10.1016/j.theriogenology.2009.06.030.
- Cervera R.P., Silvestre M.A., Marti N., Garcia-Mengual E., Moreno R., Stojkovic M. Effects of different oocyte activation procedures on development and gene expression of porcine pre-implantation embryos. *Reprod. Domest. Anim.* 2010;45:12-20. DOI 10.1111/j.1439-0531.2009.01509.x.
- Cheng W.M., Sun X.L., An L., Zhu S.E., Li X.H., Li Y., Tian J.H. Effect of different parthenogenetic activation methods on the developmental competence of *in vitro* matured porcine oocytes. *Anim. Biotechnol.* 2007;18:131-141. DOI 10.1080/10495390601096148.
- Cordova A., King W.A., Mastromonaco G.F. Choosing a culture medium for SCNT and iSCNT reconstructed embryos: from domestic to wildlife species. *J. Anim. Sci. Technol.* 2017;59:24. DOI 10.1186/s40781-017-0149-1.
- Dobransky J.R., Johnson L.A., Rath D. Development of a culture medium (BECM-3) for porcine embryos: effects of bovine serum albumin and fetal bovine serum on embryo development. *Biol. Reprod.* 1996;55(5):1069-1074.
- Enright B.P., Kubota C., Yang X., Tian X.C. Epigenetic characteristics and development of embryos cloned from donor cells treated by Trichostatin A or 5-aza-2'-deoxycytidine. *Biol. Reprod.* 2003;69: 896-901. DOI 10.1095/biolreprod.103.017954.
- Ferreira E.M., Vireque A.A., Adona P.R., Meirelles F.V., Ferriani R.A., Navarro P.A. Cytoplasmic maturation of bovine oocytes: structural and biochemical modifications and acquisition of developmental competence. *Theriogenology.* 2009;71:836-848. DOI 10.1016/j.theriogenology.2008.10.023.
- Fissore R.A., Long C.R., Duncan R.P., Robl J.M. Initiation and organization of events during the first cell cycle in mammals: applications in cloning. *Cloning.* 1999;1(2):89-100. DOI 10.1089/15204559950019979.
- Fissore R.A., Robl J.M. Intracellular Ca²⁺ response of rabbit oocytes to electrical stimulation. *Mol. Reprod. Dev.* 1992;32:9-16. DOI 10.1002/mrd.1080320103.
- Fulka J., Moor R.M. Noninvasive chemical enucleation of mouse oocytes. *Mol. Reprod. Dev.* 1993;34:427-430. DOI 10.1002/mrd.1080340412.
- Garcia-Mengual E., Alfonso J., Salvador I., Duque C.C., Silvestre M.A. Oocyte activation procedures and influence of serum on porcine oocyte maturation and subsequent parthenogenetic and nuclear transfer embryo development. *Zygote.* 2008;16:279-284. DOI 10.1017/S0967199408004796.
- Gil M.A., Martinez C.A., Nohalez A., Parrilla I., Roca J., Wu J., Ross P.J., Cuello C., Izpisua J.C., Martinez E.A. Developmental competence of porcine genome-edited zygotes. *Mol. Reprod. Dev.* 2017;84(9):814-821. DOI 10.1002/mrd.22829.
- Grupe C., Mau J.C., McIlpatrick S.M., Maddocks S., Nottle M.B. Effect of 6-dimethylaminopurine on electrically activated *in vitro* matured porcine oocytes. *Mol. Reprod. Dev.* 2002;62:387-96. DOI 10.1002/mrd.10126.
- Hardarson T., Lundin K., Hamberger L. The position of the metaphase II spindle cannot be predicted by the location of the first polar body in the human oocyte. *Hum. Reprod.* 2000;15(6):1372-1376. DOI 10.1093/humrep/15.6.1372.
- Heindryckx B., Rybouchkin A., Van Der Elst J., Dhont M. Effect of culture media on *in vitro* development of cloned mouse embryos. *Cloning.* 2001;3(2):41-50. DOI 10.1089/15204550152475545.
- Hölker M., Petersen B., Hassel P., Kues W.A., Lemme E., Lucas-Hahn A., Niemann H. Duration of *in vitro* maturation of recipient oocytes affects blastocyst development of cloned porcine embryos. *Cloning Stem Cells.* 2005;7:35-44. DOI 10.1089/clo.2005.7.35.
- Hyun H., Lee S.E., Son Y.J., Shin M.Y., Park Y.G., Kim E.Y., Park S.P. Cell synchronization by rapamycin improves the developmental competence of porcine SCNT embryos. *Cell. Reprogram.* 2016; 18(3):195-205. DOI 10.1089/cell.2015.0090.
- Hyun S., Lee G., Kim D., Kim H., Lee S., Nam D., Jeong Y., Kim S., Yeom S., Kang S., Han J., Lee B., Hwang W. Production of nuclear transfer-derived piglets using porcine fetal fibroblasts transfected with the enhanced green fluorescent protein. *Biol. Reprod.* 2003;69: 1060-1068. DOI 10.1095/biolreprod.102.014886.
- Im G.S., Lai L., Liu Z., Hao Y., Wax D., Bonk A., Prather R.S. *In vitro* development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. *Theriogenology.* 2004;61(6):1125-1135. DOI 10.1016/j.theriogenology.2003.06.006.
- Im G.S., Samuel M., Lai L., Hao Y., Prather R.S. Development and calcium level changes in pre-implantation porcine nuclear transfer embryos activated with 6-DMAP after fusion. *Mol. Reprod. Dev.* 2007;74:1158-1164. DOI 10.1002/mrd.20492.
- Im G.S., Seo J.S., Hwang I.S., Kim D.H., Kim S.W., Yang B.C., Yang B.S., Lai L., Prather R.S. Development and apoptosis of pre-implantation porcine nuclear transfer embryos activated with different combination of chemicals. *Mol. Reprod. Dev.* 2006;73:1094-1101. DOI 10.1002/mrd.20455.
- Jaenisch R. Nuclear cloning and direct reprogramming: the long and the short path to Stockholm. *Cell Stem Cell.* 2012;11(6):744-747. DOI 10.1016/j.stem.2012.11.005.
- Jin Y., Zhang M., Ju X., Liang S., Xiong Q., Zhao L., Nie X., Hou D., Liu Q., Wang J., Wang C., Li X., Zhang L., Liu X., Wang Y., Yang H., Dai Y., Li R. Factors influencing the somatic cell nuclear transfer efficiency in pigs. *Front. Agr. Sci. Eng.* 2019;6(1):73-83. (2018. Epub ahead of print). DOI 10.15302/J-FASE-2018220.

- Ju S., Rui R., Lu Q., Lin P., Guo H. Analysis of apoptosis and methyltransferase mRNA expression in porcine cloned embryos cultured *in vitro*. *J. Assist. Reprod. Genet.* 2010;27(1):49-59. DOI 10.1007/s10815-009-9378-7.
- Kawakami M., Tani T., Yabuuchi A., Kobayashi T., Murakami H., Fujimura T., Kato Y., Tsunoda Y. Effect of demecolcine and nocodazole on the efficiency of chemically assisted removal of chromosomes and the developmental potential of nuclear transferred porcine oocytes. *Cloning Stem Cells.* 2003;5(4):379-387. DOI 10.1089/153623003772032871.
- Kim Y.S., Lee S.L., Ock S.A., Balasubramanian S., Choe S.Y., Rho G.J. Development of cloned pig embryos by nuclear transfer following different activation treatments. *Mol. Reprod. Dev.* 2005;70:308-313. DOI 10.1002/mrd.20211.
- Kong Q.R., Luo Y.B., Tian J.T., Wang Z.K., Zhang L., Liu Z.H. Production of porcine reconstructed embryos by whole-cell intracytoplasmic microinjection (Article in Chinese). *Fen Zi Xi Bao Sheng Wu Xue Bao.* 2008;41(1):70-74.
- Koo D.B., Chae J.I., Kim J.S., Wee G., Song B.S., Lee K.K., Han Y.M. Inactivation of MPF and MAP kinase by single electrical stimulus for parthenogenetic development of porcine oocytes. *Mol. Reprod. Dev.* 2005;72:542-549. DOI 10.1002/mrd.20382.
- Kragh P.M., Vajta G., Corydon T.J., Purup S., Bolund L., Callesen H. Production of transgenic porcine blastocysts by hand-made cloning. *Reprod. Fertil. Dev.* 2004;16:315-318. DOI 10.10371/RD04007.
- Kues W.A., Carnwath J.W., Paul D., Niemann H. Cell cycle synchronization of porcine fetal fibroblasts by serum deprivation initiates a nonconventional form of apoptosis. *Cloning Stem Cells.* 2002;4(3):231-243. DOI 10.1089/15362300260339511.
- Lee E., Estrada J., Piedrahita J. Comparative study on the efficiency of two enucleation methods in pig somatic cell nuclear transfer: effects of the squeezing and the aspiration methods. *Anim. Biotechnol.* 2008;19(2):71-79. DOI 10.1080/10495390701839264.
- Lee J.W., Wu S.C., Tian X.C., Barber M., Hoagland T., Riesen J., Lee K.H., Tu C.F., Cheng W.T., Yang X. Production of cloned pigs by whole-cell intracytoplasmic microinjection. *Biol. Reprod.* 2003;69:995-1001. DOI 10.1095/biolreprod.103.015917.
- Liu Y., Zeng B.H., Shang H.T., Cen Y.Y., Wei H. Bama miniature pigs (*Sus scrofa domestica*) as a model for drug evaluation for humans: comparison of *in vitro* metabolism and *in vivo* pharmacokinetics of lovastatin. *Comp. Med.* 2008;58:580-587.
- Marteil G., Richard-Parpaillon L., Kubiak J.Z. Role of oocyte quality in meiotic maturation and embryonic development. *Reprod. Biol.* 2009;9(3):203-224. DOI 10.1016/S1642-431X(12)60027-8.
- McGrath J., Solter D. Nuclear transplantation in the mouse embryo by microsurgery and cell fusion. *Science.* 1983;220:1300-1302. DOI 10.1002/jez.1402280218.
- Miao Y.L., Kikuchi K., Sun Q.Y., Schatten H. Oocyte aging: cellular and molecular changes, developmental potential and reversal possibility. *Hum. Reprod. Update.* 2009;15(5):573-585. DOI 10.1093/humupd/dmp014.
- Miao Y., Ma S., Liu X., Miao D., Chang Z., Luo M., Tan J. Fate of the first polar bodies in mouse oocytes. *Mol. Reprod. Dev.* 2004;69:66-76. DOI 10.1002/mrd.20148.
- Nickoloff J.A. (Ed.) *Animal Cell Electroporation and Electrofusion Protocols*. In: Ser. "Methods in Molecular Biology". Vol. 48. Totowa, NJ: Humana Press, 1995.
- Niemann H., Kues W.A., Lucas-Hahn A., Carnwath J.W. Somatic cloning and epigenetic reprogramming in mammals. In: Atala A., Lanza R., Thompson J., Nerem R. (Eds.) *Principles in Regenerative Medicine*. 2nd edn. Academic Press, 2011:148-167.
- Obach B., Wiersma A.T., Gaynor P., Laible G., Tucker F.C., Oliver J.E., Miller A.L., Troskie H.E., Wilson K.L., Forsyth J.T., Berg M.C., Cockrem K., McMillan V., Tervit H.R., Wells D.N. Cloned cattle derived from a novel zona-free embryo reconstruction system. *Cloning Stem Cells.* 2003;5(1):3-12. DOI 10.1089/153623003321512111.
- Onishi A., Iwamoto M., Akita T., Mikawa S., Takeda K., Awata T., Hanada H., Perry A.C. Pig cloning by microinjection of fetal fibroblast nuclei. *Science.* 2000;289:1188-1190. DOI 10.1126/science.289.5482.1188.
- Park H.J., Koo O.J., Kwon D.K., Kang J.T., Jang G., Lee B.C. Effect of roscovitine-treated donor cells on development of porcine cloned embryos. *Reprod. Domest. Anim.* 2010;45(6):1082-1088. DOI 10.1111/j.1439-0531.2009.01499.x.
- Peng H., Liu F.J., Zhang X.F., Zhuang Y.F., Wang X.A., Li H.X., Hong Z.Y., Lin X.J., Zhang W.C. Study of electro-fusion/activation in somatic cell nuclear transfer to obtain cloned putian black pig embryos. *J. Anim. Vet. Adv.* 2013;12(4):497-503. DOI 10.3923/javaa.2013.497.503.
- Polejaeva I.A., Chen S.H., Vaught T.D., Page R.L., Mullins J., Ball S., Dai Y., Boone J., Walker S., Ayares D.L., Colman A., Campbell K.H. Cloned pigs produced by nuclear transfer from adult somatic cells. *Nature.* 2000;407:8690. DOI 10.1038/35024082.
- Popova E., Bader M., Krivokharchenko A. Efficient production of nuclear transferred rat embryos by modified methods of reconstruction. *Mol. Reprod. Dev.* 2009;76:208-216. DOI 10.1002/mrd.20944.
- Presicce G.A., Yang X.Z. Parthenogenetic development of bovine oocytes matured *in vitro* for 24 hr and activated by ethanol and cycloheximide. *Mol. Reprod. Dev.* 1994;38:380-385. DOI 10.1002/mrd.1080380405.
- Rim C.H., Fu Z., Bao L., Chen H., Zhang D., Luo Q., Ri H.C., Huang H., Luan Z., Zhang Y., Cui C., Xiao L., Jong U.M. The effect of the number of transferred embryos, the interval between nuclear transfer and embryo transfer, and the transfer pattern on pig cloning efficiency. *Anim. Reprod. Sci.* 2013;143(1-4):91-96. DOI 10.1016/j.anireprosci.2013.10.004.
- Robl J.M., Prather R.S., Branes F., Eyestone W., Northey D., Gilligan B., First N.L. Nuclear transplantation in bovine embryos. *J. Anim. Sci.* 1987;64:642-647.
- Samiec M., Skrzyszowska M. The use of different methods of oocyte activation for generation of porcine fibroblast cell nuclear transferred embryos. *Ann. Anim. Sci.* 2010;10(4):399-411.
- Savard C., Novak S., Saint-Cyr A., Moreau M., Pothier F., Sirard M.A. Comparison of bulk enucleation methods for porcine oocytes. *Mol. Reprod. Dev.* 2004;67(1):70-76. DOI 10.1002/mrd.20011.
- Shi J., Zhou R., Luo L., Mai R., Zeng H., He X., Liu D., Zeng F., Cai G., Ji H., Tang F., Wang Q., Wu Z., Li Z. Influence of embryo handling and transfer method on pig cloning efficiency. *Anim. Reprod. Sci.* 2015;154:121-127. DOI 10.1016/j.anireprosci.2015.01.006.
- Simões R., Santos A.R., Jr. Factors and molecules that could impact cell differentiation in the embryo generated by nuclear transfer. *Organogenesis.* 2017;13(4):156-178. DOI 10.1080/15476278.2017.1389367.
- Singina G.N., Volkova N.A., Bagirov V.A., Zinovieva N.A. Cryobanking of somatic cells in conservation of animal genetic resources: prospects and successes. *Selskokhozyaistvennaya Biologiya = Agricultural Biology.* 2014;6:3-14. DOI 10.15389/agrobiology.2014.6.3eng.
- Skrzyszowska M., Samiec M., Słomski R., Lipiński D., Mały E. Development of porcine transgenic nuclear-transferred embryos derived from fibroblast cells transfected by the novel technique of nucleofection or standard lipofection. *Theriogenology.* 2008;70:248-259. DOI 10.1016/j.theriogenology.2008.04.007.
- Sugimura S., Yamanaka K., Kawahara M., Wakai T., Yokoo M., Sato E. Early metaphase II oocytes treated with dibutyl cyclic adenosine monophosphate provide suitable recipient cytoplasm for the production of miniature pig somatic cell nuclear transfer embryos. *Anim. Sci. J.* 2010;81(1):48-57. DOI 10.1111/j.1740-0929.2009.00705.x.
- Suzuki C., Iwamura S., Yoshioka K. Birth of piglets through the non-surgical transfer of blastocysts produced *in vitro*. *Reprod. Dev.* 2004;50(4):487-491.
- Tatham B.G., Dowsing A.T., Trounson A.O. Enucleation by centrifugation of *in vitro*-matured bovine oocytes for use in nuclear transfer. *Biol. Reprod.* 1995;53:1088-1094. DOI 10.1095/biolreprod53.5.1088.
- Uhm S.J., Gupta M.K., Chung H.J., Kim J.H., Park C., Lee H.T. Relationship between developmental ability and cell number of Day 2

- porcine embryos produced by parthenogenesis or somatic cell nuclear transfer. *Asian-Australas. J. Anim. Sci.* 2009;22(4):483-491. DOI 10.5713/ajas.2009.80362.
- Vacková I., Engelová M., Marinov I., Tománek M. Cell cycle synchronization of porcine granulosa cells in G1 stage with mimosine. *Anim. Reprod. Sci.* 2003;77(3-4):235-245. DOI 10.1016/S0378-4320(03)00034-4.
- Vajta G., Kragh P.M., Mtango N.R., Callesen H. Hand-made cloning approach: potentials and limitations. *Reprod. Fertil. Dev.* 2005; 17(1-2):97-112. DOI 10.1071/RD04116.
- Vajta G., Lewis I.M., Hyttel P., Thouas G., Trounson A. Somatic cell cloning without micromanipulators. *Cloning.* 2001;3:89-95. DOI 10.1089/15204550152475590.
- Verma P.J., Du Z.T., Crocker L., Faast R., Grupen C.G., McIlfatrick S.M., Ashman R.J., Lyons I.G., Nottle M.B. In vitro development of porcine nuclear transfer embryos constructed using fetal fibroblasts. *Mol. Reprod. Dev.* 2000;57(3):262-226. DOI 10.1002/1098-2795(200011)57:3<262::AID-MRD8>3.0.CO;2-X.
- Whitworth K.M., Li R., Spate L.D., Wax D.M., Rieke A., Whyte J.J., Manandhar G., Sutovsky M., Green J.A., Sutovsky P., Prather R.S. Method of oocyte activation affects cloning efficiency in pigs. *Mol. Reprod. Dev.* 2009;76:490-500. DOI 10.1002/mrd.20987.
- Wilmot I., Schnieke A.E., McWhir J., Kind A.J., Campbell K.H. Viable offspring derived from fetal and adult mammalian cells. *Nature.* 1997;385:810-813. DOI 10.1089/clo.2006.0002.
- Wimmers K., Murani E., Ponsuksili S. Functional genomics and genetical genomics approaches towards elucidating networks of genes affecting meat performance in pigs. *Brief. Funct. Genomics.* 2010;9: 251-258. DOI 10.1093/bfpg/elq003.
- Yamanaka K., Sugimura S., Wakai T., Kawahara M., Sato E. Difference in sensitivity to culture condition between *in vitro* fertilized and somatic cell embryos in pigs. *Reprod. Dev.* 2009;55(3):299-304. DOI 10.1262/jrd.20174.
- Yang F., Hao R., Kessler B., Brem G., Wolf E., Zakhartchenko V. Rabbit somatic cell cloning: effects of donor cell type, histone acetylation status and hymeric embryo-complementation. *Reproduction.* 2007;133(1):219-230. DOI 10.1530/rep.1.01206.
- Yin X.J., Tani T., Yonemura I., Kawakami M., Miyamoto K., Hasegawa R., Kato Y., Tsunoda Y. Production of cloned pigs from adult somatic cells by chemically assisted removal of maternal chromosomes. *Biol. Reprod.* 2002;67:442-446. DOI 10.1095/biolreprod.67.2.442.
- Yoshioka K., Suzuki C., Tanaka A., Anas I.M., Iwamura S. Birth of piglets derived from porcine zygotes cultured in a chemically defined medium. *Biol. Reprod.* 2002;66(1):112-119.
- Zhang Y., Pan D., Sun X., Sun G., Wang X., Liu X., Li Y., Dai Y., Li N. Production of porcine cloned transgenic embryos expressing green fluorescent protein by somatic cell nuclear transfer. *Sci. China C. Life Sci.* 2006;49(2):164-171. DOI 10.1007/s11427-005-0071-5.
- Zheng Y.M., Zhao H.Y., Zhao X.E., Quan F.S., Hua S., He X.Y., Liu J., He X.N., Lin H. Development of cloned embryos from porcine neural stem cells and amniotic fluid-derived stem cells transfected with enhanced green fluorescence protein gene. *Reproduction.* 2009; 137(5):793-801. DOI 10.1530/REP-08-0469.
- Zhu J., Telfer E.E., Fletcher J., Springbett A., Dobrinsky J.R., De Sousa P.A., Wilmot I. Improvement of an electrical activation protocol for porcine oocytes. *Biol. Reprod.* 2002;66(3):635-641. DOI 10.1095/biolreprod66.3.635.

ORCID ID

A.V. Lopukhov orcid.org/0000-0002-1284-1486
G.N. Singina orcid.org/0000-0003-0198-9757
N.A. Zinovieva orcid.org/0000-0003-4017-6863

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

Received February 5, 2019. Revised May 20, 2019. Accepted May 20, 2019.

Prospects for marker-associated selection in tomato *Solanum lycopersicum* L.

A.B. Shcherban

Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
✉ e-mail: atos@bionet.nsc.ru

The review gives a brief description of tomato, one of the main objects of olericulture for Siberia. The data on the main directions in the breeding of this culture, such as resistance to various pathogens, the nutritional properties of fruits, the timing of their maturation and storage are generalized. A separate chapter is devoted to the use of various types of DNA markers for constructing detailed genetic maps of the specified object, which, along with full-genome sequencing data, can be used to screen for genes responsible for breeding traits. Most of these traits, especially specific resistance to one or another pathogen, were transferred to the cultivated tomato by crossing with wild species, therefore, special attention was paid in the article to identifying and marking resistance genes to a variety of viral, fungal and bacterial pathogens occurring in Western Siberia and adjacent areas. Another important aspect for breeding is the nutrient content of tomato fruits, including carotenoids, vitamins, sugars, organic acids, etc. Recently, due to modern technologies of sequencing, SNP-genotyping, the development of new bioinformatic approaches, it has become possible to establish genetic cascades determining the biochemical composition of tomato fruits, to identify key genes that can be used in the future for marker-associated selection of nutritional value. And, finally, genetic works devoted to the problem of the optimal dates of fruit ripening in certain climatic conditions and their prolonged storage without loss of quality are discussed.

Key words: tomato; selection; DNA marker; pathogen; resistance; ripening time; shelf life.

For citation: Shcherban A.B. Prospects for marker-associated selection in tomato *Solanum lycopersicum* L. Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding. 2019;23(5):534-541. DOI 10.18699/VJ19.522

Перспективы маркер-ориентированной селекции томата *Solanum lycopersicum* L.

А.Б. Щербань

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: atos@bionet.nsc.ru

В обзоре представлена краткая характеристика одного из основных для Сибири объектов овощеводства – томата. Обобщены данные об основных направлениях селекции этой культуры, таких как устойчивость к различным патогенам, сроки созревания и хранения плодов, а также содержание в них биологически активных веществ (БАВ). Отдельная глава обзора посвящена использованию различных типов маркеров ДНК для построения детальных генетических карт указанного объекта, которые наряду с данными полногеномного секвенирования могут быть использованы для скрининга различных генов, отвечающих за селективируемые признаки. Большинство таких признаков, особенно специфическая устойчивость к тем или иным патогенам, перенесено в культурный томат путем скрещивания его с дикорастущими видами, поэтому особое внимание в статье уделено выявлению и маркированию генов устойчивости к целому ряду вирусных, грибных и бактериальных патогенов, распространенных в Западной Сибири и на прилегающих территориях. Другой важный аспект для селекции – содержание БАВ в плодах томата, включая каротиноиды, витамины, сахара, органические кислоты и др. За последнее время благодаря современным технологиям секвенирования, SNP-генотипирования, разработке новых биоинформатических подходов удалось установить генетические каскады, определяющие биохимический состав плодов томата; выделить ключевые гены, которые в перспективе могут быть использованы в маркер-ориентированной селекции по признакам питательной ценности. И наконец, обсуждаются генетические работы, посвященные весьма актуальной для селекции проблеме оптимального в тех или иных климатических условиях срока созревания плодов и их длительного хранения без потери качества.

Ключевые слова: томат; селекция; ДНК-маркер; патоген; устойчивость; срок созревания; лежкость.

Introduction

Tomato, *Solanum lycopersicum* L., is the second most important vegetable crop after cabbage. It belongs to the family Solanaceae, consisting of approximately 100 genera and

2500 species, including several plants of agronomic importance (potato, eggplant, pepper, tobacco). In 2012, due to the efforts of the International Consortium on sequencing the tomato genome, the genomes of the cultivar Heinz 1706 and

the wild ancestor of the tomato, *Solanum pimpinellifolium* L. were completely sequenced (DOI 10.1038/nature11119). The tomato ($2n = 2x = 24$) has a relatively compact genome of 950 MB. It contains about 35,000 genes and was subjected to two rounds of triploidization (120 million and 70 million years ago) in the course of evolution; the second round took place before the divergence of tomato and potato. It is believed that the process of polyploidization promoted the neofunctionalization of genes responsible for the ripening and chemical composition of fruits, leading to the formation of a fleshy fruit in tomato that is of great importance for the propagation of seeds (Howe, Smallwood, 1982). Sequencing data is available through the SOLGenomics Network (SGN) website (<http://solgenomics.net>). Tomato fruits are enriched with vitamins A and C, a number of minerals and other biologically active substances (BAS), including lycopene, which belongs to antioxidants (Rao A.V., Rao L.G., 2007).

The homeland of tomato is South America, where its wild and semi-cultural forms are still found. In the middle of the XVI century, the tomato came through Spain and Portugal to Europe and was first used as an ornamental plant, since the fruits of the tomato were considered inedible. At the end of the XVIII century, a tomato appeared in Russia and was also first cultivated for decorative purposes. The tomato became a vegetable crop thanks to the agronomist scientist A.T. Bolotov, who developed a seedling method of cultivation and a method of ripening (ripening of green fruits after their collection).

DNA markers

Currently, the presence of complete genomic sequences (see above) makes it possible to effectively search for various genes responsible for valuable traits, as well as corresponding DNA markers for marker-assisted selection (MAS) of new forms of tomato. A large number of these markers were developed, including: RFLP (restriction fragment length polymorphism) (Tanksley et al., 1992), as well as PCR markers, including RAPD (randomly amplified polymorphic DNA), AFLP (amplified length polymorphism fragments), SSR (simple repeating sequences) (Saliba-Colombani et al., 2000; Ohyama et al., 2009). To date, SNP (single nucleotide polymorphism) markers are the technology of choice, and within this technology methodological approaches have been successfully approved on tomato such as using of EST SNP analysis for high-performance genotyping (Shirasawa et al., 2010), wide-scale genomic sequencing to identify SNPs that affect protein functions (<http://plant1.kazusa.or.jp/tomato/>). Polymorphic markers for tomato genomic selection were developed based on DArT (DNA chip technology for studying diversity) (Van Schalkwyk et al., 2012).

However, it should be noted that, despite the many DNA markers developed, mainly markers for qualitative traits, such as specific resistance to pathogens, are currently used in practical breeding of tomato. As for quantitative traits (QTL), so far the use of appropriate markers is hindered by their weak linkage with these traits, low polymorphism, undesirable pleiotropic effects, and the lack of validation on diverse material of lines and varieties (Foolad, Panthee, 2012). In this regard, the problem of developing of new, effective molecular markers suitable for use on a wide range of varieties and populations remains actual.

Main directions of tomato breeding in Western Siberia

Tomato is a thermophilic culture and the climate of Western Siberia does not always favor to its productivity. In addition, the tomato is susceptible to numerous infectious diseases. This implies the need to obtain new varieties and hybrids capable of producing high yields and possessing a set of economically valuable traits, such as resistance to pathogens, ripening date corresponding to a short vegetation period, shelf life, etc. As known, MAS makes it possible to conduct selection for many traits simultaneously and allows significantly (2–3 times) to reduce the time of obtaining new varieties, compared with the classical breeding. However, no one variety or hybrid of tomato has been obtained in Siberia using MAS. In this regard, it seems relevant to summarize the main results obtained in the world on this culture with the help of MAS, focusing on those directions that correspond to the conditions of Western Siberia and adjacent territories.

Tomato resistance to pathogens

Most of the resistance genes were identified within wild-growing species and then, by crossing, were introduced in a cultivated tomato (Foolad, Panthee, 2012). In Siberia, fungal diseases of tomato are in the first place by importance, namely: late blight, leaf mould (in greenhouse), septoria blight (in field), fusarium wilt and verticillium wilt. Bacterial spot and bacterial canker are the most common bacterial diseases. Viral diseases are not so relevant for Siberia, although in some years epiphytotics occur.

Resistance to fungal diseases

Late blight caused by *Phytophthora infestans* oomycete, is one of the most devastating diseases of tomato in regions with high humidity and a cool climate, leading to yield loss up to 100 %. Losses can be in the form of a drop in yield, a lower quality of fruits, for example, a low specific weight, a decrease in shelf life, etc. Due to the large economic effect, the pathology and genetics of this disease have been intensively studied for many years. Three main resistance genes were identified in wild-growing tomato *S. pimpinellifolium*: *Ph-1*, *Ph-2* and *Ph-3*, which were mapped on chromosomes 7, 10 and 9, respectively (Black et al., 1996; Moreau et al., 1998). The strongest resistance gene, *Ph-3*, provides incomplete dominant resistance to a wide range of *P. infestans* isolates (Chunwongse et al., 2002). Analysis of its primary structure showed that it encodes a CC-NBS-LRR (coiled-coil nucleotide-binding leucine-rich repeat) – protein that belongs to the extensive NBS-LRR class of plant R-genes (Zhang et al., 2014). However, even this gene does not provide resistance to *Phytophthora* most aggressive isolates. In these cases, the most effective was the combination of two genes, *Ph-2* and *Ph-3*, which were successfully transferred to a number of commercial varieties using developed CAPS markers (Robbins et al., 2010; Zhang et al., 2014). Work on the isolation and analysis of new late blight resistance genes continues. In particular, a number of QTLs carrying resistance genes have been identified that have not yet been precisely localized (Merk, Foolad, 2012; Panthee et al., 2017).

Fusarium wilt. *Fusarium oxysporum* is a soil fungus that causes wilting disease in tomato. It affects all plant tissues and

can persist for a long time in the form of chlamydospores in the soil and plant residues, without losing virulence. Currently three races of this fungus were identified; in Russia, race 1 brings the most damage in greenhouses, race 2 occurs in some farms (Ignatova, 2001). Gene *I*, which provides high resistance to race 1, and gene *I-2*, which gives resistance to races 1 and 2, were mapped on the short and long arms of chromosome 11, respectively (Ori et al., 1997; Scott et al., 2004). These genes were most often used in breeding for resistance to *Fusarium*, however, recently, race 3 has become very common and the corresponding resistance gene has been mapped in detail on chromosome 7 (Lim et al., 2008). There are various linked PCR markers for each of the three genes; markers of resistance to races 1 and 3 are most effective (Barillas et al., 2008; Arens et al., 2010).

A kind of *Fusarium* wilt – *Fusarium* root rot, caused by another strain of *F. oxysporum*. Resistance was established in the induced mutant *S. peruvianum* and the only resistance gene *Fr 1* was mapped on chromosome 9 near the *Tm-2²* gene (Vakalounakis et al., 1997). Subsequently, RAPD markers for *Fr 1* (Tanyolac, Akkale, 2010) were developed, however, to date there are few commercial varieties and lines resistant to this disease.

Leaf mould is common in almost all the world and most often affects plants in greenhouse conditions. Affected leaves, flowers and young fruits turn yellow and then dry. The pathogenic agent is *Cladosporium fulvum*, a highly contagious, optional saprotroph. More than 20 major resistance genes have been identified and mapped on different chromosomes (Wang et al., 2007). In Russia, the most effective resistance genes *Cf-2*, *Cf-5*, *Cf-6*, *Cf-9* give resistance to races of the fungus 1, 3 and 4, however, due to the appearance of new races, at least two genes must be combined (Ignatova, 2001). Although a number of PCR markers have been associated with *Cf* genes (Grushetskaya et al., 2007; Wang et al., 2007; Truong et al., 2011), there is no data on their use in breeding.

Verticillium wilt is a widespread disease characterized by the following symptoms: wilting, discoloration and leaf fall, vascular tissues and root system necrosis. *Verticillium* wilt is caused by *Verticillium dahlia* and *V. albo-atrum*. In tomato, resistance to *Verticillium* is controlled by the *Ve* locus mapped on the short arm of chromosome 9 and consisting of two linked genes *Ve-1* and *Ve-2*, each of which provides resistance to certain pathogen races (Kawchuk et al., 2001; Fradin et al., 2009). PCR markers were obtained to discriminate tolerant and sensitive to *Verticillium* forms of tomato (Acciarri et al., 2007; Arens et al., 2010).

Resistance to bacterial pathogens

Bacterial cancer caused by the rod-shaped bacterium *Clavibacter michiganensis*, is a common tomato disease worldwide and one of the most difficult to control. Infection occurs through mechanically damaged tissues. Greenhouse tomatoes are most at risk. Mapping using crosses between *S. lycopersicum* and the resistant specimen *S. habrochaites* LA 407 allowed to identify and accurately map two large QTLs on chromosomes 2 (*Rcm2.0*) and 5 (*Rcm5.1*), which are responsible for 68 % of expressivity variation (Kabelka et al., 2002; Coaker, Francis, 2004). There are data on markers

(Coaker, Francis, 2004), however, there is no information on their use.

Bacterial spot is a common disease of tomato (especially in Western Siberia, Kazakhstan), which is caused by four species of rod-shaped bacteria *Xanthomonas* (races T1–T5). It is characterized by spotting of leaves, stems and fruits, accompanied by leaf fall, a decrease in the size of fruits and their immaturity, which leads to yield loss up to 100 %. Chemical control is not effective enough due to the development of resistance in the pathogen and multiple ways of its inoculation. Pathogen resistance has been found in a number of *S. lycopersicum* specimens, as well as in wild species, however, its use is greatly complicated by the diversity of pathogen races and the complex nature of resistance. In many cases, it is characterized by race specificity, but some genotypes exhibit multiple quantitative resistance, depending on external conditions. For example, the resistance of Hawaii 7998 *S. lycopersicum* line to race T1 ranges from reduced field symptoms to a hypersensitivity reaction (HR) in a greenhouse. This reaction is provided by three independent genes *Rx-1*, *Rx-2* (chromosome 1) and *Rx-3* (chromosome 5) (Wang et al., 1994; Yu et al., 1995). The participation of *Rx-3* locus was most reliably confirmed to which markers were developed, including CAPS marker L3-L1, which was used in breeding (Yang, Francis, 2005). The same line has strong HR-resistance to race T3 (both in the field and in the greenhouse), which is controlled by the *Rx-4* gene mapped on chromosome 11 (Wang et al., 2011).

Resistance to viruses

Tomato mosaic virus (ToMV) is one of the most stable viruses; crop losses when infected with ToMV reach 50 % or more. The disease is characterized by the appearance of a motley (mosaic) color of leaves, stems and fruits, followed by their deformation and fading. ToMV is highly contagious and is transmitted via mechanical contact, as well as insects: thrips, aphids, etc. In tomato three major resistance genes were revealed: *Tm-1*, *Tm-2* and *Tm-2²* (Ohmori et al., 1996; Sobir et al., 2000; Scott, 2007). The first gene, localized on chromosome 5, inhibits the synthesis of viral RNA by suppressing viral RNA replicase (Meshi et al., 1988). The *Tm-2* and *Tm-2²* genes, localized on chromosome 9, block the movement of the virus from cell to cell, and also cause HR (Meshi et al., 1989). The highest efficiency is observed when all three dominant genes are combined in the homo- or heterozygous state (Puchalsky, 2007). For each of these, PCR markers were developed (Dax et al., 1998; Sobir et al., 2000; Arens et al., 2010).

Tomato spotted wilt virus. The disease is caused by the tomato spotted wilt virus, TSWV. It leads to a decrease in crop yields (over 50 %) and deterioration in product quality. The TSWV virus has an extremely wide range of host plants, which creates a high risk of infection. Eight major resistance genes are known, including the dominant genes *Sw-1a*, *Sw-1b*, *Sw-5*, *Sw-6* and *Sw-7* and the recessive genes *sw-2*, *sw-3* and *sw-4* (Stevens et al., 1992). The most effective gene for resistance to TSWV, the *Sw-5* gene, is localized on the long arm of chromosome 9, and since it is race-specific, it is often used in practical breeding. However, there is a risk of overcoming *Sw-5* with new TSWV strains; virulence to this

resistance gene has been reported in several countries (Scott, 2007). A large number of PCR markers have been developed to detect Sw-5 (Smiech et al., 2000; Langella et al., 2004; Garland et al., 2005).

Size and color of fruits, content in them biologically active substances

The trait of “uniform ripening” is determined by the genetic locus *uniform ripening* (*u*), which control the amount and distribution of chlorophyll in immature fruits (Bohn, Scott, 1945). The dominant allele *U* determines a normal, uneven maturation, in which the upper part of the immature fruit has a dark green and the lower – a light green color. Plants that are homozygous for the recessive *u* allele (*u/u*) produce uniformly ripening fruits that, in an immature state, have the same pale green color on all sides. The initial breeding led to the selection of such forms of tomato, because they are characterized by a uniform red color of ripe fruit. In 2012, localization of the *U* locus on the short arm of chromosome 10 was established using genetic mapping and the *GLK2* candidate gene was identified that encodes the Golden 2-like transcription factor, a regulator of chloroplast development (Powell et al., 2012). The authors sequenced this gene in varieties with *U/U* and *u/u* genotypes and found that in the first case, the *GLK2* gene encodes a complete regulatory protein of 310 amino acids in length, whereas in the case of the *u* allele, the synthesis of non-functional protein occurs due to premature stop codon which resulted from insertion of one nucleotide. Using genetic transformation, it was shown that this mutation blocking the *GLK2* gene is responsible for the uniform coloring phenotype and the associated decrease in the number of chloroplasts in fruits. The latter, in turn, leads to a decrease in the level of photosynthesis and a significant decrease in the content of soluble solids in the fruit juice. As a result, the cultural forms of tomato with the *u/u* genotype have lower taste and nutritional qualities, compared with the ancestral forms. In 2017, the Science published an article of D. Tieman et al. (2017), in which more than 300 modern and traditional tomato varieties were analyzed using genomic sequencing and chemical analysis. In this work, 28 compounds were identified that are responsible for the organoleptic qualities of tomato and then, based on the genome-wide analysis of associations (GWAS), a search was made for SNPs associated with the concentration of these chemical compounds. As a result, several major genes were identified that are responsible for the tomato flavor. Thus, the *Lin5* gene encodes an extracellular invertase that catalyzes the hydrolysis of sucrose to low molecular weight glucose and fructose. Alleles of this gene that are responsible for the alternative characteristics of modern and wild/old-fashioned varieties (low sugar content, large fruits vs. high content, small fruits) differ by only one SNP, leading to the substitution Asn→Asp. Another example, the *E8* gene, which regulates the synthesis of ethylene, hormone of maturation. In the overwhelming majority of modern varieties, this hormone has an increased activity, which leads to a higher concentration of methyl salicylate and guaiacol with an unpleasant smell, compared to the old varieties, while the “beneficial” aromatic substances are less concentrated. Three SNPs were identified in the regulatory regions of the *E8* gene, which appear to be responsible for

the indicated differences (Tieman et al., 2017).

The most important BAS of tomato fruits include carotenoids, a class of 40-carbon hydrocarbons, which are represented by orange, red and yellow pigments synthesized in various plant organs. These substances are involved in a variety of physiological processes of growth, development of plants, reactions to external stimuli. To date, the biosynthesis genes, as well as transcription factors and hormones that regulate the metabolism of carotenoids under the influence of external factors, have been established (Liu et al., 2015). In particular, key regulatory genes that determine the concentration of lycopene, the most common carotenoid-antioxidant of ripe tomatoes, have been identified. This substance is considered as an important biologically active component of the human diet, reducing the risk of cancer and cardiovascular diseases (Ford, Erdman, 2012). Recently, using genomic editing, the synthesis of lycopene in tomato fruits has been increased five times due to the knockout of genes responsible for the conversion of lycopene to β - and α -carotene (Li et al., 2018).

Specific polymorphisms that are responsible for particular varietal characteristics of the tomato fruit color were identified. The formation of a dark red color in the Black Cherry variety is caused by a mutation of the reading frame shift in the coding part of the lycopene- β -cyclase gene, leading to a loss of protein function. A similar mutation leading to a stop codon and shortened protein Psy 1 phytoene synthase underlies the yellow color of fruits (Aflitos et al., 2014).

The shape and size of the tomato fruit correlates with the number of seed chambers (locules). Two QTLs, *lc* and *fas*, have the maximum effect on these traits and can act synergistically, leading to an extremely high number of locules (Cong et al., 2008; Munos et al., 2011). *Fas* is the strongest gene (variation in the number of locules 2 more than 6), while *lc* acts weaker (3–4 locules). Two SNPs, T→C and A→G, are associated with the allele *lc^h* of a high number of locules. Analysis of the primary structure of the *lc* gene showed that all 2-chamber tomato varieties have the *lc^l* allele, and the 3, 4-chamber – allele *lc^h*. The *Fas* gene encodes a YABBY-like transcription factor (Cong et al., 2008). The *fas^h* allele appeared as a result of the inversion of the 294 kbp region on chromosome 11, that led to the shutdown of the *Fas* gene due to the spatial separation of exons 1 and 2 (Huang, van der Knaap, 2011).

Peculiarities of the formation of plants and fruit ripening

Determinancy. For greenhouse conditions tomato plants of an indeterminate type are most suitable. They are characterized by continuous growth and uniform ripening of fruits for several months. For field conditions of Siberia determinant genotypes are more acceptable, the main distinguishing feature of which is termination of shoot growth after the formation of 2–6 inflorescences. Such genotypes, as a rule, are early maturing, which prevents yield loss due to the short growing season.

Determinancy is controlled by the *SP* regulatory gene (*SELF PRUNING*), which controls the transition from the vegetative to the generative stage of development and is homologous to *FT* (*FLOWERING LOCUS T*) – gene of *Arabidopsis* (Pnueli et al., 1998). Determinant plants have the *sp/sp* genotype,

indeterminate – (SP–). There are at least six *SP* genes in the tomato genome. For one of them, *SP5G*, a mechanism of action was established that depends on photoperiod (Soyk et al., 2017). Like the *FT* gene, *SP5G* belongs to the flowering repressors. Under the influence of a long day, its expression is induced to a high level, which leads to suppression of flowering until the onset of a short day (indeterminant, wild phenotype). In a cultural tomato of determinant type, this effect of a long day on expression is reduced due to mutations in this gene. Using the CRISPR/Cas9 genomic editing, it was possible to obtain the null allele *SP5G* and thereby restore a determinant phenotype characterized by early flowering and increased productivity (Soyk et al., 2017).

Genes of slow ripening of fruits. Earlier, the pleiotropic genes responsible for the delayed fruit ripening period were revealed in tomato: *alcobaca* (*alc*), ripening inhibitor (*rin*) and non-ripening (*nor*) (Garg et al., 2008). In plants carrying these genes in a homozygous state, shelf life of fruits increased by 250–500 %; meanwhile they were less prone to the process of decay. However, such genotypes did not become widespread in commerce, due to the accompanying traits: pale coloring and poor taste. The fruits of heterozygous plants also had an increased shelf life (average between parental forms), resistance to decay, but at the same time they had acceptable color and taste for consumers. In addition, these plants had an increased yield, and such indicators as: the content of lycopene and dry matter, fruit consistency, ascorbic acid content were intermediate compared to their parents. As a result, the forms carrying the *alc*, *nor*, and *rin* genes are widely used in commercial tomato varieties in many countries (Garg et al., 2008).

In 2002 Science published an article devoted to the *rin* gene (Vrebalov et al., 2002). This gene is located on the short arm of chromosome 5 and encodes a MADS-box-transcription factor that regulates many different developmental genes, including those associated with ethylene biosynthesis. The *alc* and *nor* genes were also cloned and analyzed (Moore et al., 2002). The *alc* gene (synonym: *DFD*, delayed fruit deterioration) has several advantages for breeding, since it has a lower negative effect on fruit quality, color, aromatic properties and resistance to bacterial diseases (Garg et al., 2008). The *alc* recessive mutation is caused by a nonsynonymous T→A substitution at position 317 of the coding sequence, leading to the Val→Asp substitution (Casals et al., 2012). Using CRISPR/Cas9 in one of the varieties, the *ALC* allele was replaced by the *alc* allele by homologous recombination (Yu et al., 2017).

Functional male sterility

The low genetic diversity due to the mode of tomato reproduction (self-pollination) and the effect of the bottle neck during the introduction process make the successful breeding of tomato very difficult. The English scientist Ch. Rick first began to use the methods of introgression of genetic material from wild-growing to cultivated tomato (Rick, 1960) and most of the tomato varieties were obtained using hybridization.

In tomato, the production of hybrid seeds is laborious due to the need for isolation and castration of flowers, so the use of lines with the trait of functional male sterility (FMS) is the most effective way to obtain hybrid seeds. FMS is

caused by deviations in the development of the flower and in tomato includes the following types: ex, ex-2, ps, ps-2 (Kuzemensky, 2004). The latter type is most widely used in tomato breeding. The stamens of plants of the ps-2 type have the usual structure, fertile pollen grains, but the anthers are not opened. The *Ps-2* gene controlling this type of sterility was identified in chromosome 4, isolated and its primary structure was studied (Gorguet et al., 2009). It encodes the enzyme polygalacturonase, which affects the rigidity of the cell wall by digestion of pectins. The single mutation that disrupts splicing of mRNA, resulting in its aberrant forms is responsible for the ps2 phenotype. A number of markers have been developed for the *Ps-2* gene: SNP (Gorguet et al., 2009), CAPS (Staniaszek et al., 2012), etc.

Conclusion

The work on the complete sequencing of the tomato genome and the construction of high-resolution genetic maps laid the foundation for a fast and effective search for genes responsible for important selection traits, as well as the development of DNA markers corresponding to these genes that can be used in marker-assisted selection of a new forms of tomato. Especially relevant for a temperate climate are markers of such traits as resistance to a number of common pathogens of various nature, valuable biologically active substances, for example, carotenoids, lycopene, sugars, etc., as well as gene markers that determine the optimal, early fruit ripening in conditions of short summer period and risk of autumn frosts. To date, key genes responsible for these traits have been identified and characterized, which makes it possible, on the basis of molecular markers, to develop strategies for crossing and selection for these genes, to perform their pyramiding, as well as targeted modification using modern genomic editing methods.

References

- Acciarri N., Rotino G.L., Tamiotti G., Valentino D., Voltattorni S., Sabatini E. Molecular markers for *Ve1* and *Ve2* *Verticillium* resistance genes from Italian tomato germplasm. *Plant Breed.* 2007;126:617-621. DOI 10.1111/j.1439-0523.2007.01398.x.
- Aflitos S., Schijlen E., de Jong H., de Ridder D., Smit S., Finkers R., Wang J., Zhang G., Li N., Mao L., Bakker F., Dirks R., Breit T., Gravendeel B., Huits H., Struss D., Swanson-Wagner R., van Leeuwen H., van Ham R.C., Fito L., Guignier L., Sevilla M., Ellul P., Ganko E., Kapur A., Reclus E., de Geus B., van de Geest H., Te Lintel Hekkert B., van Haarst J., Smits L., Koops A., Sanchez-Perez G., van Heusden A.W., Visser R., Quan Z., Min J., Liao L., Wang X., Wang G., Yue Z., Yang X., Xu N., Schranz E., Smets E., Vos R., Rauwerda J., Ursem R., Schuit C., Kerns M., van den Berg J., Vriezen W., Janssen A., Datema E., Jahrman T., Moquet F., Bonnet J., Peters S. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant J.* 2014;80:136-148. DOI 10.1111/tpj.12616.
- Arens P., Mansilla C., Deinum D., Cavellini L., Moretti A., Roland S., van der Schoot H., Calvache D., Ponz F., Collonnier C., Mathis R., Smilde D., Caranta C., Vosman B. Development and evaluation of robust molecular markers linked to disease resistance in tomato for distinctness, uniformity and stability testing. *Theor. Appl. Genet.* 2010;120:655-664. DOI 10.1007/s00122-009-1183-2.

- Barillas A.C., Mejia L., Sanchez-Perez A., Maxwell D.P. CAPS and SCAR markers for detection of I-3 gene introgression for resistance to *Fusarium oxysporium* f. sp. *lycopersici* race 3. Rpt. Tomato Genet. Coop. 2008;58:11-17.
- Black L.L., Wang T.C., Hanson P.M., Chen J.T. Late blight resistance in four wild tomato accessions: effectiveness in diverse locations and inheritance of resistance. *Phytopathology*. 1996;86:S24.
- Bohn G.W., Scott D.H. A second gene for uniform unripe fruit color in the tomato. *J. Hered.* 1945;36(6):169-172.
- Casals J., Pascual L., Cañizares J., Cebolla-Cornejo J., Casañas F., Nuez F. Genetic basis of long shelf life and variability into Penjar tomato. *Genet. Resour. Crop Evol.* 2012;59:219-229. DOI 10.1007/s10722-011-9677-6.
- Chunwongse J., Chunwongse C., Black L., Hanson P. Molecular mapping of the *Ph-3* gene for late blight resistance in tomato. *J. Horticult. Sci. Biotechnol.* 2002;77:281-286. DOI 10.1080/14620316.2002.11511493.
- Coaker G.L., Francis D.M. Mapping, genetic effects, and epistatic interaction of two bacterial canker resistance QTLs from *Lycopersicon hirsutum*. *Theor. Appl. Genet.* 2004;108:1047-1055. DOI 10.1007/s00122-003-1531-6.
- Cong B., Barrero L.S., Tanksley S.D. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nat. Genet.* 2008;40:800-804. DOI 10.1038/ng.144.
- Dax E., Livneh O., Aliskevicius E., Edelbaum O., Kedar N., Gavish N., Milo J., Geffen F., Blumenthal A., Rabinowich H.D., Sela I. A SCAR marker linked to the ToMV resistance gene, *Tm2(2)*, in tomato. *Euphytica*. 1998;101:73-77.
- Foolad M.R., Panthee D.R. Marker-assisted selection in tomato breeding. *Crit. Rev. Plant Sci.* 2012;31(2):93-123. DOI 10.1080/07352689.2011.616057.
- Ford N.A., Erdman J.W. Are lycopene metabolites metabolically active? *Acta Biochim. Pol.* 2012;59:1-4.
- Fradin E.F., Zhang Z., Juarez Ayala J.C., Castroverde C.D., Nazar R.N., Robb J., Liu C.M., Thomma B.P. Genetic dissection of verticillium wilt resistance mediated by tomato *Ve1*. *Plant Physiol.* 2009;150: 320-332. DOI 10.1104/pp.109.136762.
- Garg N., Cheema D.S., Pathak D. Heterosis breeding in tomato involving *rin*, *nor* and *alc* alleles: A review of literature. *Adv. Hort. Sci.* 2008;22(1):54-62.
- Garland S., Sharman M., Persley D., McGrath D. The development of an improved PCR-based marker system for Sw-5, an important TSWV resistance gene of tomato. *Aust. J. Agric. Res.* 2005;56: 285-289.
- Gorguet B., Schipper D., van Lammeren A., Visser R.G.F., van Heusden A.W. *ps-2*, the gene responsible for functional sterility in tomato, due to non-dehiscent anthers, is the result of a mutation in a novel polygalacturonase gene. *Theor. Appl. Genet.* 2009;118:1199-1209. DOI 10.1007/s00122-009-0974-9.
- Grushetskaya Z.E., Lemesh V.A., Poliksenova V.D., Khotyleva L.V. Mapping of the *Cf-6* tomato leaf mould resistance locus using SSR markers. *Russ. J. Genet.* 2007;43:1266-1270. DOI 10.1134/S1022795407110099.
- Howe H.F., Smallwood J. Ecology of seed dispersal. *Annu. Rev. Ecol. Syst.* 1982;13:201-228.
- Huang Z., van der Knaap E. Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. *Theor. Appl. Genet.* 2011; 123(3):465-474. DOI 10.1007/s00122-011-1599-3.
- Ignatova S.I. The role of tomato hereditary potential for resistance in the system of integrated protection in protected ground. *Gavrish.* 2001;6:18-20. (in Russian)
- Kabelka E., Franchino B., Francis D.M. Two loci from *Lycopersicon hirsutum* LA407 confer resistance to strains of *Clavibacter michiganensis* subsp. *michiganensis*. *Phytopathology*. 2002;92:504-510. DOI 10.1094/PHYTO.2002.92.5.504.
- Kawchuk L.M., Lynch D.R., Hachey J., Bains P.S., Kulcsar F. Identification of a codominant amplified polymorphic DNA marker linked to the verticillium wilt resistance gene in tomato. *Theor. Appl. Genet.* 1994;89:661-664.
- Kuzemsky A.V. Studies of Mutant Forms of Tomato with Regard to Breeding. Kharkov, 2004. (in Russian)
- Langella R., Ercolano M.R., Monti L.M., Frusciant L., Barone A. Molecular marker assisted transfer of resistance to TSWV in tomato elite lines. *J. Horticult. Sci. Biotechnol.* 2004;79:806-810. DOI 10.1080/14620316.2004.11511846.
- Li X., Wang Y., Chen S., Tian H., Fu D., Zhu B., Luo Y., Zhu H. Lycopene is enriched in tomato fruit by CRISPR/Cas9-mediated multiplex genome editing. *Front. Plant Sci.* 2018;9:1-12. DOI 10.3389/fpls.2018.00559.
- Lim G., Wang G.P., Hemming M., McGrath D.J., Jones D.A. High resolution genetic and physical mapping of the I-3 region of tomato chromosome 7 reveals almost continuous microsynteny with grape chromosome 12 but interspersed microsynteny with duplications on *Arabidopsis* chromosomes 1, 2 and 3. *Theor. Appl. Genet.* 2008; 118:57-75. DOI 10.1007/s00122-008-0876-2.
- Liu L., Shao S.Z., Zhang Z.M. Regulation of carotenoid metabolism in tomato. *Mol. Plant.* 2015;8:28-39. DOI 10.1016/j.molp.2014.11.006.
- Merk H.L., Foolad M.R. Parent-offspring correlation estimate of heritability for late blight resistance conferred by an accession of the tomato wild species *Solanum pimpinellifolium*. *Plant Breed.* 2012; 131:203-210.
- Meshi T., Motoyoshi F., Adachi A., Watanabe Y., Takamatsu N., Okada Y. Two concomitant base substitutions in the putative replicase genes of tobacco mosaic virus confer the ability to overcome the effects of a tomato resistance gene, *Tin-1*. *EMBO J.* 1988;7:1575-1581.
- Meshi T., Motoyoshi F., Maeda T., Yoshiwoka S., Watanabe H., Okada Y. Mutations in the tobacco mosaic virus 30-kDa protein gene overcome *Tm-2* resistance in tomato. *Plant Cell.* 1989;1:515-522.
- Moore S., Vrebalov J., Payton P., Giovannoni J. Use of genomics tools to isolate key ripening genes and analyse fruit maturation in tomato. *J. Exp. Bot.* 2002;53:2023-2030. DOI 10.1093/jxb/erf057.
- Moreau P., Thoquet P., Olivier J., Laterrot H., Grimsley N. Genetic mapping of *Ph-2*, a single locus controlling partial resistance to *Phytophthora infestans* in tomato. *Mol. Plant-Microbe Interact.* 1998; 11:259-269. DOI 10.1094/MPMI.1998.11.4.259.
- Muños S., Ranc N., Botton E., Bérard A., Rolland S., Duffé P., Carretero Y., Le Paslier M.-C., Delalande C., Bouzayen M., Brunel D., Causse M. Increase in tomato locule number is controlled by two single nucleotide polymorphisms located near *WUSCHEL*. *Plant Physiol.* 2011;156:2244-2254. DOI 10.1104/pp.111.173997.
- Ohmori T., Murata M., Motoyoshi F. Molecular characterization of RAPD and SCAR markers linked to the *Tm-1* locus in tomato. *Theor. Appl. Genet.* 1996;92:151-156. DOI 10.1007/BF00223369.
- Ohyama A., Asamizu E., Negoro S., Miyatake K., Yamaguchi H., Tabata S., Fukuoka H. Characterization of tomato SSR markers developed using BAC-end and cDNA sequences from genome databases. *Mol. Breed.* 2009;23(4):685-691. DOI 10.1007/s11032-009-9265-z.

- Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S., Zamir D., Fluhr R. The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell*. 1997;9:521-532. DOI 10.1105/tpc.9.4.521.
- Panthee D.R., Piotrowski A., Ibrahim R. Mapping quantitative trait loci (QTL) for resistance to late blight in tomato. *Int. J. Mol. Sci.* 2017;18:1589. DOI 10.3390/ijms18071589.
- Pnueli L., Carmel-Goren L., Hareven D., Gutfinger T., Alvarez J., Ganai M., Zamir D., Lifschitz E. The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. *Development*. 1998; 125:1979-1989.
- Powell A.L.T., Nguyen C.V., Hill T., Cheng K.L., Figueroa-Balderas R., Aktas H., Ashrafi H., Pons C., Fernández-Muñoz R., Vicente A., Lopez-Baltazar J., Barry C.S., Liu Y., Chetelat R., Granell A., Van Deynze A., Giovannoni J., Bennett A.B. Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science*. 2012;336:1711-1715. DOI 10.1126/science.1222218.
- Pukhalskij V.A., Odintsova T.I., Izvekova L.I., Andreeva E.N., Korostyleva T.I., Istomina E.A., Slavokhotova A.A., Shiyani A.N., Kozlovskaya G.V., Obolenkova L.A., Badaeva E.D., Bilinskaya E.N. The problems of natural and induced immunity in plants. To the development of ideas N.I. Vavilova. *Informatsionny Vestnik VOGIS = The Herald of Vavilov Society for Geneticists and Breeders*. 2007;11(3/4):631-649. (in Russian)
- Rao A.V., Rao L.G. Carotenoids and human health. *Pharmacol. Res.* 2007;55:207-216. DOI 10.1016/j.phrs.2007.01.012.
- Rick C.M. Hybridization between *Lycopersicon esculentum* and *Solanum pennellii*: phylogenetic and cytogenetic significance. *Proc. Natl. Acad. Sci. USA*. 1960;46:78-82.
- Robbins M.D., Masud M.A.T., Panthee D.R., Gardner R.G., Francis D., Stevens M.R. Marker-assisted selection for coupling phase resistance to tomato spotted wilt virus and *Phytophthora infestans* (late blight) in tomato. *HortScience*. 2010;45:1424-1428.
- Saliba-Colombani V., Causse M., Gervais L., Philouze J. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome*. 2000;43:29-40. DOI 10.1139/g99-096.
- Scott J.W. Breeding for resistance to viral pathogens. Eds. M.K. Razdan, A.K. Mattoo. *Genetic Improvement of Solanaceous Crops. Vol. 2. Tomato*. Enfield: Science Publishers, 2007;457-485.
- Scott J.W., Agrama H.A., Jones J.P. RFLP-based analysis of recombination among resistance genes to *Fusarium* wilt races 1, 2, and 3 in tomato. *J. Am. Soc. Hort. Sci.* 2004;129:394-400.
- Shirasawa K., Isobe S., Hirakawa H., Nakamura Y., Sato S., Tabata S. SNP discovery and linkage map construction in cultivated tomato. *DNA Res.* 2010;17(6):381-391. DOI 10.1093/dnares/dsq024.
- Smiech M., Rusinowski Z., Malepszy S., Niemirówicz-Szczytt K. New RAPD markers of tomato spotted wilt virus (TSWV) resistance in *Lycopersicon esculentum* Mill. *Acta Physiol. Plantarum*. 2000;22: 299-303.
- Sobir O.T., Murata M., Motoyoshi F. Molecular characterization of the SCAR markers tightly linked to the *Tm-2* locus of the genus *Lycopersicon*. *Theor. Appl. Genet.* 2000;101:64-69.
- Soyk S., Müller N.A., Park S.J., Schmalenbach I., Jiang K., Hayama R., Zhang L., Van Eck J., Jiménez-Gómez J.M., Lippman Z.B. Variation in the flowering gene *SELF PRUNING* 5G promotes day-neutrality and early yield in tomato. *Nat. Genet.* 2017;49:162-168. DOI 10.1038/ng.3733.
- Staniaszek M., Szajko K., Kozik E.U., Nowakowska M., Marczewski W. The novel *ps* and *ps-2* specific markers for selection of functional male sterile tomato lines in breeding programs and hybrids seed production. *J. Agr. Sci.* 2012;4(10):61-67. DOI 10.5539/jas.v4n10p61.
- Stevens M.R., Scott S.J., Gergerich R.C. Inheritance of gene for resistance to tomato spotted wilt virus (TSWV) from *Lycopersicon peruvianum* Mill. *Euphytica*. 1992;59:9-17. DOI 10.1007/BF00025356.
- Tanksley S.D., Ganai M.W., Prince J.P., de Vicente M.C., Bonierbale M.W., Broun P., Fulton T.M., Giovannoni J.J., Grandillo S., Martin G.B., Messeguer R., Miller J.C., Miller L., Paterson A.H., Pineda O., Riider M.S., Wing R.A., Wu W., Young N.D. High-density molecular linkage maps of the tomato and potato genomes. *Genetics*. 1992;132:1141-1160.
- Tanyolac B., Akkale C. Screening of resistance genes to fusarium root rot and fusarium wilt diseases in F3 family lines of tomato (*Lycopersicon esculentum*) using RAPD and CAPS markers. *Afr. J. Biotech.* 2010;9:2727-2730.
- Tieman D., Zhu G., Resende M.F.R., Lin T., Nguyen C., Bies D., Rambla J.L., Beltran K.S.O., Taylor M., Zhang B., Ikeda H., Liu Z., Fisher J., Zemach I., Monforte A., Zamir D., Granell A., Kirst M., Huang S., Klee H. A chemical genetic roadmap to improved tomato flavor. *Science*. 2017;355(6323):391-394. DOI 10.1126/science.aal1556.
- Truong H.T.H., Choi H.-S., Cho M.C., Lee H.E., Kim J.H. Use of *Cf-9* gene-based markers in marker-assisted selection to screen tomato cultivars with resistance to *Cladosporium fulvum*. *Hort. Environ. Biotechnol.* 2011;52:204-210. DOI 10.1007/s13580-011-0164-y.
- Vakalounakis D.J., Laterrot H., Moretti A., Ligoixakis E.K., Smardas K. Linkage between *Frl* (*Fusarium oxysporum* f. sp. radialis-lycopersici resistance) and *Tm-2* (tobacco mosaic virus resistance-2) loci in tomato (*Lycopersicon esculentum*). *Ann. Appl. Biol.* 1997; 130:319-323. DOI 10.1111/j.1744-7348.1997.tb06835.x.
- Van Schalkwyk A., Wenzl P., Smit S., Lopez-Cobollo R., Kilian A., Bishop G., Hefer C., Berger D.K. Bin mapping of tomato diversity array (DARt) markers to genomic regions of *Solanum lycopersicum* × *Solanum pennellii* introgression lines. *Theor. Appl. Genet.* 2012;124:947-956. DOI 10.1007/s00122-011-1759-5.
- Vrebalov J., Ruezinsky D., Padmanabhan V., White R., Medrano D., Drake R., Schuch W., Giovannoni J. MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (*Rin*) locus. *Science*. 2002;296:343-346. DOI 10.1126/science.1068181.
- Wang A.X., Meng F.J., Xu X.Y. Development of molecular markers linked to *Cladosporium fulvum* resistant gene *Cf-6* in tomato by RAPD and SSR methods. *HortScience*. 2007;42:11-15.
- Wang H., Hutton S.F., Robbins M.D., Sim S.-C., Scott J.W., Yang W., Jones J.B., Francis D.M. Molecular mapping of hypersensitive resistance to race T3 of tomato bacterial spot from Hawaii 7981 maps to chromosome 11. *Phytopathology*. 2011;101:1217-1223. DOI 10.1094/PHYTO-12-10-0345.
- Wang J.F., Jones J.B., Scott J.W., Stall R.E. Several genes in *Lycopersicon esculentum* control hypersensitivity to *Xanthomonas campestris* pv. *vesicatoria*. *Phytopathology*. 1994;84:702-706. DOI 10.1094/Phyto-84-702.
- Yang W., Francis D.M. Marker-assisted selection for combining resistance to bacterial spot and bacterial speck in tomato. *J. Am. Soc. Hort. Sci.* 2005;130:716-721. DOI 10.21273/JASHS.130.5.716.

- Yu Q.-H., Wang B., Li N., Tang Y., Yang S., Yang T., Xu J., Guo C., Yan P., Wang Q., Asmutola P. CRISPR/Cas9-induced targeted mutagenesis and gene replacement to generate long shelf life tomato lines. *Sci. Rep.* 2017;7:11874. DOI 10.1038/s41598-017-12262-1.
- Yu Z.H., Wang J.F., Stall R.E., Vallejos C.E. Genomic localization of tomato genes that control a hypersensitive reaction to *Xanthomonas campestris* pv. *vesicatoria* (Doidge) Dye. *Genetics*. 1995;141:675-682.
- Zhang C., Liu L., Wang X., Vossen J., Li G., Li T., Zheng Z., Gao J., Guo Y., Visser R.G.F., Li J., Bai Y., Du Y. The *Ph-3* gene from *Solanum pimpinellifolium* encodes CC-NBS-LRR protein conferring resistance to *Phytophthora infestans*. *Theor. Appl. Genet.* 2014;127:1353-1364. DOI 10.1007/s00122-014-2303-1.

ORCID ID

A.B. Shcherban orcid.org/0000-0003-1000-8228

Acknowledgements. This work was supported by the project of the Ministry of Education and Science of the Russian Federation No. 0324-2019-0039.

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

Received March 25, 2019. Revised May 17, 2019. Accepted May 20, 2019.

Разнообразие механизмов устойчивости, вовлеченных в многоуровневый иммунитет пшеницы к ржавчинным заболеваниям

Е.С. Сколотнева , Е.А. Салина

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
 e-mail: sk-ska@yandex.ru

Обзор посвящен раскрытию современной концепции фитоиммунитета как многоуровневой системы защиты растения-хозяина, контролируемой комбинациями мажорных и минорных генов (локусов) устойчивости. Подробно разбирается модель «зигзаг» для описания молекулярных основ фитоиммунитета с ключевыми понятиями: ассоциированные с патогенами лиганды, запускающие врожденный иммунитет, дуальность эффекторов, способных вызывать восприимчивость растения, а при взаимодействии с продуктами генов устойчивости включать реакцию сверхчувствительности или альтернативные механизмы защиты. Выделено три различных типа устойчивости у злаков: 1) базовая устойчивость, обеспечиваемая рецепторными белками, локализованными в плазматической мембране; 2) расоспецифическая устойчивость, обеспечиваемая внутриклеточными R-рецепторами иммунного ответа; 3) частичная устойчивость, контролируемая локусами количественных признаков. Система «мягкая пшеница (*Triticum aestivum*) – возбудитель бурой ржавчины (*Puccinia tritica*)» является интересной моделью для наблюдения всех перечисленных механизмов устойчивости, так как стратегия данного патогена направлена на конститутивное использование ресурсов хозяина. Рассмотрены известные гены пшеницы, отвечающие за различные проявления устойчивости к бурой ржавчине: расоспецифические гены (*Lr1*, *Lr10*, *Lr19*, *Lr21*); гены возрастной устойчивости, запускающие реакцию сверхчувствительности (*Lr12*, *Lr13*, *Lr22a*, *Lr22b*, *Lr35*, *Lr48*, *Lr49*); и гены, реализующие альтернативные механизмы частичной устойчивости (*Lr34*, *Lr46*, *Lr67*, *Lr77*). Кроме того, недавно показано участие некоторых R-генов пшеницы в реализации прегаусториальной устойчивости к возбудителю бурой ржавчины: *Lr1*, *Lr3a*, *Lr9*, *LrB*, *Lr19*, *Lr21*, *Lr38*. Наличие в генотипе указанных генов позволяет останавливать ранний патогенез посредством следующих механизмов: дезориентация и ветвление ростковой гифы; формирование aberrантных структур проникновения гриба (аппрессорий, подустычная везикула); аккумуляция каллозы в клеточных стенках мезофилла. Эффективность селекции на иммунитет повышается за счет накопления данных о разнообразных механизмах устойчивости пшеницы к ржавчинным заболеваниям, которые обобщены в данном обзоре. Ключевые слова: пшеница; ржавчинные инфекции; расоспецифическая; частичная; прегаусториальная устойчивость; гены *Lr*; селекция на иммунитет.

Для цитирования: Сколотнева Е.С., Салина Е.А. Разнообразие механизмов устойчивости, вовлеченных в многоуровневый иммунитет пшеницы к ржавчинным заболеваниям. Вавиловский журнал генетики и селекции. 2019;23(5):542-550. DOI 10.18699/VJ19.523

Resistance mechanisms involved in complex immunity of wheat against rust diseases

E.S. Skolotneva , E.A. Salina

Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
 e-mail: sk-ska@yandex.ru

The review is devoted to the disclosure of the modern concept of plant immunity as a hierarchical system of plant host protection, controlled by combinations of major and minor resistance genes (loci). The “zigzag” model is described in detail for discussing the molecular bases of plant immunity with key concepts: pathogen-associated molecular patterns triggering innate immunity, ambivalent effectors causing susceptibility, but when interacting with resistance genes, a hypersensitive reaction or alternative defense mechanisms. There are three types of resistance in cereals: (1) basal resistance provided by plasma membrane-localized receptors proteins; (2) race-specific resistance provided by intracellular immune R-receptors; (3) partial resistance conferred by quantitative gene loci. The system ‘wheat (*Triticum aestivum*) – the fungus causing leaf rust (*Puccinia tritica*)’ is an interesting model for observing all the resistance mechanisms listed above, since the strategy of this pathogen is aimed at the constitutive use of host resources. The review focuses on known wheat genes responsible for various types of resistance to leaf rust: race-specific genes *Lr1*, *Lr10*, *Lr19*, and *Lr21*; adult resistance genes which are hypersensitive *Lr12*, *Lr13*, *Lr22a*, *Lr22b*, *Lr35*, *Lr48*, and *Lr49*; nonhypersensitive genes conferring partial resistance *Lr34*, *Lr46*, *Lr67*, and *Lr77*. The involvement of some wheat R-genes in pre-haustorial resistance to leaf rust has been discovered

recently: *Lr1*, *Lr3a*, *Lr9*, *LrB*, *Lr19*, *Lr21*, *Lr38*. The presence of these genes in the genotype ensures the interruption of early pathogenesis through the following mechanisms: disorientation and branching of the germ tube; formation of aberrant fungal penetration structures (appressorium, substomatal vesicle); accumulation of callose in mesophyll cell walls. Breeding for immunity is accelerated by implementation of data on various mechanisms of wheat resistance to rust diseases, which are summarized in this review.

Key words: wheat; rust diseases; race-specific; partial; pre-haustorial resistance; *Lr* genes; breeding for immunity.

For citation: Skolotneva E.S., Salina E.A. Resistance mechanisms involved in complex immunity of wheat against rust diseases. *Vavilovskii Zhurnal Genetiki i Selekcii* = *Vavilov Journal of Genetics and Breeding*. 2019;23(5):542-550. DOI 10.18699/VJ19.523 (in Russian)

Введение

Фитопатогенные организмы способны быстро адаптироваться к защитной системе растений-хозяев. Особенно интенсивно эволюционные процессы происходят в популяциях патогенных грибов, паразитирующих на сельскохозяйственных культурах и имеющих воздушный способ распространения (Parlevliet, Zadoks, 1977). В первую очередь это относится к возбудителям ржавчинных заболеваний пшеницы. Устойчивость к бурой ржавчине, например, сохраняется у районированных сортов не более 5–6 лет (Горленко, 1968). Однако селекция на иммунитет считается наиболее эффективной и экологичной стратегией защиты растений от болезней, вызываемых фитопатогенными грибами. Поэтому исследование механизмов устойчивости пшеницы к возбудителям ржавчины на генетическом, молекулярном и цитологическом уровне не только имеет фундаментальное значение, но и является необходимым этапом селекции на иммунитет.

Долгое время господствовала концепция принципиально различных типов устойчивости растений к фитопатогенным организмам: вертикальной и горизонтальной (Вандерпланк, 1972). При этом подразумевалось, что вертикальную или расоспецифическую устойчивость определяют гены с качественным проявлением в фенотипе – большие, или главные (*major genes*), а горизонтальную или неспецифическую – малые, или минорные, гены (*minor genes*) со слабым количественным выражением. Параллельно высказывалось предположение, что разделение на две независимые системы (вертикальную и горизонтальную) некорректно, поскольку в определении как расоспецифической, так и неспецифической устойчивости участвуют одни и те же группы генов. Тогда стабильное и длительное сохранение устойчивости зависит не от количества генов, но от характера взаимодействий между генами вирулентности паразита и генами устойчивости хозяина в определенных условиях среды (Parlevliet, Zadoks, 1977; Jonson, 1981; Одинцова, Шеломова, 1983; Дьяков, 2005; Михайлова, 2005).

На сегодняшний день успехи в области изучения фитоиммунитета позволяют выделить три различных типа устойчивости у зерновых растений: 1) базовая устойчивость, обеспечиваемая рецепторными белками, локализованными в плазматической мембране; 2) расоспецифическая устойчивость, обеспечиваемая внутриклеточными рецепторами иммунного ответа; 3) неспецифическая устойчивость, контролируемая локусами количественных признаков. Система «мягкая пшеница (*Triticum aestivum*) – возбудитель бурой ржавчины (*Puccinia tritica*)» является интересной моделью для наблюдения всех перечисленных

механизмов устойчивости, так как стратегия данного патогена направлена на конститутивное использование ресурсов хозяина. В отличие от возбудителей стеблевой и желтой ржавчины (*P. graminis* и *P. striiformis*), которые могут привести неиммунное растение к гибели, воздействие *P. tritica* можно охарактеризовать как относительно мягкое. Поражение бурой ржавчиной локализуется на листьях пшеницы и может значительно снижать эффективность ассимиляции, однако при сбалансированном взаимодействии патогена и хозяина пораженное растение способно развиваться, проходя стадии выколашивания, цветения и созревания.

Существующие представления о молекулярных основах фитоиммунитета, а также многочисленные генетические исследования позволяют рассмотреть разнообразие механизмов проявления мажорных (главных) и минорных генов устойчивости пшеницы к ржавчинным заболеваниям, и в частности к патогену *P. tritica*. В настоящем обзоре обобщены данные о различных механизмах устойчивости пшеницы к ржавчинным заболеваниям, позволяющих повысить эффективность ее селекции на иммунитет.

Молекулярные основы фитоиммунитета

Растения, в отличие от животных организмов, выстраивают взаимодействия с биотическими факторами преимущественно с помощью врожденного иммунитета, на что указывал в своих трудах еще Н.И. Вавилов (1935). Наиболее распространенным или базовым механизмом является процесс узнавания белковыми рецепторами хозяина разнообразных производных патогена. Последние двадцать лет иммунные рецепторы растений стали объектом интенсивных исследований, благодаря чему среди них были выделены две разные категории по локализации относительно клетки: рецепторы плазматической мембраны с экстраклеточными лиганд-связывающими доменами и внутриклеточные иммунные рецепторы (Jones, Dangl, 2006; Dodds, Rathjen, 2010; Thomma et al., 2011; Cook et al., 2015).

Большинство поверхностно локализованных рецепторов реагируют на попавшие в апопласт структуры патогена или продукты патогенных организмов, свойственных широкому кругу патогенов. Ими могут быть, например, компоненты грибной клеточной стенки (хитин, хитозан, β -глюканы) или гидролитические ферменты. Для таких молекулярных структур, сигнализирующих о факте проникновения какого-либо патогена, закрепилась аббревиатура PAMPs (pathogen-associated molecular patterns – молекулярные структуры, ассоциированные с патогенами), которая используется без перевода (Nürnberger,

Kemmerling, 2009). Среди поверхностно локализованных иммунных рецепторов различают рецепторы, лигандами которых служат соединения, специфические одному или нескольким близкородственным видам патогенов (Thomma et al., 2011; Cook et al., 2015), а также рецепторы, связывающиеся с метаболитами самого растения, продуктами инфекционного процесса (мономеры кутина, целлюлозы, пектина). В зависимости от структуры выделяют рецептор-подобные белки (receptor-like proteins – RLP) и рецептор-подобные киназы (receptor-like kinases – RLK) (van Ooijen et al., 2008). Обязательными элементами таких рецепторов являются внеклеточный и трансмембранный домены. Передача сигнала рецептором типа RLK осуществляется путем автофосфорилирования киназного домена, тогда как активированные рецепторы типа RLP взаимодействуют с адаптерными молекулами.

В результате узнавания мембранным рецептором соответствующего лиганда запускаются механизмы врожденного иммунитета (PTI, PAMPs-triggered immunity – иммунитет, запускаемый лигандами PAMP) или базовой устойчивости (basal resistance), например биосинтез защитных молекул (этилена, жасмоновой кислоты). Для ржавчинных грибов, так же как для возбудителя мучнистой росы и некоторых других грибов из класса оомицеты, ключевым этапом патогенеза является формирование гаустория, специализированной структуры, через которую устанавливается контакт с клеткой хозяина. Поэтому применительно к этой группе патогенов в литературе вместо термина «базовая устойчивость» часто используют термин «прегаусториальная устойчивость» (Niks, 1983).

Ростковые трубки *P. tritici* на растениях-хозяевах ориентированы строго перпендикулярно длинной стороне эпидермальных клеток и растут по направлению к устьицам, практически не образуя разветвлений, в то время как на растениях-нехозяевах они образуют утолщения и ответвления, что расходует энергетический потенциал ищущей структуры патогена. На устьицах риса отмечается разрастание аппрессориев возбудителей стеблевой ржавчины пшеницы *P. graminis* и ячменя *P. hordei*. Такие aberrantные аппрессории не способны ассоциировать с устьицами растения-нехозяина (Ayliffe et al., 2011). Базовой или прегаусториальной устойчивостью к возбудителям бурой и стеблевой ржавчины пшеницы обладает ячмень, у которого в ответ на проникновение этих патогенов наблюдается локальное упрочнение клеточных стенок за счет вторичного образования утолщений и выростов, что препятствует продвижению инфекционных гиф по апопласту (Atienza et al., 2004; Niks, Marcel, 2009). Развитие в тканях риса штаммов *P. graminis* f. sp. *tritici*, пшеничной формы стеблевой ржавчины, приводит к отложению каллозы и синтезу перекиси водорода в месте колонизации мицелия (Ayliffe et al., 2008).

Специализированный патоген способен преодолевать барьер врожденного PTI-иммунитета. Например, механизмы прегаусториальной устойчивости ячменя успешно преодолеваются специализированными ржавчинными грибами *P. hordei* и *P. persistens*. Описано широкое разнообразие факторов вирулентности белковой природы или эффекторов, с помощью которых передача сигнала через мембранные рецепторы хозяина может контроли-

роваться патогеном вплоть до полного подавления защитного ответа, например, путем изменения структуры рецептируемого лиганда (Schulze-Lefert, Panstruga, 2011). При этом наблюдается реакция восприимчивости, обусловленная эффектором (ETS – effector-triggered susceptibility). Однако у большинства растений существует вторая линия защиты, эволюционно более поздняя ETI-система (effector-triggered immunity – запускаемый эффектором иммунитет). Секретируемые факторы вирулентности патогенов, эффекторы, в таких системах будут сигнальными молекулами, а звеном, передающим сигнал в ядро, – иммунные внутриклеточные рецепторы из семейства консервативных нуклеотид-связывающих лейцин-богатых белков (NLR – nucleotide-binding leucine-rich repeat receptor, или R-белки) (Dodds, Rathjen, 2010).

Важно отметить, что генетика факторов вирулентности до конца не изучена, однако моногенный контроль (гены авирулентности, или *Avr*) установлен для тех из них, которым соответствует R-белок растения (Flor, 1971; Hogenhout et al., 2009). Взаимодействие продуктов *Avr*-R генов запускает сложный каскад физиологических и биохимических реакций, направленных на предотвращение развития гриба, с наиболее распространенной реакцией сверхчувствительности (СВЧ) как формы апоптоза клеток хозяина, контактирующих с патогеном. К настоящему времени обнаружено значительное разнообразие механизмов устойчивого ответа. Кроме классической реакции сверхчувствительности за счет закисления цитоплазмы ионами Ca^{2+} совместно с накоплением активных форм кислорода (АФК), результатом рецепции *Avr*-эффектора может быть изменение окислительно-восстановительного (REDOX) потенциала на мембранах, запуск биосинтеза салициловой кислоты, ауксина и различных антимикробных PR-белков (pathogen-related proteins), среди которых хитиназы, β -глюканазы, ингибиторы грибных протеиназ и дополнительные протеиназы самого растения, пероксидазы, LTP-белки, транспортирующие липиды, и низкомолекулярные белки дефензины, богатые цистеином (Дьяков, 2017).

При исследовании многоуровневого фитоиммунитета обнаружено, что коэволюция привела к развитию у патогена способности преодолевать ETI-иммунитет различными способами. Наиболее распространенный способ – мутации генов *Avr*, когда рецепция продуктов генов вирулентности (*avr*) становится невозможной. Кроме того, могут появиться новые эффекторы, пока не имеющие соответствующего R-рецептора (рис. 1) (Jones, Dunl, 2006). Такое развитие событий возможно благодаря способности некоторых патогенов, включая ржавчинные грибы, уходить из зоны действия реакции СВЧ с помощью специализированных неветвящихся поисковых гиф (runner hypha), которые растут в сторону от подвергшихся апоптотическому коллапсу клеток растения (Moldenhauer et al., 2006). Таким образом, наблюдаемая реакция восприимчивости может обуславливаться эффектором патогена при отсутствии у растения соответствующих генов устойчивости (ETS), а может быть результатом мутации генов авирулентности (преодоление ETI). В последнем случае говорят о потере эффективности гена устойчивости, для которого продукт *Avr*-гена был сигнальной молекулой.

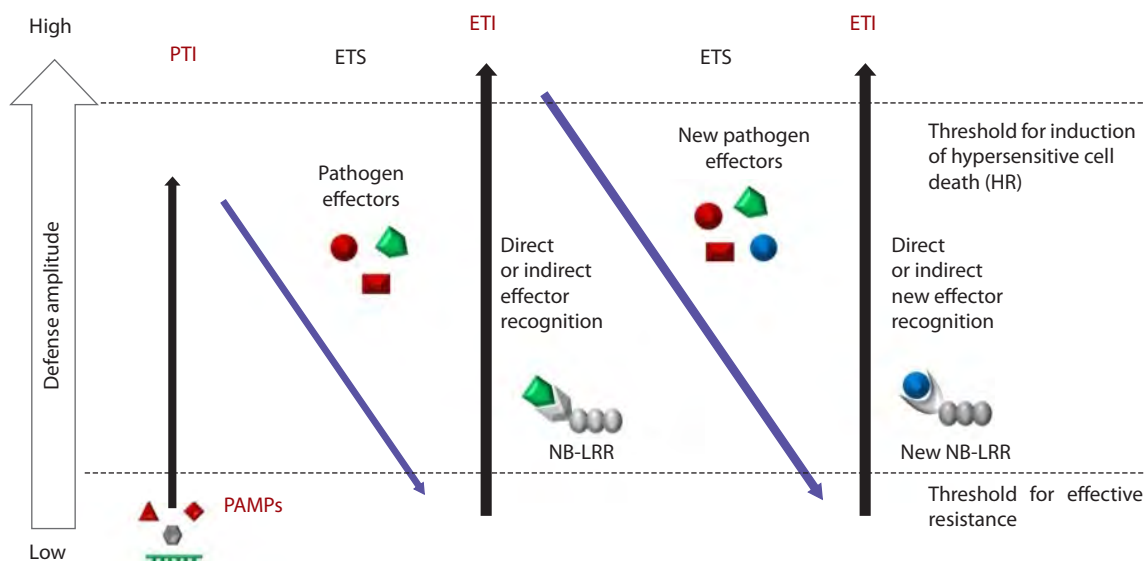


Fig. 1. The “zigzag” model representing the hierarchical plant immune system, according to (Jones, Dangl, 2006).

PAMP, pathogen-associated molecular pattern; PTI, PAMPs-triggered immunity; ETS, effector-triggered susceptibility; ETI, effector-triggered immunity; NB-LRR (NLR), nucleotide-binding leucine-rich repeat receptor.

Для описания многоуровневого фитоиммунитета была предложена модель «зигзаг», которая отражает зависимость экспрессии защитной реакции от степени специализации патогена (Jones, Dangl, 2006) (см. рис. 1). Она наглядно демонстрирует, что полигенный контроль устойчивого ответа на заражение является наивысшим уровнем взаимоотношений растения с патогеном, когда «новые» эффекторы патогена активируют «новые» R-рецепторы растения. Так, описаны линии пшеницы с промежуточным типом устойчивости к возбудителю желтой ржавчины, которые ограничивают развитие патогена с помощью запуска второй фазы реакции сверхчувствительности (Boskurt et al., 2010). Выявлены зоны первичного и вторичного некроза мезофила, появление которых, по данным анализа транскриптов ДНК, обусловлено работой различных генов, вовлеченных в каскад реакции сверхчувствительности. Однако важно отметить, что терминирующая фаза взаимодействия хозяина и патогена в реакциях несовместимости может обуславливаться альтернативными защитными механизмами, не запускающими клеточную смерть, а реализующимися, например, через накопление фенольных соединений в мезофиле, что показано на линиях овса с промежуточным типом устойчивости к *P. coronata* f. sp. *avenae* (Graichen et al., 2011).

В последнее время возрос интерес к механизмам защиты, альтернативным реакции гиперчувствительности, которые способны обеспечивать более длительную устойчивость хозяина к патогену, легко преодолевающему действие распецифических генов устойчивости благодаря высокому потенциалу популяционной изменчивости. Популяции ржавчинных грибов, в том числе возбудителей бурой ржавчины злаков, представляют собой набор физиологических рас, разных по набору генов авирулентности. По отношению к высокоспециализированным патогенам

со сложной внутривидовой структурой принято говорить о расоспецифической либо о неспецифической устойчивости. Первая является частным случаем устойчивости, реализуемым по схеме взаимодействия *Avr-R* генов. Неспецифическая или частичная устойчивость (partial resistance) имеет не качественную, а количественную (QR – quantitative resistance) генетическую основу и объединяет защитные механизмы растения, направленные не на остановку развития инфекции, а на снижение ее интенсивности. У такой устойчивости длительный эффект (durable resistance), так как она приводит к стабилизации эволюционных процессов в популяции патогена: в условиях лимитированного роста отбор направлен против агрессивных клонов, в популяции увеличивается соотношение генов авирулентности (*Avr*) к генам вирулентности (*avr*) и снижается вероятность возникновения редких мутаций вирулентности (Дьяков, 1998; Lagudah, 2011; Ellis et al., 2014; Niks et al., 2015; Krattinger, Keller, 2016). Количественная или частичная устойчивость к болезням достигается совместным действием нескольких минорных генов или количественных локусов (QTL – quantitative traits locus), поэтому их фенотипическое проявление не поддается качественной оценке, применимой при изучении R-генов, и описывается как slow rusting (Caldwell, 1968): несмотря на восприимчивый тип реакции, болезнь развивается медленно.

Мажорные (R-) гены устойчивости пшеницы к бурой ржавчине

Продукты R-генов растений находятся преимущественно в цитоплазме, куда эффекторы патогена попадают с помощью эндосом через гаустории. Большинство изученных у пшеницы R-генов и генных локусов устойчивости к ржавчинным болезням кодируют внутриклеточные ре-

цепторы семейства NLR (R-рецепторы), которые имеют общую доменную организацию. Единственным известным исключением является протеин-киназа, кодируемая локусом *Rpg1*, описанным у ячменя и определяющим устойчивость к стеблевой ржавчине при интрогрессии в геном пшеницы (Brueggeman et al., 2002).

Разнообразие генов специфической устойчивости обеспечивает структурное различие доменов, образующих R-рецепторы. LRR-область (leucine rich repeat – обогащенная лейцином область), являясь центром узнавания эффектора патогена, обладает наибольшим полиморфизмом, что выражается в варьирующем количестве обогащенных лейцином повторов, от 14 до 37, и наборе аминокислот, дополняющих лейцин в консенсусной последовательности (Dangle, Jones, 2001). NBS-домен (nucleotide binding site) представлен сайтами связывания нуклеотидов (Р-петля или киназа 1а, киназа 2, киназа 3а) и обеспечивает специфическую трансдукцию сигнала по киназному пути с гидролизом АТФ. Здесь же недавно обнаружен мотив NB-ARC, сходный с фактором активации апоптоза животных Araf-1 (apoptotic protease-activating factor 1), что указывает на участие NBS-домена в индукции реакции СВЧ как формы апоптоза у растений. В случае, когда R-рецептор связан с транскрипционным фактором WRKY ядра, возможна регуляция экспрессии генов без промежуточных киназ (Wu et al., 2008). Например, у пшеницы описана положительная регуляция промоторов *PR4* генов, кодирующих фунгицидные белки, при передаче сигнала от R-рецептора на транскрипционный фактор TaWRKY78 (Proietti et al., 2010). Кроме областей связывания эффектора (LRR-) и трансдукции сигнала (NBS-), молекула R-рецептора однодольных растений может содержать несколько копий суперскрученного (суперспирализованного) домена (CC-, coiled-coil domain) или участок, гомологичный толл-подобному рецептору цитокинина ИЛ-1 млекопитающих (TIR-, toll-interleukin 1 receptor), функции которых связаны с индукцией СВЧ и взаимодействием с белками сигнального пути.

В каталоге пшеничных генов устойчивости к ржавчинным заболеваниям приводится более 100 генов и локусов устойчивости к бурой ржавчине (McIntosh et al., 2014), из которых 76 имеют формальные названия с нумерацией генов *Lr* (leaf rust – бурая ржавчина). Однако структура и функции кодируемых R-продуктов хорошо изучены только у некоторых клонированных генов: *Lr21* (Huang et al., 2003), *Lr10* (Feuillet et al., 2003), *Lr1* (Cloutier et al., 2007), а также двух кандидатов гена *Lr19*: *CIN14* (Zhang et al., 2011) и *Lr19-Ag15* (Gennaro et al., 2009). При филогенетическом анализе аминокислотных последовательностей обнаруживается значительное сходство между белками, кодируемыми генами *Lr1*, *CIN14* и *Lr19-Ag15*, несмотря на то что ген *Lr19* происходит от пырея *Thinopyrum elongatum* (Zhang et al., 2011). Все описанные к настоящему моменту рецепторы пшеницы, связывающие эффекторы возбудителя бурой ржавчины, являются CC-NBS-LRR белками. Однако у гена *Lr10* клонирован фрагмент *Lrk10* размером 770 п. н., кодирующий четвертый редкий класс рецептороподобных киназ. В своем составе *Lrk10* имеет специфический цитоплазматический домен и небольшую трансмембранную последовательность без CC-домена в

С-терминальном участке (Feuillet et al., 1997). Наибольшим отличием обладает белковый продукт гена *Lr21*, молчащие аллели которого представлены в пшеничном геноме мягкой пшеницы. Для функционального аллеля локуса *Lr21*, интрогрессированного в геном синтетической пшеницы от дальнего сородича *Aegilops tauschii*, установлено химерное происхождение в результате межвидовой рекомбинации при участии пшеничных гаплотипов с молчащими локусами *Lr21* типа Fielder (H1) и Wichita (H2) (Huang et al., 2009). Подобная ситуация описывается как death–recycle модель эволюции генома (Michelmore, Meyers, 1998), т. е. появление новых генов устойчивости путем возвращения активности рецессивным аллелям в геноме пшеничных сородичей, и указывает на ценность исходного материала, представленного в центрах происхождения культурных растений, описанных Вавиловым.

Минорные гены и количественные локусы (QTL-) устойчивости пшеницы к бурой ржавчине

Предполагается, что большинство известных минорных генов и количественных локусов кодируют белки, не являющиеся иммунными рецепторами, но выполняющие иные защитные функции (Krattinger et al., 2009; Gou et al., 2015; Moore et al., 2015). Степень проявления комплекса защитных мер, контролируемых минорными генами, часто зависит от стадии развития растения, что сформировало концепцию возрастной устойчивости (APR – adult plant resistance), имеющей значение для селекции на иммунитет к возбудителям ржавчинных заболеваний злаков. Возрастная устойчивость характеризуется восприимчивыми инфекционными типами реакций на заражение у проростков и резко возрастающей эффективностью на следующих за ювенильной стадиях развития растения (Park, McIntosh, 1994).

Различают две категории APR-генов: запускающие реакцию СВЧ (hypersensitive genes) и реализующие альтернативные механизмы устойчивости (nonhypersensitive genes). Фенотипическое проявление действия генов возрастной устойчивости к бурой ржавчине *Lr12*, *Lr13*, *Lr22a*, *Lr22b*, *Lr35*, *Lr48* и *Lr49*, относящихся к первой категории, описывается как некротические пятна на листьях или крошечные пустулы, заметные уже у проростков пшеницы. При заражении растений с APR-генами, не запускающими СВЧ, на листьях появляются инфекционные типы без некрозов, от восприимчивых до средневосприимчивых, при этом характерны сниженная частота инфекционных очагов и отсроченные стадии развития патогена. По отношению к бурой ржавчине пока идентифицировано немного таких генов или количественных локусов устойчивости: *Lr34*, *Lr46*, *Lr67*, *Lr77* (Dyck, 1987; Singh et al., 1998; Herrera-Foessel et al., 2014; Kolmer et al., 2018), не кодирующих рецепторы семейства NLR (R-рецепторы). Механизмы их действия изучают с привлечением QTL-анализа, и уже известны некоторые продукты QR-генов. Ген или генный локус *Lr34* кодирует АТФ-зависимый белковый транспортер, отвечающий за частичную устойчивость к нескольким видам ржавчины (Krattinger et al., 2009). Аллель *Lr67*, обеспечивающий частичную устойчивость к бурой ржавчине (*Lr67 resistant*, *Lr67res*), отвечает за прекращение транспорта гексоза из клетки хозяина в апопласт,

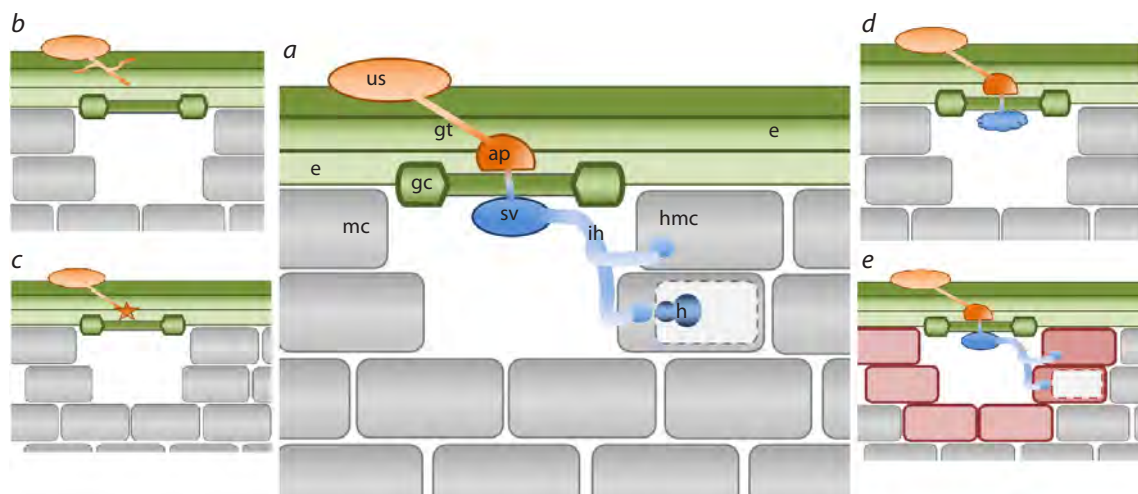


Fig. 2. Mechanisms of prehaustorial resistance to rusts in wheat as compiled from (Hu, Rijkenberg, 1998; Leonard, Szabo, 2005; Plotnikova, 2009; Wang et al., 2013; Pozherukova et al., 2015).

(a) Ectophytic (orange) and early endophytic (blue) stages of pathogenesis. Abbreviations: us, urediniospore; gt, germ tube; ap, appressorium; sv, substomatal vesicle; ih, infection hypha; hmc, haustorial mother cell; h, haustorium; e, epidermal cells; gc, guard cells; mc, mesophyll cells. Arrest of the development of infection at different stages of pathogenesis: (b) germ tube disorientation and branching; (c) aberrant appressorium on stoma; (d) distorted or collapsed substomatal vesicle; (e) accumulation of callose in mesophyll cell walls.

тогда как аллель *LR67sus* (susceptible – восприимчивый) является Н⁺/гексозным транспортером и обладает высоким средством к глюкозе (Moore et al., 2015).

Прегаусториальная устойчивость пшеницы к возбудителю бурой ржавчины

Большинство идентифицированных к настоящему моменту генов *Lr* являются расоспецифическими *R*-генами и кодируют внутриклеточные NLR-рецепторы. Они обеспечивают гиперчувствительную реакцию растения (СВЧ и альтернативные механизмы устойчивости) в ответ на поражение расами патогена, несущими соответствующие гены авирулентности. Подчеркнем, что взаимодействие продуктов генов *Avr* и *R* происходит уже в цитоплазме клетки, после формирования гаустория. Поэтому специфическую устойчивость с гиперчувствительной реакцией в иностранной литературе часто называют гаусториальной или развивающейся после проникновения устойчивостью (post-penetration resistance) (Heath, 2002; Niks, Marcel, 2009). Однако в последнее время накоплены данные об участии некоторых *R*-генов пшеницы в реализации механизмов прегаусториальной устойчивости, останавливающих развитие ржавчинных грибов на ранней эндифитной и даже эктофитной стадии. Функциональными мишенями механизмов прегаусториальной устойчивости являются ростковая трубка и аппрессорий (эктофитная фаза патогенеза), а также подустычная везикула и инфекционная гифа (ранняя эндифитная фаза патогенеза ржавчинных грибов) (рис. 2).

Описан окислительный взрыв при контакте возбудителя бурой ржавчины с устьищами на линиях с расоспецифическим геном *Lr38* (от *A. intermedium*), сходный по времени и характеру с реакцией видов-нехозяев. При заражении набором рас *P. triticina* на устьищах изогенной линии пшеницы с геном *Lr38* и на растении-нехозяине *T. timopheevii*

через 24 ч после инокуляции были отмечены некротизированные маленькие аппрессории гриба размером до 20 мкм, тогда как на устьищах контрольной линии Тэтчер формировались нормальные крупные аппрессории размером около 50 мкм (Плотникова, 2009; Пожерукова и др., 2015). У линий с генами *Lr19* и *Lr21* со специфической устойчивостью к возбудителю бурой ржавчины описаны дополнительные к реакции сверхчувствительности механизмы блокирования развития патогена. На первых стадиях развития эндифитной фазы патогена часто наблюдалась деформация или коллапсирование подустычной везикулы, что снижало эффективность колонизации в ткани растения (Hu, Rijkenberg, 1998). Были отмечены аккумуляция каллозы и отложение лигнина при заражении изогенных линий пшеницы, несущих гены специфической устойчивости *Lr1*, *Lr3a*, *Lr9*, *LrB*, набором вирулентных рас *P. triticina* (Wang et al., 2013).

Таким образом, современные подходы к исследованию механизмов устойчивости известных и новых *R*-генов должны отражать обозначенную вероятность активации цитоплазматических *R*-рецепторов на первых стадиях патогенеза, когда высвобождаются только лиганды РАРМ. Трансдукция сигнала, вероятно, может быть опосредована возбуждением механо-чувствительных ионных каналов (MSC – mechanosensitive-ion channel), описанных в клеточной стенке эпидермиса растений как небольшие трансмембранные белковые рецепторы, активация которых приводит к изменению окислительно-восстановительного потенциала клетки (Hamilton et al., 2015).

Заключение

Накопление данных о разнообразных механизмах устойчивости пшеницы к ржавчинным заболеваниям создает возможность оптимально использовать генетические ресурсы при селекции на иммунитет. Фитоиммунитет пред-

ставляет собой многоуровневую систему защиты растения-хозяина, контролируемую комбинациями мажорных и минорных генов (локусов) устойчивости, так называемыми полигенными семействами (multigene families) (Marone et al., 2013), способными влиять на различные стадии патогенеза. Интересно, что количественные локусы устойчивости к грибным заболеваниям в геноме растений часто локализованы рядом с *R*-генами, располагаясь во фланкирующей области в пределах 2 Mb, поэтому с большой вероятностью косегрегируют (Shang et al., 2009; Kang et al., 2012). Изучение кластера мажорных *R*-генов показало, что они являются наиболее динамичной группой генов, способных к быстрой эволюции (Krattinger, Keller, 2016). При этом в природных ценозах полиморфизм *R*-белков создается между индивидуальными организмами на популяционном уровне, что отличает фитоиммунитет от животной системы защиты с огромным полиморфизмом рецепторов внутри индивидуальных организмов. Углубленное изучение многоуровневого фитоиммунитета привело к переосмыслению агрономической стратегии: селекцию чистых линий с единичными мажорными *R*-генами, долгое время бывшую традиционной в растениеводстве, заменили программы создания мультилинейных сортов-популяций, различающихся генами или аллелями генов устойчивости (Дьяков, 2017).

Другим важным подходом, интенсивно используемым в последнее время при создании новых сортов и селекционных линий, является пирамидирование генов устойчивости в одном генотипе путем близкородственной и отдаленной гибридизации. Созданы сорта мягкой пшеницы, устойчивые к ржавчинным заболеваниям за счет пирамиды из нескольких *R*-генов, а также комбинации мажорных и минорных генов устойчивости, что позволило добиться эффекта долгосрочной защиты от патогенов. Наибольшее значение в селекции пшеницы на иммунитет к ржавчинным заболеваниям имеют пирамиды генов, содержащие локус устойчивости *Lr34* (*Lr34 complex*) (Singh, Rajaram, 1991) и минорный ген устойчивости *Sr2* (*Sr2 complex*) (Singh et al., 2000).

Задачей, стоящей перед современной селекцией, является разработка эффективных методов трансгенного переноса полигенных семейств растений, которому должны предшествовать точный анализ выявленных генных групп устойчивости на молекулярном уровне. Это позволит не только сократить влияние нежелательного генетического материала, приобретаемого в процессе обычной гибридизации, но и ввести в реципиентный сорт целевые гены из неродственных организмов, в том числе полигенные семейства мажорных и минорных генов (локусов) устойчивости. Так, уже созданы трансгенные линии пшеницы, устойчивость которых к стеблевой ржавчине обеспечивается геном *Rpg1* из ячменя (Collinge et al., 2010). Предложены бактериальные системы переноса генов, совместимые с пшеницей (*Pseudomonas*, *Xanthomonas* spp.), для функционального анализа эффекторов ржавчинных грибов (Yin, Hulbert, 2011; Upadhyaya et al., 2014). Перспективные работы с самим патогеном по созданию трансгенных штаммов с *Avr*-генами для исследования защитных механизмов растения, что уже сделано, например, для ржавчины льна *Melampsora lini* (Lawrence et al., 2010).

Список литературы / References

- Вавилов Н.И. Учение об иммунитете растений к инфекционным заболеваниям (Применительно к запросам селекции). Теоретические основы селекции растений. Т. 1. М.; Л., 1935;893-900.
[Vavilov N.I. The theory of plant immunity to infectious diseases. In: Theoretical Bases of Plant Breeding. Vol. 1. Moscow; Leningrad, 1935;893-900. (in Russian)]
- Вандерпланк Я.Е. Устойчивость растений к болезням. Пер. с англ. М.: Колос, 1972.
[Van der Plank J.E. Disease Resistance in Plants. New York: Acad. Press, 1968. (Russ. ed.: Vanderplank J.E. Ustoychivost' Rasteniy k Boleznyam. Moscow: Kolos Publ., 1972. (in Russian))]
- Горленко М.В. Сельскохозяйственная фитопатология. М.: Высш. шк., 1968.
[Gorlenko M.B. Agricultural Plant Pathology. Moscow: Vysshaya Shkola Publ., 1968. (in Russian)]
- Дьяков Ю.Т. Популяционная биология фитопатогенных грибов. М.: Изд. дом «Муравей», 1998.
[Diakov Yu.T. Population Biology of Phytopathogenic Fungi. Moscow: Muravey Publ., 1998. (in Russian)]
- Дьяков Ю.Т. На пути к общей теории иммунитета. Журн. общ. биологии. 2005;66(6):451-458.
[Diakov Yu.T. Towards the general theory of immunity. Zhurnal Obshchey Biologii = Journal of General Biology. 2005;66(6):451-458. (in Russian)]
- Дьяков Ю.Т. Фитоиммунитет. М.: ИНФА-М, 2017.
[Diakov Yu.T. Plant Disease Resistance. Moscow: INFА-M Publ., 2017. (in Russian)]
- Михайлова Л.А. Устойчивость пшеницы к бурой ржавчине. Идентифицированный генофонд растений и селекция. СПб.: ВИР, 2005;513-526.
[Mikhaylova L.A. Genetics of wheat resistance to brown rust. In: Identified Plant Gene Pool and Breeding. St. Petersburg: VIR Publ., 2005;513-526. (in Russian)]
- Одинцова И.Г., Шеломова Л.Ф. Горизонтальная устойчивость: генетика и возможности преодоления паразитом. Изменчивость фитопатогенных микроорганизмов. М.: Колос, 1983;51-60.
[Odintsova I.G., Shelomova L.F. Horizontal resistance: genetics and breachability by parasites. In: Variability of Pathogenic Microorganisms. Moscow: Kolos Publ., 1983;51-60. (in Russian)]
- Плотникова Л.Я. Участие активных форм кислорода в защите линий пшеницы с генами устойчивости рода *Agropyron* от бурой ржавчины. Физиология растений. 2009;56(2):200-209.
[Plotnikova L.Ya. The involvement of reactive oxygen species in defense of wheat lines with the genes introgressed from *Agropyron* species contributing the resistance against brown rust. Russ. J. Plant Physiol. 2009;56(2):181-189. DOI 10.1134/S102144370902006X.]
- Пожерукова В.Е., Плотникова Л.Я., Дегтярев А.И. Устойчивость пшеницы Тимофеева к бурой ржавчине определяется генерацией активных форм кислорода и подавлением образования гаусторий гриба *Puccinia tritica*. Фундам. исследования. 2015;2(2):285-292.
[Pozherukova V.E., Plotnikova L.Ya., Degtyarev A.I. Timofeevi wheat resistance to leaf rust is determined by generation of reactive oxygen species and inhibition of *Puccinia tritica* fungus haustoria development. Fundamentalnye Issledovaniya = Fundamental Research. 2015;2(2):285-292. (in Russian)]
- Atienza S.G., Jafary H., Niks R.E. Accumulation of genes for susceptibility to rust fungi for which barley is nearly a nonhost results in two barley lines with extreme multiple susceptibility. Planta. 2004;220:71-79.
- Ayliffe M., Jin Y., Kang Z., Persson M., Steffenson B., Wang S., Leung H. Determining the basis of nonhost resistance in rice to cereal rusts. Euphytica. 2011;179:33-40. DOI 10.1007/s10681-010-0280-2.
- Ayliffe M., Singh R., Lagudah E. Durable resistance to wheat stem rust needed. Curr. Opin. Plant Biol. 2008;11:187-192.

- Bozkurt T.O., McGrann G.R.D., MacCormack R., Boyd L.A., Akkaya M.S. Cellular and transcriptional responses of wheat during compatible and incompatible race-specific interactions with *Puccinia striiformis* f. sp. *tritici*. *Mol. Plant Pathol.* 2010;11:625-640. DOI 10.1111/j.1364-3703.2010.00633.x.
- Brueggeman R., Rostoks N., Kudrna D., Kilian A., Han F., Chen J., Druka A., Steffenson B., Kleinhofs A. The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. USA.* 2002;9(14):9328-9333. DOI 10.1073/pnas.142284999.
- Caldwell R.M. Breeding for general and specific plant disease resistance. *Proc. 3rd Int. Wheat Genetic Symp. Australia.* 1968;263-272.
- Cloutier S., McCallum B.D., Loutre C., Banks T.W., Wicker T., Feuillet C., Keller B., Jordan M.C. Leaf rust resistance gene *Lr1*, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large *psr567* gene family. *Plant Mol. Biol.* 2007;65:93-106.
- Collinge D.B., Jørgensen H.J.L., Lund O.S., Lyngkjær M.F. Engineering pathogen resistance in crop plants: current trends and future prospects. *Annu. Rev. Phytopathol.* 2010;48:269-291.
- Cook D.E., Mesarich C.H., Thomma B.P.H.J. Understanding plant immunity as a surveillance system to detect invasion. *Annu. Rev. Phytopathol.* 2015;53:541-563.
- Dangle J.L., Jones J.D.G. Plant pathogens and integrated defense responses to infection. *Nature.* 2001;411:826-833.
- Dodds P.N., Rathjen J.P. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat. Rev. Genet.* 2010;11:539-548.
- Dyck P.L. The association of a gene for leaf rust resistance with the chromosome 7D suppressor of stem rust resistance in common wheat. *Genome.* 1987;29:467-469.
- Ellis J.G., Lagudah E.S., Spielmeier W., Dodds P.N. The past, present and future of breeding rust resistant wheat. *Front. Plant Sci.* 2014;5: 641. DOI 10.3389/fpls.2014.00641.
- Feuillet C., Schachermayr G., Keller B. Molecular cloning of a new receptor-like kinase gene encoded at the *Lr10* disease resistance locus of wheat. *Plant J.* 1997;11(1):45-52.
- Feuillet C., Travella S., Stein N., Albar L., Nublat A., Keller B. Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc. Natl. Acad. Sci. USA.* 2003;100:15253-15258.
- Flor H. Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* 1971;9:275-296.
- Gennaro A., Koebner R.M.D., Ceoloni C. A candidate for *Lr19*, an exotic gene conditioning leaf rust resistance in wheat. *Funct. Integr. Genomics.* 2009;9:325-334. DOI 10.1007/s10142-009-0115-1.
- Gou J.Y., Li K., Wu K., Wang X., Lin H., Cantu D., Uauy C., Dobon-Alonso A., Midorikawa T., Inoue K., Sánchez J., Fu D., Blechl A., Wallington E., Fahima T., Meeta M., Epstein L., Dubcovsky J. Wheat stripe rust resistance protein WKS1 reduces the ability of the thylakoid-associated ascorbate peroxidase to detoxify reactive oxygen species. *Plant Cell.* 2015;27:1755-1770.
- Graichen F.A.S., Martinelli J.A., Federizzi L.C., Chaves M.S. Epidemiological and histological components of crown rust resistance in oat genotypes. *Eur. J. Plant Pathol.* 2011;131:497-510.
- Hamilton E.S., Jensen G.S., Maksaev G., Katims A., Sherp A.M., Haswell E.S. Mechanosensitive channel MSL8 regulates osmotic forces during pollen hydration and germination. *Science.* 2015;350(6259): 438-441. DOI 10.1126/science.aac6014.
- Heath M.C. Cellular interactions between biotrophic fungal pathogens and host or nonhost plants. *Can. J. Plant Pathol.* 2002;24:259-264. DOI 10.1080/0706060209507007.
- Herrera-Foessel S.A., Singh R.P., Lillemo M., Huerta-Espino J., Bhavani S., Singh S., Lan C., Calvo-Salazar V., Lagudah E.S. *Lr67/Yr46* confers adult plant resistance to stem rust and powdery mildew in wheat. *Theor. Appl. Genet.* 2014;127:781-789.
- Hogenhout S.A., Van der Hoorn R.A., Terauchi R., Kamoun S. Emerging concepts in effector biology of plant-associated organisms. *Mol. Plant Microbe Interact.* 2009;22(2):115-122. DOI 10.1094/MPMI-22-2-0115.
- Hu G.G., Rijkenberg F.H.J. Ultrastructural studies of the intercellular hyphae and haustorium of *Puccinia recondita* f. sp. *tritici*. *J. Phytopathol.* 1998;146:39-50.
- Huang L., Brooks S.A., Li W., Fellers J.P., Trick H.N., Gill B.S. Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. *Genetics.* 2003;164:655-664.
- Huang L., Brooks S.A., Li W., Fellers J.P., Nelson J.C., Gill B.S. Evolution of new disease specificity at a simple resistance locus in a crop-weed complex: reconstitution of the *Lr21* gene in wheat. *Genetics.* 2009;182:595-602.
- Jones J.D., Dangl J.L. The plant immune system. *Nature.* 2006;444: 323-329. DOI 10.1038/nature05286.
- Jonson R. Durable resistance: definition of, genetic control, and attainment in plant breeding. *Phytopathology.* 1981;71:567-568.
- Kang Y.J., Kim K.H., Shim S., Yoon M.Y., Sun S., Kim M.Y., Van K., Lee S.H. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. *BMC Plant Biol.* 2012; 12:139.
- Kolmer J.A., Su Z., Bernardo A., Bai G., Chao S. Mapping and characterization of the new adult plant leaf rust resistance gene *Lr77* derived from Santa Fe winter wheat. *Theor. Appl. Genet.* 2018;131:1553-1560. DOI 10.1007/s00122-018-3097-3.
- Krattinger S.G., Keller B. Molecular genetics and evolution of disease resistance in cereals. *New Phytol.* 2016;212:320-332.
- Krattinger S.G., Lagudah E.S., Spielmeier W., Singh R.P., Huerta-Espino J., McFadden H., Bossolini E., Selter L.L., Keller B. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science.* 2009;323:1360-1363.
- Lagudah E.S. Molecular genetics of race non-specific rust resistance in wheat. *Euphytica.* 2011;179:81-91.
- Lawrence G.J., Dodds P.N., Ellis J.G. Transformation of the flax rust fungus, *Melampsora lini*: selection via silencing of an avirulence gene. *Plant J.* 2010;61:364-369. DOI 10.1111/j.1365-3113X.2009.04052.x.
- Leonard K.J., Szabo L.J. Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol. Plant Pathol.* 2005;6:99-111.
- Marone D., Russo M.A., Laidò G., Leonardi A.M., Mastrangelo A.M. Plant nucleotide binding site-leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. *Int. J. Mol. Sci.* 2013; 14(4):7302-7326.
- McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers W.J., Morris C., Appels R., Devos K.M. Catalogue of gene symbols for wheat. KOMUGI Integrated Wheat Science Database. 2014. Available at <http://www.shigen.nig.ac.jp/wheat/komugi/ge-nes/symbolClassList.jsp>
- Michelmores R.W., Meyers B.C. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 1998;8:1113-1130.
- Moldenhauer J., Moerschbacher B.M., van der Westhuizen A.J. Histological investigation of stripe rust (*Puccinia striiformis* f. sp. *tritici*) development in resistant and susceptible wheat cultivars. *Plant Pathol.* 2006;55(4):469-474. DOI 10.1111/j.1365-3059.2006.01385.x.
- Moore J.W., Herrera-Foessel S., Lan C., Schnippenkoetter W., Ayliffe M., Huerta-Espino J., Lillemo M., Vicens L., Milne R., Periyannan S., Kong X., Spielmeier W., Talbot M., Bariana H., Patrick J.W., Dodds P., Singh R., Lagudah E. A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 2015;47:1494-1498.
- Niks R.E. Haustorium formation by *Puccinia hordei* in leaves of hypersensitive, partially resistant, and nonhost plant genotypes. *Phytopathology.* 1983;73:64-66.
- Niks R.E., Marcel T.C. Nonhost and basal resistance: how to explain specificity? *New Phytol.* 2009;182:817-828.
- Niks R.E., Qi X., Marcel T.C. Quantitative resistance to biotrophic filamentous plant pathogens: concepts, misconceptions, and mechanisms. *Annu. Rev. Phytopathol.* 2015;53:445-470.
- Nürnberg T., Kemmerling B. Pathogen-associated molecular patterns (PAMP) and PAMP-triggered immunity. *Annu. Plant Rev.* 2009;34: 16-47.

- Park R.F., McIntosh R.A. Adult plant resistances to *Puccinia recondita* f. sp. *tritici* in wheat. N. Z. J. Crop Hortic. Sci. 1994;22:151-158.
- Parlevliet J.E., Zadoks J.C. The integrate concept disease resistance: a new view including horizontal and vertical resistance in plants. Euphytica. 1977;26:5-21.
- Proietti S., Bertini L., Van der Ent S., Leon-Reyes A., Pieterse C.M.J., Tucci M., Caporale C., Caruso C. Cross activity of orthologous WRKY transcription factors in wheat and *Arabidopsis*. J. Exp. Bot. 2011;62(6):1975-1990. DOI 10.1093/jxb/erq396.
- Schulze-Lefert P., Panstruga R. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci. 2011;16:117-125.
- Shang J., Tao Y., Chen X., Zou Y., Lei C., Wang J., Li X., Zhao X., Zhang M., Lu Z., Xu J., Cheng Z., Wan J., Zhu L. Identification of a new rice blast resistance gene, *Pid3*, by genome wide comparison of paired nucleotide-binding site leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes. Genetics. 2009;182:1303-1311.
- Singh R.P., Mujeeb-Kazi A., Huerta-Espino J. *Lr46*: a gene conferring slow-rusting resistance to leaf rust in wheat. Phytopathology. 1998; 88:890-894.
- Singh R.P., Nelson J.C., Sorrells M.E. Mapping *Yr28* and other genes for resistance to stripe rust in wheat. Crop Sci. 2000;40:1148-1155.
- Singh R.P., Rajaram S. Resistance to *Puccinia recondita* f. sp. *tritici* in 50 Mexican bread wheat cultivars. Crop Sci. 1991;31:1472-1479.
- Thomma B.P.H.J., Nurnberger T., Joosten M.H.A.J. Of PAMPs and effectors: the blurred PTI-ETI dichotomy. Plant Cell. 2011;23: 4-15.
- Upadhyaya N.M., Mago R., Staskawicz B.J., Ayliffe M.A., Ellis J.G., Dodds P.N. A bacterial type III secretion assay for delivery of fungal effector proteins in to wheat. Mol. Plant Microbe Interact. 2014; 27:255-264.
- van Ooijen G., Mayr G., Kasiem M.M., Albrecht M., Cornelissen B.J., Takken F.L. Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. J. Exp. Bot. 2008;59(6):1383-1397.
- Wang X., McCallum B.D., Fetch T., Bakkeren G., Marais G.F., Saviile B.J. Comparative microscopic and molecular analysis of Thatcher near-isogenic lines with wheat rust resistance genes *Lr2a*, *Lr3*, *LrB* or *Lr9* upon challenge with different *Puccinia triticina* races. Plant Pathol. 2013;62:698-707.
- Wu H., Ni Z., Yao Y., Guo G., Sun Q. Cloning and expression profiles of 15 genes encoding WRKY transcription factors in wheat (*Triticum aestivum* L.). Prog. Nat. Sci. 2008;18:697-705.
- Yin C., Hulbert S. Prospects for functional analysis of effectors from cereal rust fungi. Euphytica. 2011;179:57-67.
- Zhang N., Wang S., Wang H., Liu D. Isolation and characterization of NBS-LRR class resistance homologous gene from wheat. Agric. Sci. China. 2011;10(8):1151-1158. DOI 10.1016/S1671-2927(11) 60105-3.

ORCID ID

E.S. Skolotneva orcid.org/0000-0001-8047-5695
E.A. Salina orcid.org/0000-0001-8590-847X

Acknowledgements. This work was supported by the Russian Foundation for Basic Research, project 17-04-00507, and State Budgeted Project 0259-2019-0001.

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

Received May 13, 2019. Revised June 10, 2019. Accepted June 10, 2019.

Характеристика генофонда яровой мягкой пшеницы (*Triticum aestivum* L.) по устойчивости к пыльной головне в условиях лесостепи Западной Сибири

Е.А. Орлова , Н.П. Бехтольд

Сибирский научно-исследовательский институт растениеводства и селекции – филиал Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
 e-mail: Orlova.Lena10@yandex.ru

Среди многочисленных болезней яровой пшеницы, вызываемых патогенными грибами, пыльная головня *Ustilago tritici* (Pers.) Jens. по-прежнему остается опасным заболеванием. На полях, где отсутствует контроль над проявлением и распространением болезни, недобор урожая может составлять до 10 %, а при возделывании высоковосприимчивых сортов достигать 40–50 %. С учетом возрастающей стоимости протравителей семян и их экологической небезопасности, самым доступным способом защиты растений, снижающим пестицидную нагрузку на агроценозы, является возделывание сортов, невосприимчивых к пыльной головне. Решающий момент в селекции на устойчивость – использование в качестве родительских форм сортов, резистентных к патогену. Целью наших исследований было выделение невосприимчивых к пыльной головне образцов яровой пшеницы на фоне искусственного заражения растений популяцией, специфичной для Западно-Сибирского региона. Приведены результаты многолетних исследований 350 генотипов яровой пшеницы различного эколого-географического происхождения на устойчивость к болезни. Физиологическую специализацию рас *U. tritici* осуществляли на дифференцирующем наборе, состоящем из шести сортов мягкой и трех сортов твердой яровой пшеницы. Полученные результаты в сочетании с литературными данными свидетельствуют об изменениях в расовом составе, произошедших в культуре патогена за последние 30–35 лет. Выделены сорта иностранной и отечественной селекции, резистентные к западносибирской популяции пыльной головни. На основе анализа родословных образцов, высоко- и практически устойчивых к головне, сделано предположение, что в селекции на иммунитет к *U. tritici* чаще всего используют одни и те же источники генов резистентности. Среди генофонда яровой пшеницы иностранной селекции наибольшее количество генотипов, резистентных к пыльной головне, выделяется в странах Североамериканской географической зоны (США, Канада, Мексика). В основном это образцы, несущие гены *Ut1*, гены от сорта пшеницы Thatcher и ее сестринской линии DC II-21-44. В сортах российской селекции прослеживаются гены от сортов Белотурка, Полтавка, Селивановский Русак (через Саратовскую 29 и ее производные), гены от пшенично-пырейных линий АГИС 1 и Грекум 114.

Ключевые слова: яровая мягкая пшеница; пыльная головня; источники устойчивости; расы; популяция.

Для цитирования: Орлова Е.А., Бехтольд Н.П. Характеристика генофонда яровой мягкой пшеницы (*Triticum aestivum* L.) по устойчивости к пыльной головне в условиях лесостепи Западной Сибири. Вавиловский журнал генетики и селекции. 2019;23(5):551-558. DOI 10.18699/VJ19.524

Characteristics of the gene pool of spring wheat (*Triticum aestivum* L.) for resistance to loose smut in the forest-steppe of Western Siberia

E.A. Orlova , N.P. Baechtold

Siberian Research Institute of Plant Production and Breeding – Branch of the Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
 e-mail: Orlova.Lena10@yandex.ru

Among the many diseases of spring wheat caused by pathogenic fungi, loose smut *Ustilago tritici* (Pers.) Jens. remains to be a dangerous disease with a wide range of distribution. In fields where there is no control over the emergence and spread of the disease, the yield reduction can be up to 10 %, and in the case of highly susceptible varieties, up to 40–50 %. Taking into account the increasing cost of seed protectants and their environmental damage, the cultivation of varieties resistant to loose smut is still the most affordable way to protect plants, reducing the pesticide load on agroecosystems. The crucial point in breeding for resistance is the use of resistant varieties as parental forms. The aim of our research was to isolate samples of spring wheat that are immune to loose smut against the background of artificial infection of plants with a population specific to the West Siberian region. The article presents the results of long-lasting studies of 350 genotypes of spring wheat of different ecological and geographical origin for resistance to disease. Physiological specialization of races was carried out on the basis of a differentiating set consisting of six varieties of soft wheat and three varieties of durum spring wheat. The obtained results in combination with literature

data reveal changes in the racial composition of the pathogen population over the past 30–35 years. Varieties of foreign and domestic selection resistant to the West Siberian population of loose smut have been identified. Based on the analysis of pedigree samples, highly and practically resistant to loose smut, we concluded that in breeding for immunity to *U. tritici*, the same sources of resistance genes are most often used. Among the gene pool of spring wheat of foreign selection, the largest number of genotypes resistant to loose smut is assigned to the countries of the North American geographical zone (USA, Canada, Mexico). These are largely samples containing *Ut1* genes, genes from spring wheat 'Thatcher' and its sister line 'DC II-21-44'. Resistance genes in Russian wheat varieties can be traced from cultivars Beloturka, Poltavka, Selivanovsky Hare (using Saratovskaya 29 and its derivatives), and genes from wheat-grass lines AGIS 1 and Grecum 114.

Key words: spring soft wheat; loose smut; sources of stability; races; population.

For citation: Orlova E.A., Baechtold N.P. Characteristics of the gene pool of spring wheat (*Triticum aestivum* L.) for resistance to loose smut in the forest-steppe of Western Siberia. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5):551–558. DOI 10.18699/VJ19.524 (in Russian)

Введение

Среди многочисленных болезней яровой пшеницы, вызываемых патогенными грибами, опасным заболеванием, имеющим широкий ареал распространения, является пыльная головня. Несмотря на то что возбудитель не вызывает эпифитотий, потери урожая зерновых от головневых болезней могут широко варьировать. На полях, где отсутствует контроль над проявлением и распространением болезни, недобор урожая может достигать 10 %, а при возделывании высоковосприимчивых сортов – 40–50 % (Nielsen, Thomas, 1996). Кроме явных потерь, заключающихся в полном разрушении колоса, возбудители головневых вызывают и скрытые потери урожая. Угнетающее действие патогена проявляется на всех этапах развития растений: уменьшается количество взошедших растений, продуктивная кустистость, масса 1000 зерен (Степановских, 1990; Nielsen, Thomas, 1996; Дружин, 2001; Орлова и др., 2015).

Существует прямая корреляционная зависимость между эмбриональной инфекцией и проявлением пыльной головни в полевых условиях. Фунгицидное протравливание семян не снимает остроту этой проблемы, решение вопроса зараженности посевов головней по-прежнему остается актуальным. Так, по данным Россельхозцентра 2016 г., в посевах яровых зерновых культур по Сибирскому федеральному округу пыльная головня яровой пшеницы была зарегистрирована на 33.9 тыс. га, ячменя – 11.14 тыс. га. В сравнении с Кемеровской, Новосибирской и Омской областями развитие заболевания на посевах яровой пшеницы в большей степени проявилось в Алтайском крае (23.5 тыс. га). В Вологодской области головней заражено более 7 тыс. га посевов, в Уральском федеральном округе пыльная головня пшеницы отмечена на площади 16.9 тыс. га. По данным А.В. Хариной (2013), потери урожая пшеницы от пыльной головни в Кировской области составляют в среднем 28 %. С учетом все возрастающей стоимости протравителей семян и их экологической небезопасности, самый доступный, снижающий пестицидную нагрузку на агроценозы способ защиты растений – это возделывание сортов, устойчивых к пыльной головне. Таким образом, проблема изучения и выявления эффективных источников и доноров генов устойчивости яровой пшеницы к пыльной головне весьма актуальна.

В России и за рубежом многие годы ведутся фитопатологические исследования по изучению мировой коллекции пшеницы и выделению источников резистентности

к пыльной головне. Согласно литературным данным, существует достаточно большой набор источников устойчивости к головневым болезням, выделенных в различных эколого-географических районах мира. Так, в работе (Плахотник и др., 2014) отмечается высокая устойчивость в условиях Поволжья у образцов к-17146 (Сирия), к-33809 (Мексика), к-60584 (Канада), Прохоровка, Тамбовчанка, Мерцана и др. Выделены источники устойчивости к популяции пыльной головни, распространенной на территории Красноярского края; это сорта Жигулевская (Россия), Московская 21 (Россия), Димитровка 5-18, Димитровка 5-2 (Болгария), Suanders, Pembina, Monitau (Канада) (Сидоров, 2001; Нешумаева и др., 2016). J. Nielsen (1983), оценивая коллекцию мягкой яровой пшеницы, выделял иммунные сорта из Африки, Британии, США, Канады: Sterling, New Pusa 201, Dominator, Glenlea, H-44-24, Hope и др. Однако чаще всего выделенные сорта обладают расоспецифическим типом устойчивости, и варьирование этого признака в значительной мере зависит от того, какие расы использовали исследователи для фитопатологических оценок.

Впервые существование рас *Ustilago tritici* (Pers.) Jens. различной вирулентности убедительно доказал F. Grevel (1930). Используя сорта с неодинаковой степенью устойчивости к различным спорным образцам возбудителя, он идентифицировал 4 расы. В Северной Америке W. Hanna и W. Popp (1932) были первыми, кто продемонстрировал существование разных рас в популяциях *U. tritici*. Hanna (1937) удалось выделить четыре физиологические формы пыльной головни из коллекций, собранных в Манитобе, с использованием дифференцирующего набора, состоящего из 11 сортов пшеницы. W. Cherewick (1953) описал 10 рас *U. tritici* из коллекций, собранных в Канаде, с помощью аналогичного, но не идентичного набора дифференциаторов. J. Nielsen (1987) на тест-наборе из 19 сортов и линий пшениц дифференцировал у *U. tritici* 41 расу (Menzies et al., 2003).

В России исследования по определению рас пыльной головни ведутся на дифференцирующем наборе яровой пшеницы, предложенном В.И. Кривченко (1984). Набор включает 6 сортов мягкой и 3 сорта твердой яровой пшеницы (табл. 1). Это связано с физиологической специализацией возбудителя к видам пшеницы – мягкой и твердой. Расы, специализированные к сортам *T. aestivum*, вирулентны на сортах и образцах вида *T. durum*, и наоборот: специализированные только к сортам твердой пшеницы не вызывают заражения образцов *T. aestivum*. Встречаются

широко специализированные расы, вирулентные на сортах мягких и твердых пшениц. При этом дифференциаторы к пыльной головне – это не изогенные линии, различающиеся по генам устойчивости, а генотипы, показывающие дифференцирующую реакцию различных изолятов *U. tritici* к тест-сортам.

Основная работа по идентификации рас, многие из которых установлены В.И. Кривченко и Д.В. Мягковой, проводилась в ВИР (Санкт-Петербург). В разные годы исследованиями занимались А.П. Шестакова, А.Е. Дружин (в Поволжье), В.С. Пенчукова, В.В. Плахотник, Л.А. Троицкая (в Казахстане), А.Н. Селянова (в Рязанской области), А.А. Грязнов (в Кустанайской области), Л.Ф. Тымченко (в Московской области) и др.

В Восточной Сибири подробные исследования расового состава пыльной головки пшеницы проведены В.Д. Тихомировым (1981). Им установлено 18 рас, три из которых (48, 51 и 54) широко распространены в регионах Красноярского края. В Западной Сибири Ж.А. Бахарева (1978) зарегистрировано 17 рас, из них расы 61, 62, 63, 64, 66 и 67 могут паразитировать как на мягкой, так и на твердой пшенице. Всего на российском наборе дифференциаторов установлена 71 раса.

В патосистеме пшеница–*U. tritici* взаимодействие патогена и растения-хозяина подчиняется теории Флора «ген на ген». Опираясь на эту теорию, J. Nielsen (1977, 1982) определил основные гены устойчивости к пыльной головне, внесенные в каталог генных символов (McIntosh et al., 2013) (табл. 2).

В настоящее время каталог дополнен новыми генами: у сорта Glenlea были определены четыре гена устойчивости – *Ut4*, *Ut5*, *Ut6*, *Ut8*, которые обеспечивают резистентность к канадским расам T-2, T-9, T-15 и T-27, но не эффективны против рас T-19, T-39 и T-10 (Knox et al., 2014). Донорами гена *Ut6* являются сорта AC Foremost, AC Karma, AC Vista, Chinese Spring, Glenlea, Oasis, NY 320, ген *Utx* предложено использовать как синоним гена *Ut5* (McIntosh et al., 2017).

Во всем мире селекционные программы, направленные на получение генотипов, не восприимчивых к болезням, опираются на генетический контроль растений к патогенам. Резистентность растений к возбудителям пыльной головки не является исключением, при этом многие сорта, тестируемые в одних зонах как устойчивые, другими исследователями могут идентифицироваться как восприимчивые. По словам В.И. Кривченко (1984), «продолжительность “полезной жизни” гена устойчивости в сорте в значительной мере зависит от состава популяций возбудителя болезни на определенной площади, и значение одного и того же гена устойчивости может быть неодинаково в различных зонах». Поэтому, изучая резистентность сортов к пыльной головне, в первую очередь необходимо знать патотипный состав популяции, ее стабильность или изменчивость. Понимание генетики устойчивости растения-хозяина в контексте с вирулентностью патогена позволяет выявить эффективные гены резистентности к доминирующим в определенной зоне расам.

Целью наших исследований было выделение среди генетического разнообразия яровой пшеницы невосприимчивых к пыльной головне образцов на фоне искусствен-

Table 1. Spring wheat varieties used for the differentiation of *Ustilago tritici* (Pers.) Jens. races (VIR collection)

VIR accession no.	Variety	Country
Soft wheat		
38519	Moskovka	Russia
14924	Kota	United States
8091	Preston	Canada
31238	Rümkers Dickkopf	Germany
8097	Reward	Canada
25019	Diamant	Sweden
Durum wheat		
39781	Akmolinka 5	Kazakhstan
5971	Mindum	United States
38514	Narodnaya	Ukraine

Table 2. Wheat genes for loose smut resistance

Gene	Donor(s)
<i>Ut1</i>	Florence/Aurore, Renfrew, Red Bobs
<i>Ut2</i>	Kota, Little Club
<i>Ut3</i>	Carma
<i>Ut4</i>	Thatcher/Regent
<i>Ut6</i>	D93213, P9163-BJ08*B, VIR 51658
<i>Utx</i>	Biggar

ного заражения растений популяцией, специфичной для Западно-Сибирского региона.

Материалы и методы

Материалом исследований служили 350 генотипов яровой пшеницы различного эколого-географического происхождения из мировой коллекции ВИР (Санкт-Петербург), образцы с идентифицированными генами устойчивости, а также сорта, районированные по Западно-Сибирскому региону (Приложение 1)¹. Изучение сортообразцов проводилось в период с 2011 по 2016 г. на искусственном инфекционном фоне фитопатологического участка СибНИИРС – филиала ИЦиГ СО РАН, расположенного в лесостепи Приобья.

Исследования по пыльной головне осложняются биологией гриба. В первый год растения, зараженные спорами гриба, формируют нормальное зерно. При этом мицелий гриба локализуется в щитке зародыша. И только на следующий год из зараженного зерна формируется поврежденный колос.

Инфицирование сортообразцов яровой пшеницы осуществляли вакуумным методом по методике ВИР. В качестве инокулюма использовали синтетическую популяцию патогена, которую создавали путем сбора телиоспор с зараженных колосов. Суспензию готовили непосредствен-

¹ Приложения 1–3 см. по адресу:
<http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx9.pdf>

но перед заражением, оптимальная инфекционная нагрузка 1–1.5 г телиоспор на 1 л воды. Растения яровой пшеницы инокулировали в период цветения, когда пыльники цветка в середине колоса желтого цвета. Зараженные растения отмечали этикетками. После созревания зараженные колосья обмолачивали на колосковой молотилке МКС-1.

Посев инокулированных семян пшеницы проводили в оптимальные для культуры сроки (вторая декада мая) в количестве 100 зерен на один метровый рядок. Высевали 200–300 зерен каждого образца. Учет осуществляли путем подсчета больных и здоровых колосьев с вычислением процента поражения.

Для классификации устойчивости использовали шкалу ВИР (1981): 0 – высокая устойчивость, поражение отсутствует; I – практическая устойчивость, поражение не превышает 5 %; II – слабая восприимчивость, поражение не превышает 25 %; III – средняя восприимчивость, поражение не превышает 50 %; IV – сильная восприимчивость, поражение более 50 % (Кривченко и др., 1987).

Сорта с поражением выше 25 % браковали в первый год изучения, устойчивые и слабо восприимчивые изучали в течение двух-трех лет. Об устойчивости сорта судили по максимальному поражению образца за годы исследований. В качестве восприимчивого стандарта были использованы сорта Скала и Сибирская 12.

Изучение физиологических рас возбудителя пыльной головки яровой пшеницы проводили согласно методическим рекомендациям ВИР. Специализацию рас *U. tritici* осуществляли на наборе сортов-дифференциаторов, который включал 6 сортов мягкой и 3 сорта твердой яровой пшеницы (см. табл. 1). Дифференциацию по устойчивости проводили по шкале: 0 – поражение отсутствует; 1 – поражение до 10 %; 2 – поражение свыше 10 %. Полученные результаты сравнивали с ключом для определения физиологических рас.

Результаты

Для пыльной головки характерен цветковый тип заражения. Возбудитель поражает растение в период цветения, и немаловажным фактором, оказывающим влияние на качество инокуляции, является окружающая среда. Несмотря на то что метеоусловия вегетационного периода за годы исследований различались по тепло- и влагообеспеченности, в целом для инфицирования растений складывались благоприятные условия. Исключение составил 2012 г., вегетационный период которого характеризовался как засушливый, с высокими температурами почвы и воздуха (табл. 3). Высокие температуры 2012 г. способствовали более быстрому прохождению растениями фаз развития и более короткому периоду, благоприятному для заражения. Для 2013 г. были характерны пониженные среднесуточные температуры воздуха (в среднем на 2–3 °C) и обилие осадков в виде проливных дождей начиная со второй декады июля, что также снизило качество инокуляции.

За 2012–2016 гг. в популяции пыльной головки идентифицировано 6 рас: расы 12, 23, 40, 53, 65 и 66 (табл. 4). В 2012, 2014 и 2016 гг. по реакции сортов-дифференциаторов зарегистрирована раса 12. Она не поражает сорта Московка и Preston, но вирулентна к Rümkers Dickkopf, Reward и Diamant. Сходную с ней реакцию показывает

раса 40, идентифицированная в 2015 г. и отличающаяся от расы 12 тем, что незначительно поражает сорт Московка (не более 5 %). Для расы 65, выделенной в 2013 г., характерны проявление слабой реакции на дифференциаторах Московка, Kota, Reward и неспособность к заражению сортов Preston и Rümkers Dickkopf. Раса 53, зарегистрированная в 2014 и 2016 гг., обладает вирулентностью к сортам Kota и Reward, не поражает дифференциатор Preston. В 2015 и 2016 гг. идентифицированы расы 23 и 66. Из дифференциаторов к ним восприимчивы Kota, Reward, Diamant. Поражение сортов Московка, Preston, Rümkers Dickkopf при инокуляции расой 23 не превышало 10 %. Раса 66 не поражает Московку и Preston, но вирулентна к Rümkers Dickkopf. Все расы обладают четко выраженной специализацией к сортам *T. aestivum* и не поражают сорта твердых пшениц Акмолинка 5, Народная и Mindum.

На основе анализа реакции дифференциаторов на инокуляцию возбудителем *U. tritici* В.И. Кривченко (1984) объединил все зарегистрированные расы в девять групп. Первые шесть групп поражают сорта мягких пшениц, седьмая и восьмая группы вирулентны для твердых, а девятая группа объединяет расы, специализированные к мягким и твердым пшеницам.

Идентифицированные нами расы 12 и 40 вошли во вторую группу, их объединяет то, что они агрессивны на средне- и сильновосприимчивых дифференцирующих образцах Rümkers Dickkopf, Reward, Diamant. Расы 23, 53 и 66 вошли в третью группу, у них четко выражена агрессивность к сорту Kota, и они, как правило, поражают сорта Rümkers Dickkopf, Reward и Diamant. Раса 65 входит в шестую группу, представители которой при сравнительной вирулентности обладают слабой агрессивностью.

Из табл. 4 видно, что высокий иммунитет к западносибирской популяции пыльной головки сохраняет сорт Preston, который, по данным В.И. Кривченко (1984), несет два гена устойчивости, условно обозначенные им как *R1R2*. В отдельные годы отмечается незначительное поражение сорта Московка, одним из родителей которого является Тулун 70 – отбор сорта Preston. Московка, как и Preston, обладает двумя доминантными генами устойчивости. Поражение сортов Kota с геном устойчивости *Ut2* и Rümkers Dickkopf, несущего один доминантный ген, колеблется по годам от слабо восприимчивых до восприимчивых. Сорта Reward и Diamant неустойчивы к местной популяции.

Известно, что генотип сорта в значительной мере оказывает влияние на расовый состав патогена. Это, в частности, подтверждается сравнительным анализом данных, полученных Ж.А. Бахарева (1978), с результатами современных исследований. Если в 1976–1978 гг. на территории Новосибирской области наиболее распространенными были расы 12, 23, 54, 61, 64, то в настоящее время – 12, 23, 40, 53, 65, 66. Интересно отметить, что расы 12 и 23 сохраняются в природе уже более 40 лет.

Для выявления эффективности генов *Ut* к западносибирской популяции было проведено изучение сортов с идентифицированными генами устойчивости, внесенных в Международный каталог генных символов. Ген *Ut1* контролирует устойчивость ко всем расам, присутствующим в западносибирской популяции пыльной головки. За годы исследований образцы Renfrew, Florence/Aurore,

Table 3. Weather conditions during inoculation of plant ears (data from the weather station Ogurtsovo)

Year	Number of days from sowing to infection	Temperature in July, °C				Amount of precipitation in July, mm			
		1st third	2nd third	Normal annual		1st third	2nd third	Normal annual	
				1st third	2nd third			1st third	2nd third
2011	41	16.0	19.2	19.3	19.7	21.0	13.0	16.0	19.0
2012	38	20.6	22.5	19.3	19.7	3.7	0	16.0	19.0
2013	51	16.7	20.6	19.3	19.7	11.3	30.6	16.0	19.0
2014	50	21.3	20.7	19.4	19.7	12.3	37.0	16.0	19.0
2015	43	18.0	21.0	19.3	19.6	61.0	4.8	15.0	19.0
2016	41	19.9	21.0	19.3	19.6	46.7	15.6	16.0	19.0

Table 4. Response scores of VIR differentiator varieties to *U. tritici* infection (2010–2016)

Race	Moskovka	Kota	Preston	Rümkers Dickkopf	Reward	Diamant	Akmolinka 5	Mindum	Narodnaya	Year
12	0	1	0	1–2	2	1–2	0	0	0	2012, 2014, 2016
23	1	2	1	1	2	2	0	0	0	2015, 2016
40	1	1	0	2	2	2	0	0	0	2015
65	1	1	0	0	1	2	0	0	0	2013
53	1	2	0	1	2	1	0	0	0	2014, 2016
66	0	2	0	2	2	2	0	0	0	2015, 2016

Red Bobs (*Ut1*) не поражаются пыльной головней. Раса 66 обладает вирулентностью к сортам Kota (*Ut2*), Thatcher/Regent (*Ut4*), Carma (*Ut3*) и Biggar (*Utx*). Максимальное поражение сортов за годы исследований составило 100, 20, 62.5 и 83.3 % соответственно. Сорт Carma неустойчив к расе 53.

Проведенные на фоне искусственного заражения исследования 350 коллекционных образцов яровой мягкой пшеницы позволили выделить генотипы, устойчивые к западносибирской популяции пыльной головни. Среди генофонда иностранной селекции наибольшее количество резистентных сортов выделяется в странах Североамериканской географической зоны (США, Канада, Мексика) (табл. 5).

Особый интерес представляют сорта Thatcher, DC II-21-44, Red Bobs, Renfrew, Preston, Fox, SWS “A” N80, AC Glenavon, AC Drummonga, Hoffman, Hope/Thatcher, Mercury, Neerawa, AC Minto, Rick, которые вошли в группу высоко- и практически устойчивых.

Сорта Marquis, Red Bobs, Renfrew, Preston – производные от сорта Red Fife, который, по одним источникам, был выделен из украинского сорта как Halychanka, по другим – Ostka Galicyjska. Этот сорт попал в Канаду из Гданьска (Польша), где в 1842 г. был переименован в сорт Red Fife (Maat, 2001).

R.D. McIntosh для сортов Renfrew и Red Bobs (производных Red Fife) определил ген устойчивости *Ut1* (McIntosh et al., 1998). Сорт Preston обладает двумя доминантными генами устойчивости, одним из которых является ген *Ut1* (Кривченко, 1984). Можно предположить наличие этого гена и у сорта Marquis, участвовавшего в создании сортов Hope, Kitchener, Thatcher (Symko, 1999) (Прил. 2). Сорт

Thatcher и его сестринская линия DC II-21-44 довольно часто вовлекались в селекцию при создании резистентных к пыльной головне сортов. Полученные от гибридной комбинации с участием Marquis и итальянского сорта твердой пшеницы Iumillo (Maat, 2001; McCallum, DePaum, 2008), эти образцы могут нести ген *Ut4* (Дружин, Крупнов, 2008). Другие исследователи (Сюков, Поротькин, 2014) предполагают, что сорта группы Thatcher, возможно, имеют ген устойчивости *UtTh*, идентичный, аллельный или тесно сцепленный с *Ut1*.

Общностью происхождения с Thatcher связаны сорта Neerawa, AC Glenavon, AC Barrie, BW-90 (Lupton, 1987; McCallum, DePaum, 2008). Сорту Hope устойчивость передана от российской полбы Yaroslav Emmer (Мартынов, Доброворская, 2003).

Сорта яровых пшениц немецкой селекции генетически базировались на озимых с использованием французского сорта Noe. Источником устойчивости и качества у них служил Heine Kolben, выделенный в XIX в. из французского сорта Saumur de Mars. Этот образец участвовал в родословной многих немецких сортов, в том числе Carma, несущего ген устойчивости *Ut3*, однако для условий Новосибирской области этот ген не эффективен. Кроме Heine Kolben, в качестве источников устойчивости к болезням в немецких сортах часто использовали канадские образцы Garnet и Thatcher (Lupton, 1987).

В сортах индийского происхождения, по данным С.П. Мартынова и Т.В. Доброворской (2003), основными источниками к пыльной головне были сорта Ostka Galicyjska, Thatcher и их производные.

Родоначальниками шведских сортов пшеницы были сорта Zeeuwse White и английский сорт Red Squarehead.

Table 5. Non-Russian spring soft wheat varieties resistant to loose smut (2012–2016)

Ecogeographical zone	Country	Resistant varieties
North America	United States	Cadet, DC II-21-44, H-44-24, Hope, Hope/Thatcher, Thatcher, k 31224, Mercury, Morris, Russell, Rick, Fox
	Canada	AC Barrie, Biggar, BW-90, Cuttler, AC Drummonga, Hoffman, Garnet, AC Glenavon, Kitchener, AC Minto, Marquis, Neepawa, Park, Preston, Red Bobs 222, Renfrew, RL 4277, Roblin, SWS "A" N80, Wisc 245
	Mexico	CBRD, k 31470, k 34636,
South America	Chile	CBP 566, k 34452,
	Peru	La Molina 82
Western Europe	Germany	Cardinal, Claudius, Diablon, Gzenos, Kvintus, Melissos, Probat, Sertori, Thasa, Sonet, Nachzos
	Switzerland	Calanda
	Netherlands	Blonda, Bastion
Northern Europe	Sweden	WWW Algot, Canon, WWW Sober
Eastern Europe	Belarus	Belorusskaya 928
	Ukraine	Kardinal, Kharkovskaya 18, Kharkovskaya 22
Southern Asia	India	Giriza, NP 790
Far East	China	Dong nong, PS-65, PS-87, PS-89, PS-90, Xing mai, Zong chin
Central Asia	Kazakhstan	Karagandinskaya 2, Karagandinskaya 29
Australia	Australia	Cunningham, Gabo, Yarralinka, Nardo, RAC-610, Stiletto

В австралийских сортах первоисточниками в селекции пшеницы являлись Fife из Канады и сорта из Индии. В результате их скрещивания был получен сорт Yandillac, который вместе с Purple Straw вошел в родословную сорта Federation. Этот сорт долгое время возделывался в производстве и использовался в селекции пшениц разных стран (Maat, 2001; Genetic Resources Information System for Wheat and Triticale GRIS).

В сортах украинской и белорусской селекции резистентность к пыльной головне возможна от твердой пшеницы Белотурка, российской полбы Yaroslav Emmer и мягкой пшеницы Полтавка через сорт Саратовская 29. Также в селекции на устойчивость участвовали американские источники: Thatcher и ее сестринская линия DC II-21-44 (через сорт Безенчукская 98), Red river 68, Hope, Marquis, Selkirk (Rabinovich et al., 1996; Мартынов, Добротворская, 2003).

По результатам фитопатологической оценки широкий набор резистентных генотипов к западносибирской популяции пыльной головки был выделен среди образцов отечественной селекции. В группы с высокой и практической устойчивостью к патогену вошли сорта: Московка (Немчиновка), Ленинградка, ГДС-24 (Северо-Западный НИИСХ), сорта самарской селекции – Тулайковская 5, Тулайковская 100, Тулайковская 10, Куйбышевская 1, Пирамида; Юлия (Пензенский НИИСХ); Саратовская 74, Фаворит, Юго-Восточная 2 (НИИСХ Юго-Востока); Воронежская 8, Воронежская 16 (НИИСХ ЦЧП им. В.В. Докучаева); Маргарита (Ульяновский НИИСХ), Кинельская нива (Поволжский НИИСХ им. Н.П. Константинова).

Проведенный С.П. Мартыновым и Т.В. Добротворской (2003) анализ родословных российских сортов позволил сделать вывод, что устойчивость к пыльной головне во многие отечественные сорта передана от стародавних об-

разцов Селивановский Русак, Полтавка (мягкая пшеница) и Белотурка (твердая пшеница) через сорт Саратовская 29, который прослеживается в генеалогии большинства российских сортов. Кроме того, в отечественной селекции использовались канадские источники. Так, в сорте Московка можно предположить наличие генов от Red Fife, полученных через канадские сорта Preston и Kitchener (селекция от Marquis). Сорт Ленинградка выведен с участием сирийского сорта Хоранка. В сортах ульяновской селекции в качестве источников резистентности были использованы канадские Pembina, DC II-21-44 и Red river (Genetic Resources Information System for Wheat and Triticale; Госреестр..., 2016).

Практическую устойчивость к пыльной головне самарские сорта Тулайковская 5 (происхождение – Эритросперм 865 (изогенная линия Саратовской 29)/АГИС 1) и ее производные Тулайковская 100 и Тулайковская 10 несут от *Ag. intermedium* через пшенично-пырейную замещенную линию Агис 1 (Мартынов и др., 2016).

В основе саратовских сортов лежит генетический материал от межвидовых скрещиваний твердой и мягкой пшениц Белотурка и Полтавка. Сорт Фаворит, являясь производной от Беянки, несет замещенную пырейную хромосому 6D/6Ai, не идентичную Тулайковской 5 и ее производным (Сибикеев, Дружин, 2017).

Сибирский генофонд яровых мягких пшениц включает также большое количество образцов, высокоустойчивых к местным расам пыльной головки. Это сорта Новосибирская 15, Памяти Вавенкова, Новосибирская 29, Новосибирская 22, Новосибирская 44, Баганская 95. Среди образцов омской селекции – Омская 20, Омская 29, Омская 33, Омская 36, Омская краса, Мелодия, Казанская юбилейная; сорта алтайской селекции – Алтайская 325, Алтайская 110, Алтайская жница, Сибирский альянс,

Алтайская 530; Памяти Афродиты (Кемерово), Минуса (Красноярск).

Устойчивость к пыльной головне в сорта Новосибирская 15 и Новосибирская 44 передана от DC II-21-44 (сестринская линия Thatcher) через Безенчукскую 98. Практическая устойчивость сортов Новосибирская 22 и Новосибирская 29, по-видимому, получена от сорта Красноярская, который в своей родословной также несет генетический материал от Thatcher через сорт Saunders. Кроме того, Новосибирская 29 несет гены устойчивости от Грекум 114. Резистентность к пыльной головне в сорт Памяти Вавенкова передана от итальянского Funello, несущего один доминантный и один рецессивный ген или два гена, взаимодействующих по типу доминантного эпистаза (Бахарева, Христов, 2003). Это сорт устойчив к расам пыльной головни 21, 37 и 38 (Дружин, Крупнов, 2008).

Близки по происхождению сорта омской селекции, выделившиеся по устойчивости. Основным источником устойчивости у них является пшенично-пырейный гибрид Грекум 114, донор генов пырея *Ag. intermedium* (Першина и др., 2013).

У резистентных сортов алтайской селекции в качестве источника использовалась селекционная линия Лютеценс 281, имеющая происхождение Грекум 114/Лют. 1210//Катерва (Коробейников и др., 2006).

Кемеровский сорт яровой мягкой пшеницы Памяти Афродиты несет гены от Саратовской 29 (через Кантегирскую 89) и от Грекум 114 (через Омскую 24). В родословной красноярского сорта Минуса присутствуют источники Красноярская и Грекум 114 (Нешумаева и др., 2016) (Прил. 3).

Заключение

В результате проведенных на искусственном инфекционном фоне исследований была получена информация об устойчивости 350 генотипов яровой пшеницы к возбудителю пыльной головни. Анализ родословной резистентных образцов позволил сделать вывод, что, несмотря на большое количество источников устойчивости, состав генов, определяющий признак невосприимчивости к патогену, небольшой. Эффективную защиту к местной популяции обеспечивают сорта, несущие гены *U1*, гены от Thatcher и ее сестринской линии DC II-21-44. В сортах российской селекции прослеживаются гены резистентности от Белотурки, Полтавки, Селивановского Русака (через Саратовскую 29 и ее производные), гены от пшенично-пырейных линий АГИС 1 и Грекум 114. Они контролируют признак резистентности к большинству рас патогена, распространенных в Западно-Сибирском регионе.

Сужение генотипического разнообразия вызывает необходимость поиска источников с новыми факторами устойчивости, так как массовое использование одних и тех же генов приводит к возникновению и накоплению в популяции вирулентных форм гриба, способных преодолевать устойчивость сортов.

Список литературы / References

Бахарева Ж.А. Изучение расового состава головневых заболеваний зерновых культур в Западной Сибири. Проблемы селекции и семеноводства полевых культур в Сибири: Сб. науч. тр. ВАСХНИЛ. Новосибирск, 1978;3-96.

[Bakhareva J.A. Study of the racial composition of smut diseases of grain crops in West Siberia. In: Problems of Breeding and Seed Production of Field Crops in Siberia: SB. science. Proceedings of the Lenin All-Union Academy of Agricultural Sciences. Novosibirsk, 1978;3-96. (in Russian)]

Бахарева Ж.А., Христов Ю.А. Создание сортов зерновых культур, устойчивых к головневым заболеваниям в Западной Сибири. Новосибирск: ПАСХН. Сиб. отд-ние, 2003.

[Bakhareva J.A., Khristov Yu.A. Creating Varieties of Grain Crops Resistant to Loose Smut Diseases in West Siberia. Novosibirsk: Siberian Branch of RAAS, 2003. (in Russian)]

Государственный реестр селекционных достижений, допущенных к использованию. 2016. <http://reestr.gossort.com>

[State Register of Breeding Achievements Admitted to Use. 2016. Available at: <http://reestr.gossort.com> (in Russian)]

Дружин А.Е. Влияние пыльной головни на хозяйственно полезные признаки яровой мягкой пшеницы. Селекция и семеноводство полевых культур: Сб. материалов Всерос. науч.-практ. конф. Пенза, 2001;24-26.

[Druzhin A.E. Influence of loose smut on commercially valuable traits of spring soft wheat. In: Breeding and Seed Industry of Field Crops: Proc. of the All-Russia scientific and practical conference. Penza, 2001;24-26. (in Russian)]

Дружин А.Е., Крупнов В.А. Пшеница и пыльная головня. Саратов, 2008.

[Druzhin A.E., Krupnov V.A. Wheat and Loose Smut. Saratov, 2008. (in Russian)]

Кривченко В.И. Устойчивость зерновых колосовых к возбудителям головневых болезней. М., 1984.

[Krivchenko V.I. The Resistance of Cereals to Agents of Loose Smut Diseases. Moscow, 1984. (in Russian)]

Кривченко В.И., Мягкова Д.В., Жукова А.Н., Хохлова А.П. Методические указания по изучению головнеустойчивости зерновых колосовых культур. Л., 1987.

[Krivchenko V.I., Myagkova D.V., Zhukova A.N., Khokhlova A.P. Guidelines for the Study of Resistance to Smut in Cereal Crops. Leningrad, 1987. (in Russian)]

Коробейников Н.И., Розова М.А., Кривогаарницын Б.И., Бородулина В.А. Принципы и результаты селекции зерновых культур на устойчивость к грибным заболеваниям на Алтае. Селекция на устойчивость растений к биотическим и абиотическим факторам среды: Материалы науч.-метод. конф. Новосибирск, 2006;41-59.

[Korobeynikov N.I., Rozova M.A., Krivogarnitsyn B.I., Borodulina V.A. Principles and results of breeding of cereals for resistance to fungal diseases in Altay. In: Breeding for Plant Resistance to Biotic and Abiotic Factors in the Environment: Proc. sci.-meth. conf. Novosibirsk, 2006;41-59. (in Russian)]

Мартынов С.П., Добротворская Т.В. Сравнительный анализ устойчивости яровой мягкой пшеницы к пыльной головне, основанный на генеалогическом подходе. Генетика. 2003;39(7):956-968. [Martynov S.P., Dobrotvorskaya T.V. Genealogy-based comparison of loose smut resistance for spring common wheat cultivars. Russ. J. Genet. 2003;39(7):799-810. DOI 10.1023/A:1024757205050.]

Мартынов С.П., Добротворская Т.В., Крупнов В.А. Генеалогический анализ использования двух видов пырея (*Agropyron*) в селекции мягкой пшеницы (*Triticum aestivum* L.) на устойчивость к болезням. Генетика. 2016;52(2):179-188.

[Martynov S.P., Dobrotvorskaya T.V., Krupnov V.A. Genealogical analysis of the use of two wheatgrass (*Agropyron*) species in common wheat (*Triticum aestivum* L.) breeding for disease resistance. Russ. J. Genet. 2016;52(2):154-163. DOI 10.1134/S1022795416020071.]

Нешумаева Н.А., Сидоров А.В., Голубев С.С. Селекция яровой пшеницы на устойчивость к пыльной головне. Достижения науки и техники АПК. 2016;30(6):22-24.

[Nechumaeva N.A., Sidorov A.V., Golubev S.S. Breeding of spring wheat for resistance to loose smut. Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC. 2016; 30(6):22-24. (in Russian)]

- Орлова Е.А., Бехтольд Н.П., Лихенко И.Е. Влияние возбудителя твердой головни ячменя на хозяйственно-полезные признаки растений. Достижения науки и техники АПК. 2015;29(3):4-6. [Orlova E.A., Bechtold N.P., Lihenko I.E. The influence of the causative agent of covered smut of barley on the economic-useful signs of the plants. Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC. 2015;29(3):4-6. (in Russian)]
- Першина Л.А., Осадчая Т.С., Бадаева Е.Д., Белан И.А., Россева Л.П. Изучение особенностей андрогенеза в культуре пыльников сортов и перспективной формы яровой мягкой пшеницы западносибирской селекции, различающихся наличием или отсутствием пшенично-чужеродных транслокаций. Вавиловский журнал генетики и селекции. 2013;17(1):40-49. [Pershina L.A., Osadchaya T.S., Badaeva E.D., Belan I.A., Rosseva L.P. Features of androgenesis in anther cultures of varieties and a promising accession of spring common wheat bred in Western Siberia differing in the presence or absence of wheat-alien translocations. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2013;17(1):40-49. (in Russian)]
- Плахотник В.В., Зеленева Ю.В., Судникова В.П. Источники и высокоэффективные доноры для селекции яровой пшеницы на устойчивость к стрессовым факторам среды. Вopr. соврем. науки и практики. 2014;1(50):109-113. [Plakhotnik V.V., Zeleneva U.V., Sudnikova V.P. Sources and high-performance donors for spring wheat breeding to increase resistance to ambient stress factors. Voprosy Sovremennoi Nauki i Praktiki = Issues of Modern Science and Practice. 2014;1(50):109-113. (in Russian)]
- Россельхозцентр. <http://rosselhoccenter.com/2016-02-06-19-01-37> [Russian Agricultural Center. <http://rosselhoccenter.com/2016-02-06-19-01-37>]
- Сибикеев С.Н., Друзин А.Е. Влияние замещения 6AG¹ (6D)-хромосомы Agropyron intermedium на хозяйственно ценные и адаптивные признаки у линий яровой мягкой пшеницы. Успехи соврем. естествознания. 2017;11:43-48. [Sibikeev S.N., Druzhin A.E. Effect of the substitution of 6AG¹ (6D) chromosome from Agropyron intermedium on agronomically valuable and adaptive traits in lines of spring bread wheat. Uspekhi Sovremennogo Estestvoznaniya = Advances in Modern Natural Science. 2017;11:43-48. (in Russian)]
- Сидоров А.В. Селекция яровой пшеницы на устойчивость к грибным болезням. Селекция и семеноводство. 2001;3:20-23. [Sidorov A.V. Breeding of spring wheat for resistance to fungal diseases. Seleksiya i Semenovodstvo = Breeding and Seed Production. 2001;3:20-23. (in Russian)]
- Степановских А.С. Головные болезни ячменя. Челябинск, 1990. [Stepanovskiyh A.S. Smut Disease of Barley. Chelyabinsk, 1990. (in Russian)]
- Сюков В.В., Поротькин С.Е. Генетика устойчивости мягкой пшеницы (*Triticum aestivum* L.) к пыльной головне (*Ustilago tritici* (Pers.) Jens.). Вавиловский журнал генетики и селекции. 2014;18(3):517-522. [Syukov V.V., Porotkin S.E. Genetics of common wheat (*Triticum aestivum* L.) resistance to loose smut (*Ustilago tritici* (Pers.) Jens.). Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2014;18(3):517-522. (in Russian)]
- Тихомиров В.Т. Расы пыльной головни *Ustilago tritici* (Pers.) Jens. в Красноярском крае. Сиб. вестн. с.-х. науки. 1981;1:36-40. [Tikhomirov V.T. Races of loose smut *Ustilago tritici* (Pers.) Jens. in the Krasnoyarsk region. Sibirskiy Vestnik Selskokhozyaystvennoy Nauki = Siberian Herald of Agricultural Sciences. 1981;1:36-40. (in Russian)]
- Харина А.В. Головные болезни яровой мягкой пшеницы в условиях Евро-Северо-Востока. Аграр. наука Евро-Северо-Востока. 2013;1:15-18. [Harina A.V. Bunt diseases of spring soft wheat under conditions of Euro-North-East. Agrarnaya Nauka Evro-Severo-Vostoka = Agricultural Science of the Euro-North-East. 2013;1:15-18. (in Russian)]
- Cherewick W.J. Smut Diseases of Cultivated Plants in Canada. Ottawa, Canada: Dept. of Agriculture, 1953.
- Hanna W.F. Physiologic forms of loose smut of wheat. Can. J. Res. 1937;15(4):141-153.
- Hanna W.F., Popp W. Physiologic forms of loose smut of wheat. Phytopathology. 1932;22(1):141.
- Genetic Resources Information System for Wheat and Triticale GRIS. Available at: <http://wheatpedigree.net>
- Grevel F.K. Untersuchungen über das Vorhandensein biologischer Rassen des Flugbrandes des Weizens (*Ustilago tritici*). Phytopathologische Zeitschrift. 1930;2(3):209-234.
- Knox R.E., Campbell H.L., Clarke F.R., Menzies J.G., Popovic Z., Procunier J.D., Clarke J.M., Depauw R.M., Cuthber R.D., Somers D.J. Quantitative trait loci for resistance in wheat (*Triticum aestivum*) to *Ustilago tritici*. Can. J. Plant Pathol. 2014;36(2):187-201.
- Lupton F. Wheat Breeding: Its scientific basis. Dordrecht, Netherlands: Springer, 1987.
- Maat H. A History of Agricultural Science in the Netherland and its Colonies, 1863–1986. Kluwer Acad. Publ., 2001.
- McCallum B.D., DePaum R.M. A review of wheat cultivars grown in the Canadian prairies. Can. J. Plant Sci. 2008;88(4):649-665.
- McIntosh R.A., Dubcovsky J., Rogers W.J., Morris C., Xia X.C. Catalog of Gene Symbols for Wheat: 2017 Supplement. 12th Int. Wheat Genetics Symp., Yokohama, Japan, 2017.
- McIntosh R.A., Hart C.E., Devos K.M., Gale M.D., Rogers W.J. Catalogue of Gene Symbols for Wheat. Proc. 9 Int. Wheat Genet. Symp., Saskatchewan, Canada, 1998;5.
- McIntosh R.A., Yamazaki Y., Dubcovsky J., Rogers J., Morris C., Appels R., Xia X.C. Catalogue of Gene Symbols for Wheat. 12th Int. Wheat Genet. Symp., Sept. 8–13, 2013;195.
- Menzies J.G., Knox R.E., Thomas P.L. Virulence of Canadian isolates of *Ustilago tritici*: 1964–1998, and the use of the geometric rule in understanding host differential complexity. Can. J. Plant Pathol. 2003;25:62-72.
- Nielsen J. Inheritance of virulence of loose smut of wheat *Ustilago tritici* on the differential cultivars Renfrew, Florence × Aurore, Kota and Little Club. Can. J. Bot. 1977;55:260-263.
- Nielsen J. Inheritance of virulence of loose smut of wheat *Ustilago tritici* on the differential cultivars Carma, Red Bobs, and derivative of the cross Thatcher × Regent. Can. J. Bot. 1982;60:1191-1193.
- Nielsen J. Spring wheat's immune or Highly resistant to *Ustilago tritici*. Plant Dis. 1983;67(8):860-863. DOI 10.1094/PD-67-860.
- Nielsen J. Races of *Ustilago tritici* and techniques for their study. Can. J. Plant Pathol. 1987;9:91-105.
- Nielsen J., Thomas P., Loose smut. In: Wilcoxson R.D., Saari E.E. (Eds.) Bunt and Smut Diseases of Wheat: Concepts and methods of disease management. Mexico, CIMMYT, 1996;33-47.
- Rabinovich S.V., Chernyaeva I.N., Afonskaya E.J., Dolgova E.M. Genetic basis of loose smut resistance in Ukrainian, Russian, and U.S. winter and spring wheats. Ann. Wheat Newslett. 1996;42: 210-212.
- Symko S. From a Single Seed: Tracing the Marquis wheat success story in Canada to its roots in the Ukraine. Canada: Agriculture and Agri-Food Canada Publ., 1999.

ORCID ID

Е.А. Орлова orcid.org/0000-0001-5084-375X

Acknowledgements. This work was supported by State Budgeted Project 0324-2019-0039 for ICG.

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

Received April 5, 2019. Revised April 25, 2019. Accepted May 6, 2019.

Происхождение, селекция и адаптация российских пород крупного рогатого скота по данным полногеномных исследований

Н.С. Юдин¹, Д.М. Ларкин^{1, 2} 

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

² Королевский ветеринарный колледж, Лондонский университет, Лондон, Великобритания

 e-mail: dmlarkin@gmail.com

В обзоре рассмотрены недавние работы по истории происхождения и следам отбора в геномах исконно российских пород крупного рогатого скота (КРС). Большинство из этих работ недоступно для неанглоязычной аудитории. Показана история происхождения российских пород с точки зрения генетики и связь многих из них с породами европейского, а некоторых – с породами азиатского происхождения. Даны описания основных кластеров пород и указаны породы, сохранившие свой уникальный генофонд и, соответственно, нуждающиеся в охране. Кроме того, приведены результаты поиска следов селекции и адаптации в геномах российских пород КРС. Описаны уникальные следы отбора, представленные у азиатских «турано-монгольских» пород, в частности у якутской, которые, вероятно, позволили этой породе адаптироваться к суровым условиям жизни в Заполярье. Приведены также описания генов, по-видимому, способствовавших адаптации и других местных пород КРС к климатическим условиям Российской Федерации. Показано, что геномы российских пород КРС содержат следы отбора в районах известных генетических маркеров доместикиции, которые претерпели изменения в результате одомашнивания КРС человеком 8–10 тыс. лет назад. Наиболее яркие из них – это гены изменения окраски шерстного покрова, который во многих случаях служил признаком, различавшим примитивные породы. Полногеномный анализ ассоциаций у российских пород позволил выявить новый ген-кандидат, связанный с фенотипом «белая голова», и ген, связанный с поддержанием температуры тела при сильном холодном стрессе. Обобщенные в обзоре данные могут быть использованы для выбора генетических маркеров для селекции (и/или геномного редактирования) при создании нового поколения высокопродуктивных пород, адаптированных к условиям Российской Федерации и других стран со схожими климатическими условиями.

Ключевые слова: крупный рогатый скот; порода; адаптация; отбор; Россия; однонуклеотидный полиморфизм; климат; полногеномный анализ ассоциаций.

Для цитирования: Юдин Н.С., Ларкин Д.М. Происхождение, селекция и адаптация российских пород крупного рогатого скота по данным полногеномных исследований. Вавиловский журнал генетики и селекции. 2019;23(5): 559-568. DOI 10.18699/VJ19.525

Whole genome studies of origin, selection and adaptation of the Russian cattle breeds

N.S. Yudin¹, D.M. Larkin^{1, 2} 

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

² Royal Veterinary College, University of London, London, United Kingdom

 e-mail: dmlarkin@gmail.com

Our review presents several recent studies on the genetic history and signatures of selection in genomes of the native Russian cattle breeds. Most of these works are not easily accessible for the Russian-speaking audience. We describe the origins of appearance of the Russian cattle breeds from the genetics perspective. We point to the links between most of the Russian breeds with the taurine breeds of the European origin and for some Russian breeds with the breeds of the Asian origin. We describe major phylogenetic clusters of the Russian breeds and point to those that still maintain their unique genetics, meaning that their preservation is a priority. In addition, we review the results of the search for signatures of selection in genomes of the Russian cattle breeds. Some unique signatures of selection present in the genomes of so-called “turano-mongolian” cattle (i.e. the Yakut cattle) are described which allowed the Yakut cattle to adapt to harsh environments found above the Polar Circle. Signatures of selection which could help other cattle breeds of the Russian origin to adapt to various climatic condition of the Russian Federation are reviewed. The Russian cattle genomes also contain known signatures of selection related to cattle domestication about 8–10 thousand years ago. The most profound ones include genes related to changes of the coat colour. This phenotype in many cases could be related to the distinction of the first domesticated populations and lead to the formation of so-called land races (primitive breeds). Whole-genome association studies of Russian cattle breeds pointed to a novel gene which could be related to the “white-faced” phenotype and to a gene which is related to body temperature support under the acute cold stress. The data presented in our review could be used for identification of genetic markers to focus on in future

efforts on designing new highly productive cattle breeds adapted to climates of the Russian Federation and other countries with similar climates.

Key words: cattle; breed; adaptation; selection; Russia; single nucleotide polymorphism; climate; genome-wide association study.

For citation: Yudin N.S., Larkin D.M. Whole genome studies of origin, selection and adaptation of the Russian cattle breeds. *Vavilovskii Zhurnal Genetiki i Selekcii* = Vavilov Journal of Genetics and Breeding. 2019;23(5):559-568. DOI 10.18699/VJ19.525 (in Russian)

Введение

Тысячелетия искусственного отбора в сочетании с направляемой людьми миграцией и адаптацией к различным условиям окружающей среды привели к появлению около 1000 пород крупного рогатого скота¹ по всему миру (Buchanan, Lenstra, 2015). Эти породы приспособлены к местным экономическим потребностям, эстетическим требованиям и обладают уникальными генетическими профилями (Дмитриев, 1978). В течение последних двухсот лет ряд примитивных пород крупного рогатого скота был значительно улучшен, в результате чего появилось несколько коммерческих пород, которые при правильном уходе демонстрируют рекордные характеристики (Voichard, Brochard, 2012). В настоящее время существует тенденция к замене местных пород или к их улучшению генетическим материалом от выдающихся коммерческих пород. Это означает, что генетическое разнообразие, следы адаптации к местным условиям и информация об исторических связях, закодированные в геномах местных пород, часто бывают безвозвратно утеряны еще до того, как они были должным образом описаны и изучены (Gao et al., 2015). С другой стороны, геномы местных популяций (мы используем термины «порода» и «популяция» как синонимы, поскольку определение породы в разных странах может варьировать) могут быть использованы как источники комбинаций генетических вариантов, неоценимых в развитии нового поколения коммерческих пород, которые лучше вписывались бы в разнообразие условий окружающей среды (Gao et al., 2017). Первым шагом на пути к раскрытию этой информации является понимание происхождения, генетической структуры и событий гибридизации местных пород, а также их включение в контекст обширного набора мировых пород (Bovine HarMap Consortium et al., 2009; Matukumalli et al., 2009; Beynon et al., 2015).

Появление недорогих ДНК-чипов для генотипирования однонуклеотидных полиморфизмов (SNP) позволило выявить генетические профили различных пород одомашненных видов, разработать обоснованные стратегии для их улучшения, с одной стороны, и узнать о генетических процессах, сопровождающих одомашнивание и формирование породы, с другой. Хотя большинство усилий направлено на изучение популярных коммерческих пород, например тексель у овец (Mucha et al., 2015) и голштино-фризской у крупного рогатого скота (van Binsbergen et al., 2015), растет интерес к генетике малочисленных местных пород из-за уникальных адаптаций и их потенциального вклада в решение проблем сельского хозяйства, связанных с изменением окружающей среды (например, гло-

бальное потепление) и локальной устойчивостью к патогенам (Beynon et al. 2015).

Генетическое разнообразие домашнего скота связано с двумя независимыми событиями domestikации вымершего в настоящее время тура (*Bos primigenius*), которые примерно в VIII–V тысячелетии до н. э. (Loftus et al., 1994) привели к возникновению современных видов домашнего скота – *Bos indicus* (горбатого или индийского скота) и *Bos taurus* (безгорбого или тауринного скота) в долине Инда и на Ближнем Востоке соответственно. Некоторые современные популяции в Китае, Африке и Америке произошли от скрещивания между *B. taurus* и *B. indicus* (Ларкин, Юдин, 2016). Формирование других популяций на островах Юго-Восточной Азии происходило с участием бантенга (*B. javanicus*) (Decker et al., 2014).

При анализе генетического разнообразия 134 мировых пород крупного рогатого скота с помощью полногеномного генотипирования на ДНК-чипе Illumina BovineSNP50 BeadChip были выделены три основные группы пород: азиатские индийские, евразийские тауринные и африканские тауринные (Decker et al., 2014). Своеобразие африканских тауринных пород, вероятно, связано с их происхождением в результате гибридизации domestikцированных тауринных предков с диким африканским туром. Европейские породы крупного рогатого скота состоят в основном из животных, произошедших от *B. taurus*, без значительного вклада генов *B. indicus*, за исключением нескольких итальянских пород. Кроме того, у итальянских и иберийских популяций крупного рогатого скота имеется также значительный генетический компонент, восходящий к африканским тауринам (Decker et al., 2014). Многие европейские породы, включая линкольнскую красную, бельгийскую голубую, санта-гертруда, бифмастер и др., имеют общих предков с шортгорнской породой. Однако до последнего времени масштабные исследования мировых пород крупного рогатого скота не включали породы из России.

Первые полногеномные исследования отдельных российских пород скота

Недавние исследования по полногеномному генотипированию коренных европейских пород скота включали несколько коренных российских пород (Зиновьева и др., 2016; Iso-Touru et al., 2016; Upadhyay et al., 2017). Однако при этом авторы, как правило, не проводили исчерпывающего сравнения между российским скотом и мировыми породами.

В упомянутых работах была показана высокая дивергенция якутского крупного рогатого скота по сравнению с девятью другими породами (Iso-Touru et al., 2016), что хорошо соответствует полученным ранее данным по по-

¹ В данной статье под этим термином подразумеваются популяции двух подвидов: *Bos taurus* и *Bos indicus*.

лиморфизму мтДНК, Y-хромосомы и других микросателлитных ДНК-маркеров (Kantanen et al., 2009; Li, Kantanen, 2010). Животные калмыцкой породы оказались близки к сербской породе буша (Iso-Touru et al., 2016). Исследование генетических профилей пяти молочных пород (бестужевской, холмогорской, костромской, красной горбатовской и ярославской) показало, с одной стороны, их обособленность от европейской голштино-фризской популяции, а с другой – значительную интрогрессию голштинского скота в генофонд бестужевской породы (Зиновьева и др., 2016). По результатам анализа SNP маркеров девять российских пород и 45 пород Евразии были сгруппированы по генетическому сходству в три кластера (Sermuyagin et al., 2018). Первый кластер состоял из турано-монгольских пород (якутской и калмыцкой), которые оказались наиболее удаленными от всех остальных пород. Черно-пестрая, костромская и суксунская породы образовали второй кластер, поскольку в их формировании участвовал ряд пород из Западной Европы (голландская, бурая швицкая, красная датская). В третий кластер вошли бестужевская, холмогорская, красная горбатовская и ярославская породы, которые имели наименьшую интрогрессию чужеродного генетического материала и потому могут считаться чистокровными породами российского происхождения.

Из-за уникального географического положения России (и в Европе, и в Азии), ее огромной территории с разнообразными климатическими условиями и богатой истории можно ожидать, что отечественный рогатый скот будет демонстрировать множество уникальных адаптаций и, вероятно, покажет генетическую связь с европейскими и азиатскими популяциями крупного рогатого скота. Согласно историческим данным, современные российские породы происходят от древнего евразийского крупного рогатого скота, в том числе степного (Li, Kantanen, 2010). Позднее, с начала XVIII в., они подвергались неконтролируемому скрещиванию с многочисленными европейскими популяциями крупного рогатого скота (голландский, тирольский, симментальский, швицкий скот и т.д.) (Dmitriev, Ernst, 1989). В настоящее время в России разводят 16 отечественных пород скота (Дунин, Данкверт, 2013).

Результаты полногеномного генотипирования основных российских пород в контексте мировых пород крупного рогатого скота

Целью нашего недавнего исследования было проанализировать в контексте мировых пород популяции крупного рогатого скота из России и породы из ближнего зарубежья, разводимые в России (Yurchenko et al., 2018). Мы использовали ДНК-чипы GGP HD150K и Illumina Bovine 50K для генотипирования животных 18 пород, разводимых в России, отобранных на основе вероятного исторического вклада местных популяций скота в их современные геномы. Наряду с популярными российскими породами (например, черно-пестрой или холмогорской) были включены высокоспециализированные породы, которые демонстрируют хорошую адаптацию к специфической внешней среде (например, якутская) и/или почти исчезли (например, бурятская). Таким образом, был представлен самый большой и самый полный набор пород крупного рогатого

скота из России, доступный для популяционно-генетических исследований до сих пор.

Наши данные генотипирования мы объединили с набором данных, содержащим дополнительные 129 пород крупного рогатого скота, собранных по всему миру и описанных в работе Декера с соавторами (Decker et al., 2014), и десятью породами из России и Европы, которые были генотипированы ранее и описаны в статье Исо-Тору с соавторами (Iso-Touru et al., 2016). Мы стремились использовать опубликованные ресурсы для выявления генетической структуры и филогенетической истории отечественного скота, а также для выработки гипотез о его взаимоотношениях с породами по всему миру. Чтобы выявить сложную историю российских пород крупного рогатого скота, использовались многочисленные взаимодополняющие методы популяционной генетики, а гипотезы, касающиеся происхождения и структуры современных пород, были построены на основе интеграции этих результатов.

Филогенетический анализ российских и зарубежных пород крупного рогатого скота

Методом объединения ближайших соседей (neighbor-joining) было построено филогенетическое дерево российских пород крупного рогатого скота совместно с данными о 134 мировых породах (рис. 1). В соответствии с географическим положением России и ее историческими и торговыми связями с другими странами, большинство российских пород демонстрировали общее происхождение с породами тауринного крупного рогатого скота из Европы. Для большинства российских пород мы не могли четко идентифицировать на филогенетическом дереве их сестринские зарубежные породы, за исключением того, что европейский и российский скот находится на одном и том же широком филогенетическом узле. Однако несколько примеров, когда наши данные подтвердили уже известные исторические взаимоотношения между российским крупным рогатым скотом и некоторыми иностранными породами, говорят о надежности наших результатов.

Из связей, выявленных в нашем исследовании и подтвержденных известными историческими отношениями, самые выраженные оказались между казахской белоголовой и герефордской породами, как из России, так и из Европы, что хорошо подкрепляется историей происхождения казахской белоголовой породы. Эта порода была сформирована в период с 1930 по 1950 г. путем скрещивания турано-монгольского казахского и калмыцкого скота с герефордами в Казахской ССР (Dmitriev, Ernst, 1989). Другим примером выявленных генетических взаимоотношений, подтвержденных известной историей формирования породы, была кластеризация костромской, бурой швицкой, бурой немецкой и алатауской пород (Dmitriev, Ernst, 1989).

Отдельный узел на филогенетическом дереве, образованный холмогорской, голштино-фризской, черно-пестрой и несколькими родственными европейскими породами, вероятно, отражает исторические взаимоотношения, которые прослеживаются вплоть до XVII в., когда начала формироваться и затем была скрещена с голландским скотом холмогорская порода (Dmitriev, Ernst, 1989). Однако анализ структуры популяций показывает, что генетическая

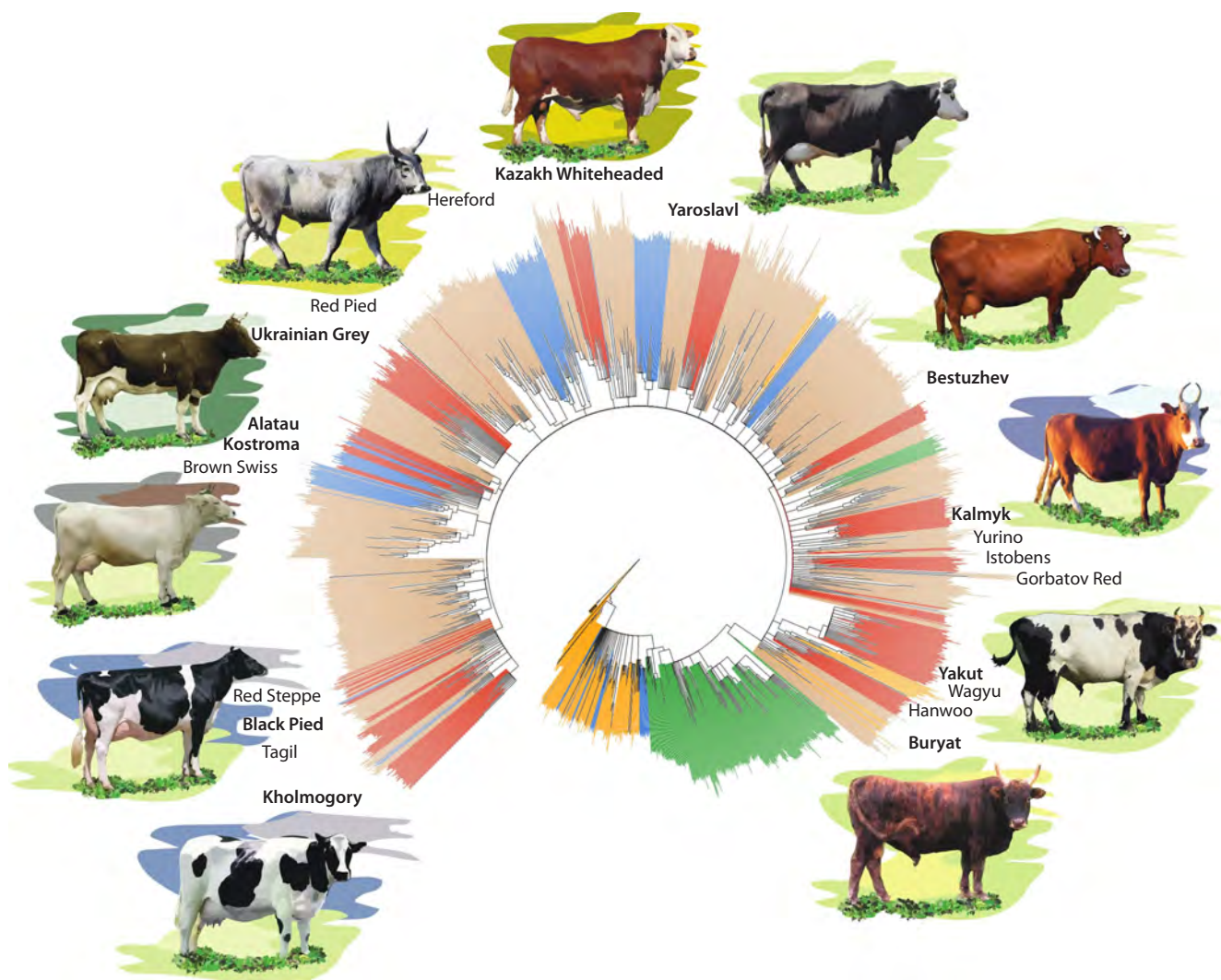


Fig. 1. Phylogenetic tree of the genetic diversity of Russian and non-Russian cattle breeds built using the neighbour-joining approach (Yurchenko et al., 2018).

Asian breeds are shown in yellow; African, in green; American, in blue; European, in beige; Russian, in red. Names of the Russian breeds and their sister breeds are indicated. Breeds shown in the images are indicated in boldface. Dairy breeds are represented by cows, and beef and dual-purpose, by oxen.

составляющая современной голштино-фризской породы в холмогорской породе относительно невелика и что холмогорскую породу следует считать генетически обособленной, что подтверждается и предыдущими наблюдениями (Зиновьева и др., 2016). С другой стороны, наши образцы российской черно-пестрой породы демонстрируют очень низкую дифференциацию от голштино-фризского скота. Это позволяет предположить, что использование импортных голштино-фризских быков или их семени в России могло существенно повлиять на генетику отечественной черно-пестрой породы.

Следует отметить отдельный кластер, сформированный турано-монгольской якутской породой, породой ханву из Кореи и вагу из Японии. Эти породы кластеризуются и с бурятским скотом.

Проведенный нами анализ общих гаплотипов, который был основан на коротких гаплотипных блоках (предположительно отражающих взаимоотношения предков), еще раз подтвердил сложную историю формирования пород

российского скота европейского происхождения, выявил описанные выше взаимоотношения, а также показал, что серый украинский крупный рогатый скот филогенетически связан с подольской и романьольской породами, подтверждая отнесение серого украинского скота к группе примитивных подольских пород (Кушнир, Глазко, 2009).

Анализ популяционной структуры российских пород крупного рогатого скота

В дополнение к обширным генетическим связям с крупным рогатым скотом европейского происхождения, все методы популяционного анализа наводят на мысль, что в России существуют породы, имеющие общих предков с крупным рогатым скотом из Азии. В соответствии с этим якутский, бурятский и калмыцкий крупный рогатый скот группировался с турано-монгольскими и другими азиатскими тауриновыми породами как на филогенетическом дереве, так и на графиках программы STRUCTURE, которая объединяет данные генотипирования отдельных

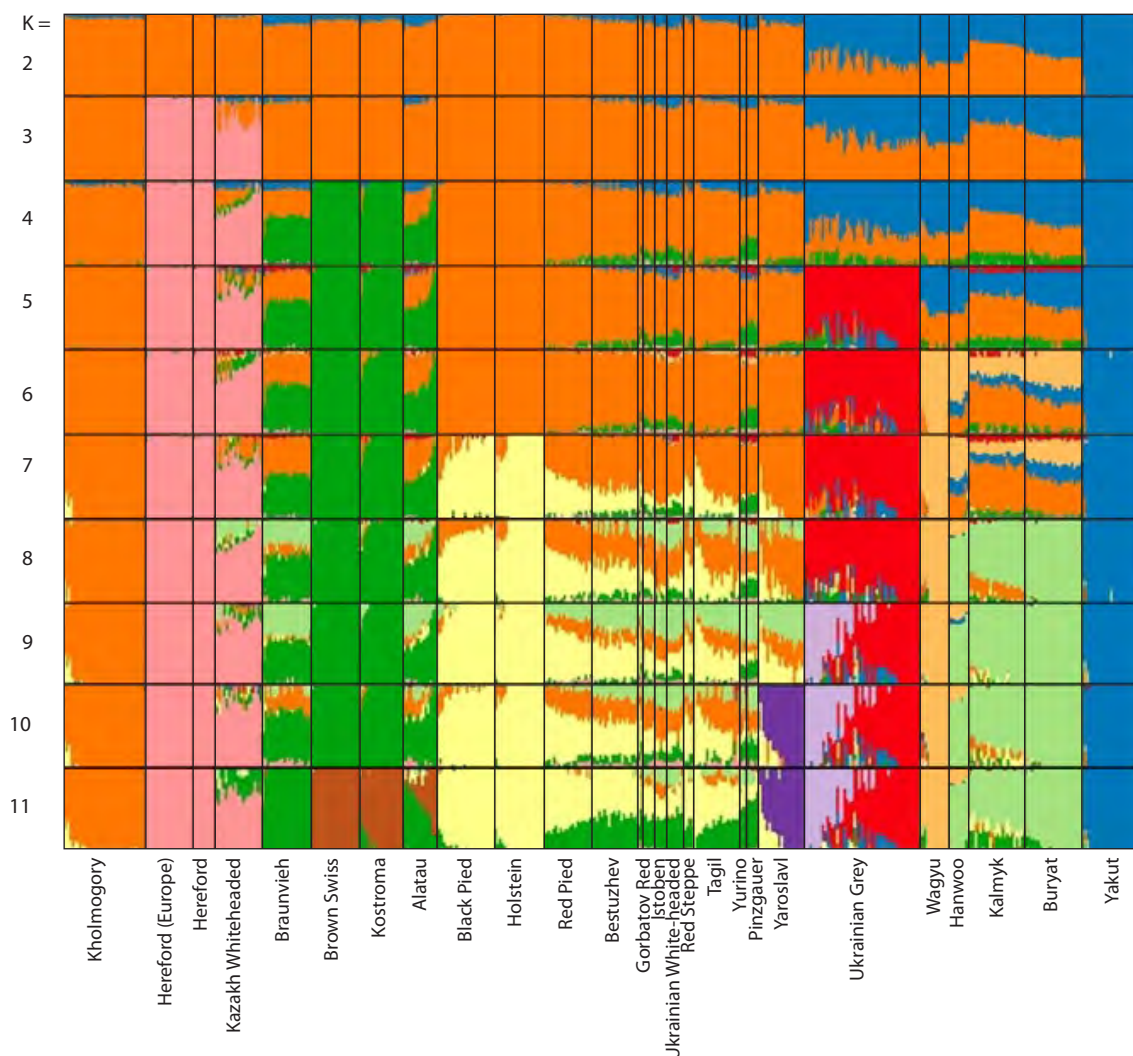


Fig. 2. Results of STRUCTURE analysis for Russian and several closely related non-Russian cattle breeds: Hereford, Braunvieh, Brown Swiss, Holstein, Red Pied Lowland, Pinzgauer, Wagu, Hanwoo (Yurchenko et al., 2018).

Different colors indicate different populations to which the program classified individuals according to the chose number of populations (K). Breeds shown with several colors are likely to have originated by crossing to other populations.

особей в группы согласно заданному пользователем набору популяций. При этом выявляются как «чистокровные», так и «смешанные» особи (рис. 2). Так, на графике STRUCTURE глобального набора данных якутская порода формирует первый породоспецифический кластер после наблюдаемой дивергенции *B. indicus* и африканских тауринов, а на филогенетическом дереве она находится в узле с бурятским крупным рогатым скотом и другими азиатскими тауриновыми породами. Точная причина того, что результаты программы STRUCTURE и филогенетического анализа настолько расходятся для якутского крупного рогатого скота, в настоящее время не ясна. Возможно, это обусловлено низким историческим эффективным размером популяции в сочетании с длительной изоляцией от других пород. Более близкое родство якутского крупного рогатого скота с другими дивергировавшими азиатскими турано-монгольскими породами может означать их раннее отделение от остальной части тауринного генофонда или даже независимую domestикацию в Азии (Mannen et al.,

2004). Результаты анализа общих гаплотипов тоже подтверждают эти взаимоотношения в пределах выборки турано-монгольских пород, помещая якутский крупный рогатый скот в один кластер с бурятской породой и японской вагу, а также указывая на связи с корейской ханву и монгольским крупным рогатым скотом.

Бурятский скот считался вымершим в России до недавнего времени, когда сохранившееся чистокровное стадо было обнаружено в Монголии и завезено обратно в Россию, чтобы начать восстановление этой породы. Наши результаты продемонстрировали, что бурятский скот действительно имеет больше общих гаплотипов с якутским скотом и вагу, чем с породами из Монголии, что указывает на его отдельное от монгольского крупного рогатого скота происхождение.

Положение третьей турано-монгольской породы в нашем списке, калмыцкой, остается неясным. Она формирует отдельный кластер в пределах европейского скота на филогенетическом дереве, однако результаты анализа

STRUCTURE наводят на мысль о ее общем предке с бурятской породой и ханву. Выявление общих гаплотипов показало недавнюю обширную гибридизацию калмыцкой породы с шортгорнской, что может объясняться использованием шортгорнов для улучшения калмыцкого скота в бывшем СССР (Dmitriev, Ernst, 1989). Вероятно, это повлияло на генофонд породы и ее положение на филогенетическом дереве, маскируя ожидаемые предковые взаимоотношения, которые были выявлены только программой STRUCTURE.

Интересно отметить, что по результатам выявления общих гаплотипов и анализа программой Treemix, которая определяла наиболее вероятные случаи и направления исторического смешения популяций, мы не обнаружили каких-либо существенных свидетельств гибридизации российских пород и зебувидного скота. Однако график STRUCTURE для глобальной выборки при заданном наборе популяций (в нашем случае равном двум, т. е. *B. indicus* и *B. taurus*) свидетельствовал о небольшом вкладе зебувидных предков в турано-монгольские породы. Это наблюдение позволяет предполагать очень древнее и, вероятно, незначительное событие гибридизации, не детектируемое другими методами. Также возможно, что среди локусов SNP на ДНК-чипе BovineSNP50K (и, как результат, в наборе SNP, используемом в настоящей работе) преобладают локусы, свойственные тауринному скоту, и предковые SNP, общие между популяциями тауринного и зебувидного скота (McKay et al., 2008), что могло в некоторой степени повлиять на наши результаты и замаскировать гибридизацию с *B. indicus*. Для того чтобы пролить дополнительный свет на причины наблюдаемой дивергенции якутского скота, потребуются более подробные исследования, включающие полногеномное секвенирование российских пород скота и их сравнение как с тауринным, так и с зебувидным референсными геномами. При внутривидовом анализе или в контексте только самых родственных мировых пород российские породы крупного рогатого скота демонстрировали умеренный уровень генетического разнообразия и сопоставимые с другими евразийскими породами оценки эффективных размеров популяций (Iso-Touru et al., 2016). Холмогорская и якутская породы были подтверждены как наиболее генетически различающиеся в пределах набора пород из России и родственных им евразийских пород по результатам STRUCTURE, поддерживаемым высокими значениями индекса фиксации, который указывает на различие частот минорных аллелей SNP между породами. В отдельном кластере пород было очевидным сильное влияние голштино-фризской породы.

Ярославская порода отделилась от этого кластера при заданном количестве популяций, равном 10, будучи последней русской породой крупного рогатого скота, которая имела уникальный генофонд, в то время как другие породы (исходя из данных наших образцов), в том числе черно-пестрая, тагильская, бестужевская, истобенская, юрьинская и украинская белоголовая, демонстрировали различный уровень вклада голштино-фризской породы в их генофонд. Из этого можно предполагать, что в перечисленных породах сохранились относительно небольшие доли предковых аллелей от местных популяций, что под-

тверждается низким уровнем популяционной дифференциации в этой группе. Однако нельзя исключать, что на результаты могли повлиять небольшие размеры выборки для некоторых пород в нашем списке.

Таким образом, нами проведен первый подробный анализ популяционной генетики и филогенетических связей практически всех основных пород крупного рогатого скота, разводимых в настоящее время в России, которые, вероятно, возникли из местных популяций крупного рогатого скота и адаптировались к суровым внешним условиям, включая климат. Некоторые из изученных пород (холмогорская, якутская, ярославская) имеют обособленные генетические профили, что делает их приоритетными объектами для выявления признаков отбора и адаптаций, связанных с местными условиями среды. С другой стороны, большая группа пород (коостромская, казахская белоголовая, истобенская и др.) испытывала как прошлое, так и недавнее влияние со стороны коммерческих европейских пород, например голштино-фризской. По-видимому, геномы этих пород содержат лишь незначительное число предковых аллелей, которые важны для выживания в местных условиях. Поэтому такие породы могут быть использованы для генетического анализа экономически важных признаков методом картирования примесей (admixture mapping) (Kassahun et al., 2015).

Филогенетические связи между российскими породами и породами из других стран, выявленные в нашем исследовании, станут основой для будущей работы по изучению аллелей или гаплотипов, непосредственно влияющих на фенотип. Полученная информация позволит оптимально выбирать набор родственных и «внешних» пород для сравнения геномов, чтобы обеспечить статистическую значимость тестируемых различий и избежать ложноположительных результатов (например, когда при анализе не учитывается генетическая история пород).

Уникальность якутской породы, с одной стороны, делает ее приоритетной для дальнейших геномных исследований, но с другой – затрудняет выбор наилучших пород для сравнения с ее геномом. Поэтому для исчерпывающего анализа уникальной генетики этой породы могут потребоваться дополнительные, более подробные исследования как российских пород крупного рогатого скота азиатского происхождения, так и зарубежных азиатских тауриновых и индициновых пород. В то же время работы по анализу популяционной структуры и истории происхождения российских пород крупного рогатого скота подготовили почву для изучения изменения их геномов в ответ на селекцию человеком и адаптацию к локальным условиям среды обитания, которые будут рассмотрены в следующем разделе.

Следы селекции в геномах российских пород крупного рогатого скота

Достаточно холодный климат Российской Федерации и небогатая кормовая база позволяют предположить, что геномы местных пород домашнего скота содержат явные следы селекции, сформированные в течение веков адаптации к этим условиям. Как наши, так и данные других групп показали, что наиболее адаптированная к холодному климату якутская порода крупного рогатого скота принадлежит к

так называемой турано-монгольской группе, куда входит ряд пород, таких как бурятская, калмыцкая, вагу, ханву и янбиан (Колесник, 1936; Моисеева и др., 2006; Давыдов, 2012). Существует гипотеза, основанная на генетическом и морфологическом отличии турано-монгольских пород от пород, распространенных в Европе, что они произошли в результате независимой доместикиции популяции тура, по всей вероятности, в Северном Китае, а потом уже распространились по Азиатскому континенту (Zhang et al., 2013). Полученные нами данные подтвердили общность происхождения якутской, бурятской пород, вагу и ханву (Yurchenko et al., 2018). Исходя из этого, был проведен поиск следов отбора в геномах этой группы пород и у каждой из них в отдельности. Действительно, наиболее интересными и перспективными в связи с изучением адаптации к холодному климату оказались результаты, полученные при анализе турано-монгольских пород. Самым статистически значимым был выявленный при сравнении якутской породы с остальными породами турано-монгольской группы (бурятской, калмыцкой, ханву и вагу) район, содержащий ген *FAM134B* (*RETREG1*). Интересно, что у человека этот ген участвует в восприятии нейронами сигналов боли и холода (Islam et al., 2018). Мутации в гене *FAM134B* у человека вызывают заболевание (наследственную сенсорную автономную невропатию 2-го типа), сопровождающееся неспособностью ощущать боль и температуру окружающей среды (Kurth et al., 2009). Тем самым *FAM134B* становится одним из основных кандидатов, которые могут влиять на адаптацию якутского скота к экстремально холодным условиям обитания. Следует отметить, что слабые следы отбора в районе гена *FAM134B* наблюдались и у японской вагу и корейской ханву. Не исключено, что этот район мог находиться под давлением отбора у общего предка этих пород, а впоследствии позволил якутской породе адаптироваться к обитанию в условиях Заполярья.

Конечно, нечувствительность к температуре окружающей среды не может являться основным способом адаптации к холодному климату. Скорее всего, это лишь возможность снизить общий стресс от мощного негативного фактора. Для полной адаптации необходимы изменения в ряде биологических процессов, таких как терморегуляция, иммунный ответ и многие другие. Это предположение было подтверждено в нашей и других работах, где показано, что следы отбора у якутской породы находятся в районах генов, отвечающих за эти процессы. Так, в работе (Iso-Tougu et al., 2016) было показано, что гены, отвечающие за антибактериальную защиту, у якутской породы находятся в районах под позитивным отбором. Мы обнаружили следы селекции у якутской породы в районе рибосомального гена *RPL7*. Экспрессия этого гена значительно отличается у лягушек, хорошо переносящих замораживание, по сравнению с лягушками, которые его не переносят (Wu et al., 2008). Интересно, что слабые следы селекции в районе этого гена наблюдались опять же у ханву и вагу. Гены *TNKS* и *CERKL* также, возможно, связаны с адаптацией якутской породы к холодному климату и обитанию за полярным кругом. Первый из них, кодирующий танкиразу, имеет отношение к энергообмену, питанию и формированию жировой ткани у мышей

(Yeh et al., 2009), а второй кодирует керамид-киназу и экспрессируется в радужной оболочке глаза. Ранее была показана его роль в реакции на световой стресс и защите фоторецепторов (Mandal et al., 2013). Таким образом, этот ген может участвовать в адаптации якутского скота к световому режиму Заполярья.

Гены-кандидаты, относящиеся к адаптации к климатическим условиям, кроме турано-монгольских пород крупного рогатого скота, по-видимому, подвергались отбору и у других российских пород. Так, ген *RAD52*, ответственный за репарацию ДНК и вовлеченный в процессы противовирусной защиты, оказался под давлением отбора у целого ряда российских пород, включая холмогорскую, бестужевскую, калмыцкую и ярославскую. Из зарубежных пород следы селекции в районе этого гена наблюдались только у корейской ханву.

Ранее мы составили список из 416 генов, подверженных позитивному отбору у холодостойких млекопитающих, обитающих в условиях Арктики и Антарктики (Yudin et al., 2017). В этом списке был и ген аквапорина 5 (*AQP5*), который представляет интерес с точки зрения адаптации как к холоду, так и к жаре, поскольку он регулирует испарение воды кожным покровом (Xie et al., 2017). Среди российских пород ген *AQP5* оказался в районе позитивной селекции у холодостойкой холмогорской породы. У бестужевской породы мы наблюдали следы позитивной селекции в районе гена *RGS7*, экспрессия которого в нейронах у человека повышается при хроническом холодовом стрессе (Jedema et al., 2008). У ярославской породы следы селекции наблюдались в районе гена *SFTPD*, участвующего в защите легких от микроорганизмов и подвергающегося селекции у высокогорных популяций человека (Valverde et al., 2015).

Одной из важнейших тканей, участвующих в термогенезе и реакции на холод, является бурая жировая ткань, а ключевым геном этого процесса – термогенин (*UCP1*). Хотя мы не выявили признаков селекции в районе термогенина, ряд генов, регулирующих экспрессию термогенина и вовлеченных в процесс формирования жировой ткани, был обнаружен в районах, подверженных селекции у российских пород крупного рогатого скота. Ген деацетилазы гистонов 3 (*HDAC3*), участвующий в активации энхансеров бурой жировой ткани, находился в районе селекции у якутского скота, а ген *SYK*, который участвует в дифференцировке бурых адипоцитов и регулирует экспрессию *UCP1*, находился в районе под давлением отбора у черно-пестрой и холмогорской пород. Интересно, что ген *ARRDC3*, регулирующий экспрессию *UCP1* в белой жировой ткани, локализован в районе, находящемся под давлением отбора у четырех российских пород крупного рогатого скота, а также у пород ханву, джерси и флитвик. Однако, исходя из влияния этого гена на белую жировую ткань и того факта, что соответствующий район генома находился под давлением отбора и у ряда иностранных пород, *ARRDC3*, скорее всего, имеет отношение к такому экономически важному признаку, как качество мяса, и, соответственно, мог подвергаться искусственному отбору.

Помимо поисков следов селекции в геномах представителей различных пород, перспективным для выявления районов генома, влияющих на конкретные признаки,

является ассоциативный анализ между генотипами и фенотипом. В нашем случае, при адаптации к холодному климату популяций крупного рогатого скота, изначально выведенных в условиях более мягкого климата, можно ожидать, что аллели, отвечающие за адаптацию, не будут фиксированы в популяции, особенно на ранних этапах адаптации. Поэтому дополнительно к общему выявлению следов отбора мы провели исследование по выявлению генов, связанных с поддержанием температуры тела в условиях холодного климата, на популяции герефордов, которых разводят в Сибири на протяжении 60 лет (Igoshin et al., 2019). Кроме того, в результате подтвержденного нами родства пород (см. рис. 1) в работу были включены и особи казахской белоголовой породы, которая создавалась на основании герефордов на территории современного Казахстана в 1930–1950-е гг.

Регистрацию устойчивости к холоду проводили по модифицированной методике (Howard et al., 2014). Для этого за несколько дней до наступления экстремально холодной, по данным метеорологического прогноза, погоды в ушной канал животного вводился датчик температуры, который оставался там не менее двух недель. Датчик передавал информацию о тимпанальной температуре каждые 15 мин. Для полногеномного анализа ассоциаций в качестве фенотипа использовали площадь под кривой температуры тела в течение самых холодных пяти суток из двухнедельного периода. Генотипирование выполняли на ДНК-чипе GeneSeek Genomic Profiler High-Density (GGP HD150K), содержащем около 139000 генетических маркеров. Результаты ассоциативного анализа указали на единственный район хромосомы 15, который совпал для всех типов проведенного ассоциативного анализа. В этом районе находились два гена: *GRIA4* и *MSANTD4*. Этот же район был выявлен и в ходе независимого поиска следов отбора в нашей выборке как один из районов, подверженных селекции в сибирской выборке герефордов и казахской белоголовой породы. Наиболее вероятным кандидатом, связанным с поддержанием температуры тела на холоде, был ген *GRIA4*, поскольку глутаминовый рецептор, который кодируется этим геном, влияет на терморегуляцию у млекопитающих (крыс) (Sengupta et al., 2016).

Исследований по генетической адаптации к холодному климату, выполненных на зарубежных породах скота, крайне мало. Так, по описанному выше протоколу (регистрация площади под кривой температуры тела в течение самых холодных пяти суток за период наблюдения), J. Howard с соавторами (2014) провели полногеномный анализ ассоциаций способности поддерживать температуру тела на холоде у гибридов ангусской, симментальской и пьемонтской пород. Наибольший сигнал был выявлен в районах генов, белки которых контролируют энергетический метаболизм (*COX7C*), пентозофосфатный путь окисления глюкозы (*FBP1* и *FBP2*), васкулогенез (*RASA1*), ионный транспорт (*CACNG3* и *PRKCB*), работу рецептора тиреоидного гормона (*TRIP11*) и тепловой шок (*HSBP1*). С помощью полногеномного секвенирования были идентифицированы потенциальные гены холодовой адаптации у представителей западного (*CD14*, *COBL*, *JMJD1C*, *KCNMA1*, *PLA2G4*, *SERPINF2*, *SRA1*, *TAF7*) и восточного (*DNAJC28*, *HSP90B1*, *AGTRAP*, *TAF7*, *TRIP13*, *NPPA*,

NPPB) финского скота, традиционно разводимого в Финляндии (Weldenegodguad et al., 2019). Ранее было показано, что некоторые из этих генов, вовлеченные в энергетический метаболизм, сокращение гладкой мускулатуры, температурную рецепцию и другие физиологические реакции, могли участвовать в адаптации к холоду у коренных народностей Сибири (Cardona et al., 2014). Анализ транскриптома крови показал, что профиль экспрессии генов у северного финского скота более близок к профилю якутского, чем голштинского скота (Pokharel et al., 2019).

Доместикация и морфологические признаки

Как и ожидалось из-за выявления схожих следов селекции в других популяциях крупного рогатого скота, значительное количество следов отбора у российских пород наблюдалось в районе генов, отвечающих за окраску тела. Этот признак, скорее всего, подвергался изменению при domestикации 8–10 тыс. лет назад. Основными генами, вовлеченными в процесс изменения окраски, являются гены *KIT* (белая окраска) и *KITLG* (чалая окраска). У двух групп российского скота европейского происхождения район генома, содержащий *KITLG*, был под давлением отбора. Однако у группы пород, в которую входят казахская белоголовая, костромская, джерсейская и флитвик, этот район был значительно короче (3.3 Мб), чем у группы, куда входят бестужевская, черно-пестрая, голштинская, холмогорская, ярославская породы (13.8 Мб). При анализе следов селекции у индивидуальных пород район гена *KITLG* был идентифицирован у бестужевской и казахской белоголовой.

В случае гена *KIT* сигнал выявлялся в группе пород: казахская белоголовая, костромская, джерсейская, флитвик. При анализе следов селекции в индивидуальных породах, множественные следы селекции в районе гена *KIT* наблюдались у ярославской, казахской белоголовой и флитвик. Это дает возможность предположить, что на самом деле в районах генома, содержащих *KIT* и *KITLG*, могут находиться множественные участки ДНК, которые подвергаются селекции. Кроме генов *KIT* и *KITLG*, в районах селекции нами выявлены и другие гены, отвечающие за окраску. Например, ген *EDN3* был найден у казахской белоголовой породы. Этот ген, наряду с геном *SLC41A2*, может быть связан с фенотипом «белая голова» у российских пород крупного рогатого скота (Юдин и др., 2018). Ген *COPA*, участвующий в синтезе пигментов, находился в районе генома под давлением отбора у черно-пестрой, бестужевской породы и у голштинов. Известно, что мутации в гене *COPA* отвечают за красный окрас шерсти у голштинов (Dorshorst et al., 2015).

Заключение

Недавние работы по изучению российских пород крупного рогатого скота позволили показать их близость как к европейским, так и к азиатским тауриновым породам. Вследствие этого интересны связи якутской породы с коммерческими породами из Японии и Кореи. Некоторые российские породы до сих пор обладают уникальной генетикой и нуждаются в защите, тогда как другие подверглись значительным изменениям в результате недавнего прилития крови голштинов. Породы с уникальной генетикой

служат источником генетической информации об изменениях, происходящих при адаптации к суровым климатическим условиям. Эта информация может быть использована для создания высокопродуктивных пород, лучше адаптированных к климатическим условиям Евразии.

Список литературы / References

- Давыдов В.Н. Эколого-генетические основы распределения популяций *Bos taurus* в Байкальской Сибири. Вестн. Бурят. гос. ун-та. 2012;4:128-129.
- [Davydov V.N. Ecological and genetic bases of the *Bos taurus* population distribution in the Baikal region. Vestnik Buryatskogo Gosudarstvennogo Universiteta = Bulletin of the Buryat State University. 2012;4:128-129. (in Russian)]
- Дмитриев Н.Г. Породы скота по странам мира. Справ. книга. Л.: Колос, 1978.
- [Dmitriev N.G. Breed Cattle by Countries of the World. Leningrad: Kolos Publ., 1978.]
- Дунин И.М., Данкверт А.Г. (ред.) Справочник пород и типов сельскохозяйственных животных, разводимых в Российской Федерации. М.: ВНИИплем, 2013.
- [Dunin I.M., Dankvert A.G. (Eds.) Breeds and Types of Farm Animals in the Russian Federation. Moscow: All-Russia Research Institute of Animal Breeding, 2013. (in Russian)]
- Зиновьева Н.А., Доцев А.В., Сермягин А.А., Виммерс К., Рейер Х., Солкнер Й., Денискова Т.Е., Брем Г. Изучение генетического разнообразия и популяционной структуры российских пород крупного рогатого скота с использованием полногеномного анализа SNP. С.-х. биология. 2016;51(6):788-800. DOI 10.15389/agrobiology.2016.6.788rus.
- [Zinovieva N.A., Dotsev A.V., Sermyagin A.A., Wimmers K., Reyer H., Sölkner J., Deniskova T.E., Brem G. Study of genetic diversity and population structure of five Russian cattle breeds using whole-genome SNP analysis. Selskokhozyaistvennaya Biologiya = Agricultural Biology. 2016;51(6):788-800. DOI 10.15389/agrobiology.2016.6.788eng.]
- Колесник Н.Н. Происхождение домашних видов животных турано-монгольской группы. М., 1936.
- [Kolesnik N.N. The Origin of Domesticated Species of the Turan-Mongolian Group. Moscow, 1936. (in Russian)]
- Кушнир А.В., Глазко В.И. Серый украинский скот и его близкородственные формы. Сиб. экол. журн. 2009;16(3):495-506.
- [Kushnir A.V., Glazko V.I. Gray Ukrainian cattle and their closely related forms. Contemporary Problems of Ecology. 2009;3(3):288-295. DOI 10.1134/S1995425509030205.]
- Ларкин Д.М., Юдин Н.С. Значимость геномных исследований в истории формирования домашних животных. Молекуляр. генетика, микробиология и вирусология. 2016;4:123-128.
- [Larkin D.M., Yudin N.S. The genomes and history of domestic animals. Molekulyarnaya Genetika, Mikrobiologiya i Virusologiya = Molecular Genetics, Microbiology and Virology. 2016;31(4):197-202. DOI 10.3103/S0891416816040054.]
- Моисеева И.Г., Уханов С.В., Столповский Ю.А., Сулимова Г.Е., Каштанов С.Н. Генофонды сельскохозяйственных животных. Генетические ресурсы животноводства России. М.: Наука, 2006.
- [Moiseeva I.G., Ukanov S.V., Stolpovsky Yu.A., Sulimova G.E., Kashtanov S.N. Gene Pools of Farm Animals. Moscow: Nauka Publ., 2006. (in Russian)]
- Юдин Н.С., Белоногова Н.М., Ларкин Д.М. Выявление генов, вовлеченных в контроль белой окраски головы, с использованием полногеномного анализа ассоциаций. Вавиловский журнал генетики и селекции. 2018;22(2):217-223. DOI 10.18699/VJ18.350.
- [Yudin N.S., Belonogova N.M., Larkin D.M. Genes related to the white face colour pattern in eight Russian cattle breeds. Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding. 2018;22(2):217-223. DOI 10.18699/VJ18.350 (in Russian)]
- Beynon S.E., Slavov G.T., Farré M., Sunduimijid B., Waddams K., Davies B., Haresign W., Kijas J., MacLeod I.M., Newbold C.J., Davies L., Larkin D.M. Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. BMC Genet. 2015;16:65. DOI 10.1186/s12863-015-0216-x.
- Boichard D., Brochard M. New phenotypes for new breeding goals in dairy cattle. Animal. 2012;6(4):544-550. DOI 10.1017/S175173112000018.
- Bovine HapMap Consortium, Gibbs R.A., Taylor J.F., Van Tassell C.P., Barendse W., Eversole K.A., Gill C.A., Green R.D., Hamernik D.L., Kappes S.M., Lien S., Matukumalli L.K., McEwan J.C., Nazareth L.V., Schnabel R.D., Weinstock G.M., Wheeler D.A., Ajmone-Marsan P., Boettcher P.J., Caetano A.R., Garcia J.F., Hanotte O., Mariani P., Skow L.C., Sonstegard T.S., Williams J.L., Diallo B., Hailemariam L., Martinez M.L., Morris C.A., Silva L.O., Spelman R.J., Mulatu W., Zhao K., Abbey C.A., Agaba M., Araujo F.R., Bunch R.J., Burton J., Gorni C., Olivier H., Harrison B.E., Luff B., Machado M.A., Mwakaya J., Plastow G., Sim W., Smith T., Thomas M.B., Valentini A., Williams P., Womack J., Woolliams J.A., Liu Y., Qin X., Worley K.C., Gao C., Jiang H., Moore S.S., Ren Y., Song X.Z., Bustamante C.D., Hernandez R.D., Muzny D.M., Patil S., San Lucas A., Fu Q., Kent M.P., Vega R., Matukumalli A., McWilliam S., Sclap G., Bryc K., Choi J., Gao H., Grefenstette J.J., Murdoch B., Stella A., Villa-Angulo R., Wright M., Aerts J., Jann O., Negrini R., Goddard M.E., Hayes B.J., Bradley D.G., Barbosa da Silva M., Lau L.P., Liu G.E., Lynn D.J., Panzitta F., Dodds K.G. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science. 2009;324(5926):528-532. DOI 10.1126/science.1167936.
- Buchanan D.S., Lenstra J.A. Breeds of Cattle. In: Garrick D.J., Ruvinsky A. (Eds.) The Genetics of Cattle. 2nd edn. CAB International, 2015;33-66.
- Cardona A., Pagani L., Antao T., Lawson D.J., Eichstaedt C.A., Yngvadottir B., Shwe M.T., Wee J., Romero I.G., Raj S., Metspalu M., Vilems R., Willerslev E., Tyler-Smith C., Malyarchuk B.A., Derenko M.V., Kivisild T. Genome-wide analysis of cold adaptation in indigenous Siberian populations. PLoS One. 2014;9:e98076. DOI 10.1371/journal.pone.0098076.
- Decker J.E., McKay S.D., Rolf M.M., Kim J., Molina Alcalá A., Sonstegard T.S., Hanotte O., Götherström A., Seabury C.M., Prahara L., Babar M.E., Correia de Almeida Regitano L., Yildiz M.A., Heaton M.P., Liu W.S., Lei C.Z., Reecy J.M., Saif-Ur-Rehman M., Schnabel R.D., Taylor J.F. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. PLoS Genet. 2014;10(3):e1004254. DOI 10.1371/journal.pgen.1004254.
- Dmitriev N.G., Ernst L.K. Animal genetics resources of the USSR. Food and Agriculture Organization of the United Nations, Rome, Italy. 1989. Available at: <http://www.fao.org/docrep/009/ah759e/ah759e00.htm>
- Dorshorst B., Henegar C., Liao X., Sällman Almén M., Rubin C.J., Ito S., Wakamatsu K., Stothard P., Van Doormaal B., Plastow G., Barsh G.S., Andersson L. Dominant red coat color in Holstein cattle is associated with a missense mutation in the Coatmer Protein Complex, Subunit Alpha (COPA) gene. PLoS One. 2015;10(6):e0128969. DOI 10.1371/journal.pone.0128969.
- Gao Y., Wu H., Wang Y., Liu X., Chen L., Li Q., Cui C., Liu X., Zhang J., Zhang Y. Single Cas9 nickase induced generation of *NRAMP1* knockin cattle with reduced off-target effects. Genome Biol. 2017;18(1):13. DOI 10.1186/s13059-016-1144-4.
- Gaouar S.B., Da Silva A., Ciani E., Kdidi S., Aouissat M., Dhimi L., Lafri M., Maftah A., Mehtar N. Admixture and local breed marginalization threaten Algerian sheep diversity. PLoS One. 2015;10(4):e0122667. DOI 10.1371/journal.pone.0122667.
- Howard J.T., Kachman S.D., Snelling W.M., Pollak E.J., Ciobanu D.C., Kuehn L.A., Spangler M.L. Beef cattle body temperature during climatic stress: a genome-wide association study. Int. J. Biometeorol. 2014;58(7):1665-1672. DOI 10.1007/s00484-013-0773-5.
- Igoshin A.V., Yurchenko A.A., Belonogova N.M., Petrovsky D.V., Aitnazarov R.B., Soloshenko V.A., Yudin N.S., Larkin D.M. Ge-

- nome-wide association study and scan for signatures of selection point to candidate genes for body temperature maintenance under the cold stress in Siberian cattle populations. *BMC Genetics*. 2019; 20(Suppl.1):26.
- Islam F., Gopalan V., Lam A.K. *RETREG1 (FAM134B)*: a new player in human diseases: 15 years after the discovery in cancer. *J. Cell. Physiol.* 2018;233(6):4479-4489. DOI 10.1002/jcp.26384.
- Iso-Touru T., Tapio M., Vilkki J., Kiseleva T., Ammosov I., Ivanova Z., Popov R., Ozerov M., Kantanen J. Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe. *Anim. Genet.* 2016;47(6):647-657. DOI 10.1111/age.12473.
- Jedema H.P., Gold S.J., Gonzalez-Burgos G., Sved A.F., Tobe B.J., Wensel T., Grace A.A. Chronic cold exposure increases RGS7 expression and decreases α_2 -autoreceptor-mediated inhibition of noradrenergic locus coeruleus neurons. *Eur. J. Neurosci.* 2008;27(9):2433-2443.
- Kantanen J., Edwards C.J., Bradley D.G., Viinalass H., Thessler S., Ivanova Z., Kiselyova T., Cinkulov M., Popov R., Stojanović S., Ammosov I., Vilkki J. Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*). *Heredity (Edinb.)*. 2009;103(5):404-415. DOI 10.1038/hdy.2009.68.
- Kassahun Y., Mattiangeli V., Ameni G., Hailu E., Aseffa A., Young D.B., Hewinson R.G., Vordermeier H.M., Bradley D.G. Admixture mapping of tuberculosis and pigmentation-related traits in an African-European hybrid cattle population. *Front. Genet.* 2015;6:210. DOI 10.3389/fgene.2015.00210.
- Kurth I., Pamminger T., Hennings J.C., Soehendra D., Huebner A.K., Rothier A., Baets J., Senderek J., Topaloglu H., Farrell S.A., Nürnberg G., Nürnberg P., De Jonghe P., Gal A., Kaether C., Timmerman V., Hübner C.A. Mutations in *FAM134B*, encoding a newly identified Golgi protein, cause severe sensory and autonomic neuropathy. *Nat. Genet.* 2009;41(11):1179-1181. DOI 10.1038/ng.464.
- Li M.H., Kantanen J. Genetic structure of Eurasian cattle (*Bos taurus*) based on microsatellites: clarification for their breed classification. *Anim. Genet.* 2010;41(2):150-158. DOI 10.1111/j.1365-2052.2009.01980.x.
- Loftus R.T., MacHugh D.E., Bradley D.G., Sharp P.M., Cunningham P. Evidence for two independent domestications of cattle. *Proc. Natl. Acad. Sci. USA*. 1994;91(7):2757-2761.
- Mandal N.A., Tran J.T., Saadi A., Rahman A.K., Huynh T.P., Klein W.H., Cho J.H. Expression and localization of CERKL in the mammalian retina, its response to light-stress, and relationship with *NeuroD1* gene. *Exp. Eye Res.* 2013;106:24-33. DOI 10.1016/j.exer.2012.10.014.
- Mannen H., Kohno M., Nagata Y., Tsuji S., Bradley D.G., Yeo J.S., Nyamsamba D., Zagdsuren Y., Yokohama M., Nomura K., Amano T. Independent mitochondrial origin and historical genetic differentiation in North Eastern Asian cattle. *Mol. Phylogenet. Evol.* 2004; 32(2):539-544.
- Matukumalli L.K., Lawley C.T., Schnabel R.D., Taylor J.F., Allan M.F., Heaton M.P., O'Connell J., Moore S.S., Smith T.P., Sonstegard T.S., Van Tassell C.P. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One*. 2009;4(4):e5350. DOI 10.1371/journal.pone.0005350.
- McKay S.D., Schnabel R.D., Murdoch B.M., Matukumalli L.K., Aerts J., Coppieters W., Crews D., Dias Neto E., Gill C.A., Gao C., Mannen H., Wang Z., Van Tassell C.P., Williams J.L., Taylor J.F., Moore S.S. An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Genet.* 2008;9:37. DOI 10.1186/1471-2156-9-37.
- Mucha S., Bunger L., Conington J. Genome-wide association study of footrot in Texel sheep. *Genet. Sel. Evol.* 2015;47:35. DOI 10.1186/s12711-015-0119-3.
- Pokharel K., Weldenegodguad M., Popov R., Honkatukia M., Huu-ki H., Lindeberg H., Peippo J., Reilas T., Zarovnyaev S., Kantanen J. Whole blood transcriptome analysis reveals footprints of cattle adaptation to sub-arctic conditions. *Anim. Genet.* 2019;50(3):217-227. DOI 10.1111/age.12783.
- Sengupta T., Jaryal A.K., Mallick H.N. Effects of NMDA and non-NMDA ionotropic glutamate receptors in the medial preoptic area on body temperature in awake rats. *J. Therm. Biol.* 2016;61:1-7. DOI 10.1016/j.jtherbio.2016.07.020.
- Sermiyagin A.A., Dotsev A.V., Gladyr E.A., Traspov A.A., Denis-kova T.E., Kostyunina O.V., Reyher H., Wimmers K., Barbato M., Paronyan I.A., Plemiyashov K.V., Sölkner J., Popov R.G., Brem G., Zinovieva N.A. Whole-genome SNP analysis elucidates the genetic structure of Russian cattle and its relationship with Eurasian taurine breeds. *Genet. Sel. Evol.* 2018;50(1):37. DOI 10.1186/s12711-018-0408-8.
- Upadhyay M.R., Chen W., Lenstra J.A., Goderie C.R., MacHugh D.E., Park S.D., Magee D.A., Matassino D., Ciani F., Megens H.J., van Arendonk J.A., Groenen M.A.; European Cattle Genetic Diversity Consortium; RPMA Crooijmans. Genetic origin, admixture and population history of aurochs (*Bos primigenius*) and primitive European cattle. *Heredity*. 2017;118(2):169-176. DOI 10.1038/hdy.2016.79.
- Valverde G., Zhou H., Lippold S., de Filippo C., Tang K., López Herráez D., Li J., Stoneking M. A novel candidate region for genetic adaptation to high altitude in Andean populations. *PLoS One*. 2015; 10(5):e0125444. DOI 10.1371/journal.pone.0125444.
- van Binsbergen R., Calus M.P., Bink M.C., van Eeuwijk F.A., Schrooten C., Veerkamp R.F. Genomic prediction using imputed whole-genome sequence data in Holstein Friesian cattle. *Genet. Sel. Evol.* 2015;47:71. DOI 10.1186/s12711-015-0149-x.
- Weldenegodguad M., Popov R., Pokharel K., Ammosov I., Ming Y., Ivanova Z., Kantanen J. Whole-genome sequencing of three native cattle breeds originating from the Northernmost cattle farming regions. *Front. Genet.* 2019;9:728. DOI 10.3389/fgene.2018.00728.
- Wu S., De Croos J.N., Storey K.B. Cold acclimation-induced up-regulation of the ribosomal protein L7 gene in the freeze tolerant wood frog, *Rana sylvatica*. *Gene*. 2008;424(1-2):48-55. DOI 10.1016/j.gene.2008.07.023.
- Xie L., Jin L., Feng J., Lv J. The expression of AQP5 and UTs in the sweat glands of uremic patients. *Biomed. Res. Int.* 2017;2017:8629783. DOI 10.1155/2017/8629783.
- Yeh T.Y., Beiswenger K.K., Li P., Bolin K.E., Lee R.M., Tsao T.S., Murphy A.N., Hevener A.L., Chi N.W. Hypermetabolism, hyperphagia, and reduced adiposity in tankyrase-deficient mice. *Diabetes*. 2009;58(11):2476-2485. DOI 10.2337/db08-1781.
- Yudin N.S., Larkin D.M., Ignatieva E.V. A compendium and functional characterization of mammalian genes involved in adaptation to Arctic or Antarctic environments. *BMC Genet.* 2017;18(Suppl.1):111. DOI 10.1186/s12863-017-0580-9.
- Yurchenko A., Yudin N., Aitnazarov R., Plyusnina A., Brukhin V., Soloshenko V., Lhasaranov B., Popov R., Paronyan I.A., Plemiyashov K.V., Larkin D.M. Genome-wide genotyping uncovers genetic profiles and history of the Russian cattle breeds. *Heredity (Edinb.)*. 2018;120(2):125-137. DOI 10.1038/s41437-017-0024-3.
- Zhang H., Pajmans J.L., Chang F., Wu X., Chen G., Lei C., Yang X., Wei Z., Bradley D.G., Orlando L., O'Connor T., Hofreiter M. Morphological and genetic evidence for early Holocene cattle management in northeastern China. *Nat. Commun.* 2013;4:2755. DOI 10.1038/ncomms3755.

ORCID ID

N.S. Yudin orcid.org/0000-0002-1947-5554
D.M. Larkin orcid.org/0000-0001-7859-6201

Acknowledgements. This work was supported by the Russian Science Foundation, project 16-14-00090.

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

Received November 22, 2018. Revised May 22, 2019. Accepted May 31, 2019.

Assessment of line differentiation in the Thoroughbred horse breed using DNA microsatellite loci

L.A. Khrabrova , N.V. Blohina, O.I. Suleymanov, G.A. Rozhdestvenskaya, V.F. Pustovoy

All-Russian Research Institute for Horse Breeding, Divovo, Ryazan Region, Russia

 e-mail: l.khrabrova@yandex.ru

The Thoroughbred (TB) horse is the best racehorse breed used in the racing industry. This breed has had a closed studbook for about 300 years. In Russia TB horses have been bred since the second half of the XVIII century. The modern Russian Stud Book register of TB horses is partially presented by stallions and broodmares imported from different countries. The genealogical structure of the breed is represented by 17 lines, among which the Northern Dancer line dominates (30.9 %). The study of features of different lines of TB was carried out on 17 loci of DNA microsatellites (*VHL20*, *HTG4*, *AHT4*, *HMS7*, *HTG6*, *AHT5*, *HMS6*, *ASB23*, *ASB2*, *HTG10*, *HTG7*, *HMS3*, *HMS2*, *ASB17*, *LEX3*, *HMS1* and *CA425*) to assess genetic differentiation of the genealogical structure. The results of the DNA typing of 8091 Thoroughbred horses across microsatellite loci show that the gene pool of the domestic population is represented by 100 alleles typical of the breed. A comparative analysis of the genotypes of horses representing different lines indicates that they differ in the number of alleles (85–99), allele frequencies, the level of polymorphism A_e (2.93–3.48) and the degree of the observed heterozygosity H_o (0.653–0.739). The genetic distances between the lines varied in a wide range from 0.014 (Nasrullah – Northern Dancer) to 0.125 (Massine – Teddy). The correspondence to HWE was maintained in most lines, which is confirmed by the negative values of F_{is} . Cluster analysis demonstrated the correspondence of the obtained dendrogram of Nei's genetic distances to its genealogical scheme of lines. Genetic differentiation of lines by index F_{st} varied in a range 0.005–0.073 at the mean value $F_{st} = 0.024$. The data indicate genetic differentiation of lines of TB at the STR markers and confirm the effectiveness of the system of linear breeding for the maintenance of interbreed biodiversity.

Key words: genetic diversity; horse; linear structure; microsatellite DNA; population analysis; Thoroughbred.

For citation: Khrabrova L.A., Blohina N.V., Suleymanov O.I., Rozhdestvenskaya G.A., Pustovoy V.F. Assessment of line differentiation in the Thoroughbred horse breed using DNA microsatellite loci. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5):569-574. DOI 10.18699/VJ19.526

Оценка дифференциации линий в чистокровной верховой породе лошадей с использованием микросателлитов ДНК

Л.А. Храброва , Н.В. Блохина, О.И. Сулейманов, Г.А. Рождественская, В.Ф. Пустовой

Всероссийский научно-исследовательский институт коневодства, Дивово, Рязанская область, Россия

 e-mail: l.khrabrova@yandex.ru

Чистокровная верховая лошадь является лучшей породой для использования в скаковой индустрии. Почти около 300 лет эта порода имеет закрытую племенную книгу. В России чистокровных верховых лошадей разводят со второй половины XVIII в. Современный российский племенной регистр чистокровных верховых лошадей частично представлен жеребцами и племенными кобылами, импортированными из разных стран. Генеалогическая структура породы включает 17 линий, среди которых явно доминирует линия Northern Dancer (30.9 %). С целью оценки генетической дифференциации генеалогической структуры чистокровной верховой породы было проведено изучение особенностей разных линий с использованием 17 локусов микросателлитов ДНК (*VHL20*, *HTG4*, *AHT4*, *HMS7*, *HTG6*, *AHT5*, *HMS6*, *ASB23*, *ASB2*, *HTG10*, *HTG7*, *HMS3*, *HMS2*, *ASB17*, *LEX3*, *HMS1* и *CA425*). Результаты генотипирования 8 091 чистокровной верховой лошади по панельным микросателлитным локусам свидетельствуют, что аллелофонд отечественной популяции представлен 100 аллелями, типичными для этой породы. Сравнительный анализ лошадей разных линий показал, что они различаются по числу аллелей (85–99), частотам встречаемости аллелей, уровню полиморфности A_e (2.93–3.48) и степени фактической гетерозиготности H_o (0.653–0.739). Генетические дистанции между линиями варьировали в широком диапазоне: от 0.014 (Nasrullah – Northern Dancer) до 0.125 (Massine – Teddy). Соответствие распределению HWE наблюдали во всех линиях, что подтверждают отрицательные значения F_{is} . Кластерный анализ продемонстрировал соответствие полученной дендрограммы генетических дистанций по Nei генеалогической схеме линий. Генетическая дифференциация линий по индексу F_{st} варьировала в интервале 0.005–0.073 при среднем значении $F_{st} = 0.024$.

Полученные данные свидетельствуют о генетической дифференциации линий чистокровной верховой породы по STR-маркерам и подтверждают эффективность системы линейного разведения для поддержания внутрипородного биоразнообразия.

Ключевые слова: генетическое разнообразие; лошадь; линейная структура; микросателлит ДНК; популяционный анализ; чистокровная верховая порода.

Introduction

Thoroughbred (TB) is the fastest of the world's most valuable breeds, which is currently bred in 70 countries. The breed was created in England in the XVII–XVIII centuries as a result of crossing imported Arabian and other oriental sires with native mares. The General Stud Book of TB horses was published in 1791 and since that time only the method of purebred breeding has been practiced. Thoroughbred horses were introduced to Russia in the second half of the XVIII century. The first volume of the Russian Thoroughbred Stud Book was issued in 1836 and included 287 stallions and 366 mares. After World War I the TB population was perfected in comparative isolation with a minor exchange of genetic material. At the end of the XX century the country began to import stallions and mares from Europe and the USA for racing and breeding. Currently Russia ranks the fifth in Europe for the number of Thoroughbred mares (Suleymanov, 2016).

The breed has three foundation stallions, Byerley Turk, Darley Arabian and Godolphin Arabian. The most successful sire was Darley Arabian, the line developed through Eclipse and its descendants St. Simon, Irish Birdcatcher and Touchstone. In the mid-twentieth century the lines going back to Phalaris (1913) began to dominate in the breed and this trend has continued to the present time (Vitt, 1957; Konovalova, Klebosolova, 2016).

According to the requirements of the International Stud Book Committee (ISBC) and the International Society for Animal Genetics (ISAG), genetic laboratories must test TB horses using microsatellite loci of DNA as of 2001. The microsatellites are mainly localized in non-coding regions of the genome and are characterized by the codominant type of inheritance and a high polymorphic level. These universal markers are used for parentage verification, in genetic monitoring and population analyses (Wright, Bentzen, 1994; Bowling, Ruvinsky, 2000; Van de Goor et al., 2011; Khrabrova, Blohina, 2018).

Numerous studies indicate a high genetic similarity of TB horse populations in different countries all over the world by STR loci tested (Cunningham et al., 2001; Ling et al., 2011; Blohina, Khrabrova, 2012; Rukavina et al., 2016; Putnova et al., 2018), while Great Britain and the USA keep the leading positions among countries that breed TB horses. In a previous study (Khrabrova, 2009), genetic features of sires of different lines were studied using 13 DNA STR loci.

The aim of our work was to assess the interline differentiation of the modern structure of the Russian population of the breed using 17 microsatellite markers recommended by the ISAG for parentage testing.

Materials and methods

The object of research was the database of DNA typing of 8091 TB horses registered in the Russian Thoroughbred Stud

Book for the period of 2001–2017, which includes 97–99 % of horses raised in Russia. The reference population included 957 sires and 3194 broodmares and their offspring. Over the period as indicated, 2529 TB horses were imported, most of which (about 65 %) were used for breeding.

Genomic DNA were prepared from hair follicles or blood samples of TB horses by a standard procedure using Extra Gene™ DNA Prep 200 and Diatom™ DNA Prep 200 kits (Laboratory Isogene, Russia) in the Laboratory of Genetics of the All-Russian Research Institute for Horse Breeding, certified by the ISAG based on the Horse Comparison Test (HCT) results.

PCR products of 17 panel microsatellite markers (*AHT4*, *AHT5*, *ASB2*, *ASB17*, *ASB23*, *CA425*, *HMS1*, *HMS2*, *HMS3*, *HMS6*, *HMS7*, *HTG4*, *HTG6*, *HTG7*, *HTG10*, *LEX3* and *VHL20*) were amplified with fluorescently labeled primers (StockMarks®, Applied Biosystems) and then analyzed on the automated DNA sequencer ABI 3130. The size of the STR marker was determined using Gene Mapper version 4.0 (Applied Biosystems). Detailed information on microsatellite markers used for parentage control of horses is presented in a work by van de Goor and van Haeringen (2010).

For information on the pedigree and line of horses, the base of the storage-and-retrieval system for managing the selection process in Russian horse breeding “ISS Kony 3” was used (www.ruhorses.ru).

Genetic diversity within the line was evaluated by the basic parameters including total number of allele variants (N_a), effective number of allele (A_e), number of allele per locus (MNA), observed (H_o) and expected heterozygosity (H_e) and Hardy–Weinberg equilibrium (HWE). The coefficients of genetic similarity and genetic distances between the lines were determined by Nei's methods (Nei, 1975) using software Statistica 12 ver.10 (www.StatSoftStatistica.ru). Inbreeding and genetic differentiations among lines were estimated by methods of F-Statistics (Weir, 1996) using program FSTAT 1.2 (www2.unil.ch/popgen/fstat.htm).

Results

The preliminary analysis revealed that at present in the Russian population of Thoroughbred horses the Northern Dancer line dominates, including 2.5 thousand horses (30.9 %). The second place in number is occupied by representatives of the Nasrullah line ($n = 1353$; 16.7 %), the third place, by horses of the young popular line of Mr. Prospector ($n = 1040$; 12.9 %). In general, the genealogical structure of the national population of this breed, as elsewhere in the world, is dominated by the representatives of the lines going back to Phalaris (76.2 %).

The results of DNA typing of 8091 Thoroughbred horses by 17 panel microsatellite loci show that allele's pool of the studied group of horses is represented by 100 alleles. The number of alleles at the loci varied from 4 (*HMS1*) to 9 (*ASB2*),

Table 1. Characteristics of the Thoroughbred lines by the genetic population parameters

Line	N	N_a	A_e	H_o	H_e	F_{is}	F_{st}	MNA
A.P. Indy	196	85	3.141	0.677	0.653	-0.037	0.041	5.000
Blandford	204	89	3.348	0.684	0.676	-0.012	0.005	5.235
Dark Roland	138	87	3.300	0.666	0.670	0.009	0.018	5.118
Douglas	380	92	3.319	0.688	0.671	-0.025	0.014	5.412
Fair Trial	106	85	3.198	0.672	0.660	-0.023	0.029	5.000
Hyperion	49	85	3.341	0.724	0.666	-0.085	0.026	5.000
Man O'War	316	89	3.383	0.684	0.666	-0.030	0.028	5.235
Massine	105	87	2.934	0.653	0.631	-0.028	0.073	5.118
Mr. Prospector	1040	95	3.390	0.674	0.671	-0.004	0.018	5.588
Nasrullah	1353	97	3.355	0.680	0.673	-0.011	0.013	5.706
Native Dancer	292	91	3.280	0.692	0.657	-0.055	0.037	5.353
Nearco	678	94	3.360	0.683	0.674	-0.016	0.012	5.529
Northern Dancer	2501	99	3.423	0.680	0.678	-0.003	0.006	5.824
Prince Rose	280	92	3.373	0.682	0.666	-0.022	0.025	5.412
Ribot	173	87	3.199	0.669	0.657	-0.021	0.034	5.118
Teddy	33	85	3.478	0.739	0.687	-0.080	0.015	5.000
Tourbilon	129	86	3.103	0.678	0.650	-0.043	0.045	5.059
Others	118	95	3.461	0.661	0.670	0.020	0.020	5.588
Mean	8091	100	3.488	0.681	0.682	0.001	0.024	5.882

Note: N , number of horses; N_a , number of alleles; A_e , effective number of alleles; H_e , expected heterozygosity; H_o , observed heterozygosity; F_{is} , population inbreeding level; F_{st} , index differentiation; MNA , average number of alleles per locus.

the most of loci show high level of polymorphism. Genetic structure of the Thoroughbred breed is characterized by a high frequency of alleles *AHT4O* (0.410), *AHT5K* (0.418), *ASB17G* (0.332), *ASB23J* (0.323), *HMS1J* (0.445), *HMS2L* (0.686), *HMS3I* (0.571), *HMS6P* (0.560), *HTG4K* (0.512), *HTG6J* (0.482), *HTG10I* (0.367) and *VHL20M* (0.300). The number of alleles on average amounted to 5.88 per locus, the A_e level ranged from 1.71 (*HMS2*) to 5.70 (*ASB2*) (mean 3.49). The observed heterozygosity (H_o) was 0.681 and varied from 0.418 (*HMS2*) to 0.854 (*ASB2*).

Comparative analysis of representatives of different lines on the basis of genetic and population parameters shows (Table 1) that they differ in number of alleles (N_a), level of polymorphism (A_e) and degree of observed heterozygosity (H_o). In the main lines the number of alleles varied in the range of 85–99; polymorphism level (A_e), 2.934–3.478; the degree of observed heterozygosity (H_o), 0.653–0.739. The leading line of Northern Dancer is characterized by the maximum spectrum of alleles ($n = 99$), sufficiently high levels of polymorphism ($A_e = 3.423$) and the degree of heterozygosity ($H_o = 0.680$).

Horses from small lines of Blandford, Hyperion and Teddy also show high levels of genetic diversity. The domestic Douglas line showed an average level of genetic parameters, but significantly differed in the high frequency of alleles *VHL20I*, *AHT5J*, *ASB2K*, *HMS3O* and *CA425I* ($p > 0.999$).

Horses of the Tourbilon line (1928) going back to Herod (1758) are characterized by a low level of genetic parameters, but stand out from the general background by significantly

higher frequencies of alleles *VHL20M*, *AHT4H*, *HMS7L* and *HMS1M* ($p > 0.999$).

Another genealogical TB branch going from Matchem (1748) to Man O'War (1917) was represented by 316 horses in our study. Representatives of this line differed markedly from other lines by a high concentration of alleles *VHL20L*, *HTG4M*, *AHT4H*, *HTG6G*, *AHT5J*, *ASB2B*, *HTG10O*, *HMS3P*, *HMS2L*, *ASB17R* and *CA425N*. Significant differences in the frequency of occurrence of individual alleles of 17 STR loci were noted for all analyzed lines, while only the leading abundant Northern Dancer line had an allele frequency structure similar to the population one.

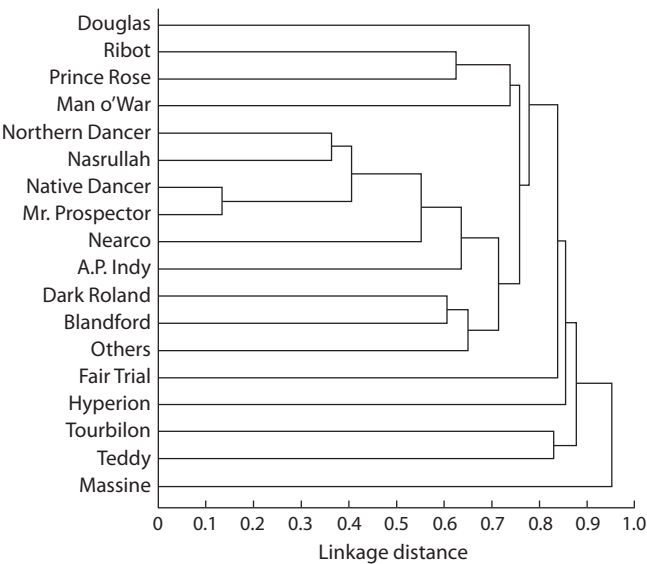
For 16 lines studied, the observed heterozygosity H_o exceeded H_e and a negative F_{is} value was recorded, indicating an excess of heterozygous genotypes. Only in the Dark Roland line and a group of others lines the Hardy–Weinberg equilibrium was disturbed ($F_{is} = 0.009–0.020$).

Analysis of genetic differentiation of TB lines based on F -statistics showed that index F_{st} varied in the range of 0.005–0.073 and averaged 0.024. Within the population index F_{st} indicates the level of genetic differentiation among subpopulations (in this case among the different lines) and always has positive values ranging from 0 to 1. The small Massine line had the highest index of fixation ($F_{st} = 0.073$). Relatively low values of the differentiation index ($F_{st} = 0.005–0.006$) were determined in the lines of Blandford and Northern Dancer.

Estimation of the genetic distances between the breeds, lines and families is important to optimize breeding strategies and

Table 2. Genetic similarity coefficients (above diagonal) and Nei’s genetic distances (below diagonal) between the lines of TB horses

Line	Mr. Prospector	Native Dancer	A.P. Indy	Nasrullah	Douglas	Tourbilon	Prince Rose	Northern Dancer	Nearco	Fair Trial
Mr. Prospector	x	0.966	0.970	0.978	0.951	0.932	0.950	0.984	0.972	0.931
Native Dancer	0.034	x	0.943	0.970	0.936	0.926	0.953	0.976	0.951	0.925
A.P. Indy	0.031	0.059	x	0.960	0.925	0.930	0.938	0.964	0.950	0.906
Nasrullah	0.022	0.030	0.041	x	0.950	0.943	0.971	0.986	0.968	0.949
Douglas	0.051	0.066	0.078	0.051	x	0.948	0.943	0.950	0.941	0.932
Tourbilon	0.070	0.077	0.072	0.059	0.053	x	0.933	0.942	0.930	0.924
Prince Rose	0.051	0.048	0.064	0.030	0.059	0.070	x	0.967	0.961	0.948
Northern Dancer	0.016	0.024	0.037	0.014	0.051	0.059	0.034	x	0.975	0.933
Nearco	0.028	0.050	0.051	0.033	0.060	0.072	0.040	0.025	x	0.943
Fair Trial	0.072	0.078	0.098	0.053	0.071	0.079	0.053	0.069	0.058	x



The dendrogram of distances between different lines of Thoroughbred horses by 17 STR loci.

long-term decisions. High coefficients of genetic similarity (Table 2) related lines of Mr. Prospector and Northern Dancer (0.984), as well as Mr. Prospector and Nasrullah (0.978), Mr. Prospector and Nearco (0.972). A minimum genetic similarity was of the lines of Massine and Teddy (0.875), continuing the male line of the ancestors of the Darley Arabian breed. The genetic distances between the lines varied in a wide range from 0.014 (Nasrullah–Northern Dancer) to 0.125 (Massine–Teddy). The data obtained once more confirm the existence of genetic differentiation of the linear structure in the Thoroughbred horse breed, marking three centuries of purebred breeding. The genetic differences revealed between the lines in the Russian population of TB horses by STR loci are graphically illustrated on the dendrogram of linkage distances (Figure) built by the method of unweighted pair-group average Euclidean distances that largely coincide with the genealogical scheme of this breed. In the center of the dendrogram is clearly visible a cluster, combining all the base branches of the old

line of Phalaris (1913) – Northern Dancer (1931), Nasrullah (1940), Native Dancer (1950), Mr. Prospector (1970), Nearco (1935) and A.P. Indy (1989) lines. A certain genetic relationship can be seen between the lines of Douglas, Ribot and Prince Rose going back to famous stallion St. Simon. A comparative evaluation of the genetic diversity of microsatellite loci in horses of different lines by the number of alleles, genotypes, polymorphism level and degree of heterozygosity, as well as the dendrogram of genetic distances show that currently among the TB horses tested the most consolidated are the lines of Douglas and Massine that are continued through a limited number of descendants of these stallions. The analysis of molecular genetic features of TB lines testifies the existence of a certain genetic differentiation of the genealogical structure of the TB horses registered in Russia on the spectrum and frequencies of alleles of satellite DNA. Breeding the lines and obtaining formation on the structure of stud subpopulations contribute to the formation and consolidation of valuable genetic complexes, which, in its turn, causes interbreed heterosis and the progressive development of breeds.

Discussion

Analysis of the results of genotyping of 8091 Thoroughbred horses shows that the group studied has a breed-specific spectrum and structure of alleles of the all 17 microsatellite loci located on 13 different chromosomes. Only minor differences from TB populations of foreign countries in the frequency of rare alleles were revealed (Jungwoo et al., 2014; Shelyov et al., 2014; Rukavina et al., 2016). The results of a comparative analysis of the genetic structure of different breeds show that the TB breed has its own pool of alleles that remains stable for generations (Khrabrova, 2008; Kalashnikov et al., 2011; Khrabrova, Blohina, 2018). The formation of the genealogical structure of TB horses in Russia is influenced by global trends in the racing industry and import of horses from the United States and European countries. In recent decades, the lines of Northern Dancer, Nasrullah and Mr. Prospector have been intensively developed due to the import of sires, but the domestic line of Douglas is still important. Of course, the flow of genes has added to the gene pool of the national population, but its influence

on the genetic structure of microsatellite loci is insignificant (Khrabrova, Blohina, 2018).

The evaluation of the genetic differentiation of the linear structure of TB revealed the presence of genetic characteristics of horses of different lines on the analyzed parameters (N_a , A_e , H_o , MNA , F_{st}) which confirms the existence of interline differences at the molecular genetic level. Genetic distances between the lines varied in a wide range of 0.014–0.125, but on average they were not so great as the differences between the breeds (Kalashnikov et al., 2011). Earlier data suggest (Khrabrova, 2009) that genetic differences between TB sires of different lines are more contrasting and Nei's genetic distances varied in a wider interval, 0.041–0.234. In the Thoroughbred horse population of Bulgaria ($n = 157$) the genetic differences between the sire lines on STR loci ($F_{st} = 0.048$ –0.302) were also established (Vlaeva, Lukanova, 2015). In the Arabian breed genetic distances between the lines varied in a range of 0.040–0.200 (Zaitceva et al., 2010). It is obvious that the genetic structure and size of the reference population have a certain influence on the assessment of linear differentiation in horse breeds.

It can be expected that the interline genetic features revealed cause – to an extent – the effect of intrabreed heterosis, and the Thoroughbred horse breed with all its consolidation still has a genetic resource for further improvement. Another mechanism for maintaining this effect is the genetic differences between the best sires and the fathers of the mares.

The dendrogram of genetic distances between the lines made on the basis of allele frequency matrices of the loci studied is quite consistent with the linear structure of TB horses. The genealogical structure of the breed cluster Darley Arabian – Eclipse – Phalaris, which includes 6 related lines and in total 76.2 % of the TB population, is dominating. Within this cluster, a new line of A.P. Indy (1989) genetically is the most isolated ($F_{st} = 0.041$). Molecular genetic and pedigree studies (Cunningham et al., 2001) indicate that the genotype of the founder of the breed, sire Darley Arabian, has played a huge role in the genetic structure of TB horses. In the genealogical scheme of the breed the small Massine line stands out for its genetic individuality and is characterized by the maximum frequency of alleles (*VHL20N*, *HMS7N*, *AHT5J*, and *ASB17R*) and coefficient F_{st} (0.073). Undoubtedly, work with the most differentiated lines of Massine and Tourbilon is important for maintaining the diversity of the breed.

Linear breeding is the basis for the reproduction of genetic material in horse breeding, pig breeding and poultry farming (Barmintsev, 1972; Ernst, Zinovieva, 2008; Kharitonov et al., 2018). As an alternative to this method, cattle breeders also practice intensive use of the most valuable bulls, without dividing the breed into groups of animals by origin (Kharitonov et al., 2018). Analysis of boars' genotypes using a panel of DNA microsatellites showed that linear differentiation ($F_{st} = 0$ –0.34) is available in all pig breeds studied (Ernst, Zinovieva, 2008). In general, the use of genetic markers associated with the productive qualities of individuals significantly increases the efficiency of linear breeding (Boev et al., 2012).

According to V.O. Vitt (1957), it is very important to have a picture of genetic differentiation of Thoroughbred lines and trace the results caused by breeding methods. Genetic certification of lines provides important information about the level

of differentiation of the genealogical structure of the breed and allows controlling the level of intrabreed diversity. The undeniable advantage of microsatellite markers is the possibility of their use for studying phylogenetic relationships and microevolution of related breeds, subpopulations or lines. The standardized system of parentage testing for horses of different breeds by DNA microsatellites (van de Goor, van Haeringen, 2010) creates a true opportunity to use the results of testing to assess the diversity of populations and provides genetic control of breeding processes in the populations (Cothran, Luis, 2005; Ernst, Zinovieva, 2008).

Conclusion

There is a continuous process of development of genealogical structure and changes in the domestic population in the Thoroughbred breed that reflects the general trends of development of the breed. The analysis of genetic features of 8091 TB horses registered in Russia and belonging to 17 basic lines shows that there are differences between them both in the spectrum and frequencies of alleles, the level of polymorphism, the degree of heterozygosity and genetic distances. The data obtained confirm that the method of linear breeding is a reliable mechanism for the preservation of interbreed diversity and is quite effective even with a Thoroughbred breeding system. The system of control of a horse's origin by DNA microsatellites makes it possible to assess effectively the diversity of the breed's genealogical structure and control the selection process at the genetic level.

References

- Barmintsev Yu.N. Horse Breeding and Equestrian Sport. Moscow: Kolos Publ., 1972.
- Blokhina N.V., Khrabrova L.A. Molecular and genetic features of subpopulations of Thoroughbred horses. Konevodstvo i Konny Sport = Horse Farming and Equestrianism. 2012;4:13-15. (in Russian)
- Boev M.M., Kakushka Ye.V., Noshchenko A.S. Evaluation of intra-linear rearing and crosses of dairy cattle lines with account for inheritance of genetic markers. Vestnik of the Russian Academy of Agricultural Sciences. 2012;4:72-75. (in Russian)
- Bowling A.T., Ruvinsky A. The Genetics of the Horse. Wallingford: CABI Publ., 2000.
- Cothran E.G., Luis C. Genetic distance as a tool in conservation of rare horse breeds. In: Bodó I., Alderson L., Langlois B. (Eds.) Conservation Genetics of Endangered Horse Breeds. EAAP Publ. No. 116. Slovenia, Bled, 2005;55-72. DOI 10.3920/978-90-8686-546-8.
- Cunningham E.P., Doley J.J., Splan R.K., Bradley D.G. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. Anim. Genet. 2001;32(6):360-364.
- Ernst L.K., Zinovieva N.A. Biological Problems of Livestock in XXI Century. Moscow: Russian Academy of Agricultural Sciences, 2008. (in Russian)
- Jungwoo E., Jeong-An G., Bong-Hwan C., Kyoung-Tag D., Byung-Wook C., Heui-Soo K. Genetic profiling of Thoroughbred racehorses by microsatellite marker analysis. Genes Genomics. 2014;36: 119-123.
- Kalashnikov V.V., Khrabrova L.A., Zaitcev A.M., Zaitceva M.A., Kalinkova L.V. Polymorphism of microsatellite DNA in horses of stud and local breeds. Selskokhozyaystvennaya Biologiya = Agricultural Biology. 2011;2:41-45. (in Russian)
- Kharitonov S.N., Melnikova E.E., Osadchaya O.Yu., Yanchukov I.N., Ermilov A.N., Sermyagin A.A. In the concern to the question about principles of line breeding in Russian dairy cattle sector. Genetika i Razvedenie Zhivotnykh = Genetics and Breeding of Animals. 2018; 2:13-19. DOI 10.31043/2410-2733-2018-2-13-19. (in Russian)

- Khrabrova L.A. Monitoring of the genetic structure of breeds in horse breeding. *Russ. Agric. Sci.* 2008;34(4):261-263. DOI 10.3103/S1068367408040150.
- Khrabrova L.A. Genetic differentiation of the linear structure of Thoroughbred horse by DNA microsatellites. In: *Proc. Int. Conf. "Advances in Genetics, Breeding and Reproduction of Agricultural Animals"*. St. Petersburg, Pushkin, 2009;2:111-114. (in Russian)
- Khrabrova L.A., Blohina N.V. Genetic monitoring of the Thoroughbred breed on loci of DNA microsatellite. *Genetika i Razvedenie Zhivotnykh = Genetics and Breeding of Animals.* 2018;3:11-16. DOI 10.31043/2410-2733-2018-3-11-16. (in Russian)
- Konovalova G.K., Khlebosolova A.V. *Thoroughbred Horse Breeding in Russia and Abroad.* Moscow: Akvarium-Print Publ., 2016. (in Russian)
- Ling Y.H., Ma Y.H., Guan W.J., Cheng Y.J., Wang Y.P., Han J.L., Mang L., Zhao Q.J., He X.H., Pu Y.B., Fu B.L. Evaluation of the genetic diversity and population structure of Chinese indigenous horse breeds using 27 microsatellite loci. *Anim. Genet.* 2011;42(1):56-63.
- Nei M. *Molecular Evolutionary Genetics.* New York: Columbia Univ. Press, 1987.
- Putnová L., Štohl R., Vrtková I. Genetic monitoring of horses in the Czech Republic: a large scale study with a focus on the Czech autochthonous breeds. *J. Anim. Breed. Genet.* 2018;135(1):73-83. DOI 10.1111/jbg.12313.
- Rukavina D., Hasanbasic D., Ramic J., Zahirovic A., Ajanovic A., Beganovic K., Durnic-Pasic A., Kalamujic B., Pojskic N. Genetic diversity of Thoroughbred horse population from Bosnia and Herzegovina based on 17 microsatellite markers. *Japan. J. Veter. Res.* 2016;64(3):215-220. DOI 10.14943/jjvr.643.215.
- Shelyov A.V., Melnyk O.V., Suprun I.O., Spyrydonov S.V., Melnychuk S.D., Dzitsiuk V.V., Gorka B.M. The comparative analysis of the allele pool of Thoroughbred horses in different countries. *Iran. J. Appl. Anim. Sci.* 2014;4(3):637-641.
- Suleymanov O.I. International standards in Thoroughbred breeding records. *Konevodstvo i Konny Sport = Horse Farming and Equestrianism.* 2016;4:6-8. (in Russian)
- Van de Goor L.H.P., van Haeringen W.A. A proposal for standardization in forensic equine DNA typing: allele nomenclature for equine-specific STR loci. *Anim. Genet.* 2010;41(2):122-127. DOI 10.1111/j.1365-2052.2009.01975.x.
- Van de Goor L.H.P., van Haeringen W.A., Lenstra J.A. Population studies of 17 equine STR for forensic and phylogenetic analysis. *Anim. Genet.* 2011;42(6):627-633. DOI 10.1111/j.1365-2052.2011.02194.x.
- Vitt V.O. *Practice and Theory of Thoroughbred Horse Breeding.* Moscow, 1957.
- Vlaeva R., Lukanova N. DNA microsatellite analysis of the Thoroughbred horse population in Bulgaria. Genetic relationships between the studied sirelines. *Trakia J. Sci.* 2015;83-87. DOI 10.15547/tjs.2015.01.011.
- Weir B.S. *Genetic Data Analysis II: Methods for Discrete Population Genetic Data.* Sunderland, MA: Sinauer Associates, Inc., 1996.
- Wright J.M., Bentzen P. Microsatellites: genetics markers for the future. *Rev. Fish. Biol. Fish.* 1994;4:384-388.
- Zaitceva M.A., Khrabrova L.A., Kalinkova L.V. Intra-breed diversity on 17 loci of microsatellite DNA in Arabian horses of different lines. *Konevodstvo i Konny Sport = Horse Farming and Equestrianism.* 2010;1:19-21.

ORCID ID

L.A. Khrabrova orcid.org/0000-0003-2590-8472
N.V. Blohina orcid.org/0000-0001-7406-6385
O.I. Suleymanov orcid.org/0000-0002-4501-6642
G.A. Rozhdestvenskaya orcid.org/0000-0002-7371-8422
V.F. Pustovoy orcid.org/0000-0002-1145-2657

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

Received February 08, 2019. Revised May 06, 2019. Accepted May 20, 2019.

Неканонические эффекты вазопрессина в ангиогенезе

И.И. Херай

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: khegay@bionet.nsc.ru

Молекулярное действие вазопрессина зависит от локализации гормональных рецепторов. Основные регуляторные эффекты вазопрессина реализуются в кровеносном сосудистом русле, мозговом веществе почки и головном мозге. В настоящее время накоплена новая информация по тканеспецифичному распределению рецепторов вазопрессина, требующая обобщения. Тромбоциты и эндотелиоциты, экспрессирующие, соответственно, рецепторы типа V1a и V2, относятся к наименее исследованным гормональным мишеням вазопрессина. Вазопрессин инициирует начальную обратимую стадию активации тромбоцитов, необходимую для взаимодействия с белками внеклеточного матрикса. Адгезия тромбоцитов на эндотелий активирует в клетках секрецию ростовых факторов и ферментов метаболизма гликозаминогликанов внеклеточного матрикса. Тромбоцитарная гиалуронидаза HYAL2 гидролизует мегаполимеры гиалуроновой кислоты, иммобилизованные на эндотелиоцитах, на более короткие фрагменты. В отличие от интактной высокомолекулярной гиалуроновой кислоты с молекулярным весом в несколько мегadalton, обладающей в целом антиангиогенными свойствами, промежуточные фракции гидролиза гиалуроновой кислоты в диапазоне от 2.5 до 200 килоdalton оказывают стимулирующий эффект на ангиогенез. Межклеточные контакты тромбоцитов и эндотелиоцитов стабилизируются за счет взаимодействия адгезивных трансмембранных гликопротеинов PECAM-1. Образующиеся гетеродимеры PECAM-1 приобретают конформацию с высоким сродством к интегрину $\alpha v \beta 3$. Активация интегринов формирует контактные связи эндотелия с фибриллярными белками. Активированные эндотелиоциты секретируют фактор фон Виллебранда и Р-селектин. Эти белки аккумулированы в тельцах Вайбеля–Паладе. Вазопрессин стимулирует цАМФ-зависимый ACAP-регулируемый экзоцитоз телец Вайбеля–Паладе. Секретируемые функционально активные мультимеры фактора фон Виллебранда имеют в своем составе множественные домены связывания с другими белками и гликопептидами и дополнительно усиливают взаимодействие клеток с внеклеточным матриксом. Адгезия на фибриллярный коллаген и мембранные гликопротеины в кооперации с эффектами PECAM-1– $\alpha v \beta 3$ интегриновых комплексов фиксирует клеточные агрегаты в окружающем интерстиции и ориентирует миграцию пролиферирующих эндотелиоцитов в направлении локальных градиентов ростовых факторов ангиогенеза. Нейрогормональная регуляция секреторной активности тромбоцитов и эндотелиоцитов функционально связывает пролиферацию и миграцию эндотелиоцитов в процессе ангиогенеза и интегрирует их с адаптивными возможностями организма.

Ключевые слова: вазопрессин; рецепторы V1a и V2; тромбоцит; эндотелиоцит; гиалуроновая кислота; гиалуронидаза HYAL2; белок PECAM-1; интегрин $\alpha v \beta 3$.

Для цитирования: Херай И.И. Неканонические эффекты вазопрессина в ангиогенезе. Вавиловский журнал генетики и селекции. 2019;23(5):575-581. DOI 10.18699/VJ19.527

Noncanonical effects of vasopressin in angiogenesis

I.I. Khegay

Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
✉ e-mail: khegay@bionet.nsc.ru

The molecular action of vasopressin depends on the localization of hormonal receptors. The basic physiological effects of vasopressin are manifested in the blood vasculature, renal inner medulla and brain. To date, new information concerning the tissue-specific spreading of vasopressin receptors has been accumulated, and it needs to be summarized. Platelets and endotheliocytes expressing V1a and V2 receptor types, respectively, are related to less investigated targets of the hormone. Vasopressin induces the initial reversible stage of platelet activation, required for interaction with intercellular matrix proteins. Platelet adhesion on endothelium activates cellular secretion of growth factors and enzymes for intercellular matrix glucosamine metabolism. Platelet hyaluronidase HYAL2 hydrolyses high-molecular hyaluronic acid to shorter fragments. Unlike intact hyaluronic acid with a molecular weight of several megadaltons, generally showing distinctive antiangiogenic properties, intermediate fractions of hyaluronan hydrolysis in a range from 2.5 to 200 kilodaltons have a stimulating effect on angiogenesis. Intercellular contacts between platelets and endotheliocytes are stabilized due to adhesive transmembrane glycoprotein PECAM-1 interaction. Resulting PECAM-1 heterodimers acquire conformation with high affinity to integrins $\alpha v \beta 3$. Integrin activation forms contact links between endothelium and fibrillar proteins. Activated endotheliocytes secrete von Willebrand factor and P-selectin. These proteins are accumulated in Weibel–Palade bodies. Vasopressin stimulates cAMP-dependent ACAP-regulated exocytosis of Weibel–Palade bodies. von Willebrand factor possesses adhesive properties and additionally accelerates interaction of cells with the intercellular matrix. Adhesion on fibrillar collagen and membrane glycoproteins in cooperation with effects of PECAM-1– $\alpha v \beta 3$ integrin complexes fixes cell aggregates in the surrounding interstitium and promotes proliferating endotheliocyte migration in according to the direction of local growth factor gradients during angiogenesis. Neurohormonal regulation of platelet and endotheliocyte secretory activity

functionally link proliferation and migration of endotheliocytes during angiogenesis and integrate it according to the adaptive capacity of the entire organism.

Key words: vasopressin; V1a- and V2-receptors; platelet; endotheliocyte; hyaluronic acid; hyaluronidase HYAL2; protein PE-CAM-1-integrin $\alpha\beta 3$.

For citation: Kheday I.I. Noncanonical effects of vasopressin in angiogenesis. *Vavilovskii Zhurnal Genetiki i Selekcii* = Vavilov Journal of Genetics and Breeding. 2019;23(5):575–581. DOI 10.18699/VJ19.527 (in Russian)

Введение

Вазопрессин – один из ключевых нейрогормональных регуляторов гомеостаза внутренней среды. Широкий спектр эффектов гормона определяется составом рецепторов вазопрессина в тканях-мишенях. Все рецепторы вазопрессина относятся к семейству трансмембранных рецепторов, сопряженных с G-белками. Действие рецепторов типа V1 опосредовано через G_q – фосфолипаза C – фосфатидилинозитол-сигнальный путь. Внутриклеточная трансдукция гормонального стимула альтернативных рецепторов типа V2 реализуется при участии G_s -белка, аденилатциклазы и цАМФ (Thibonnier et al., 2001; Bankir et al., 2017). В висцеральных системах наиболее широко представлены рецепторы типа V1a. Данные рецепторы, прежде всего, опосредуют классический сосудосуживающий эффект вазопрессина, развивающийся за счет сократительной реакции гладкомышечных волокон кровеносных сосудов на повышение уровня внутриклеточного кальция (Aisenbrey et al., 1981; Landry et al., 1997). Другие физиологические эффекты рецепторов V1a связаны с агрегацией и секреторной активностью тромбоцитов (Thibonnier et al., 1993), гликогенолизом, секрецией ростовых факторов и факторов свертывания крови в печени (Ostrowski, Young, 1993). Кинетика рецепторной активности рецепторов V1 характеризуется короткими циклами с быстрой десенситизацией и восстановлением мембранной локализации после экспозиции лигандом и в целом коррелирует со скоростью контролируемых процессов. Интернализация рецепторов V2 проходит медленнее и включает стадию аккумулялирования в перинуклеарных везикулах (Innamorati et al., 2001). Рецепторы типа V2 определяются преимущественно в дистальных канальцах нефрона и собирательных трубках почки, где они непосредственно вовлечены в цАМФ-зависимую регуляцию уровня реабсорбции воды и играют ключевую роль в функционировании системы осмотического концентрирования (Lolait et al., 1992).

В настоящее время накоплена дополнительная информация по локализации рецепторов вазопрессина. Рецепторы V1a были обнаружены в гломерулярных мезангиальных клетках (Ghosh et al., 2001), интерстициальных и люминальных эпителиальных клетках собирательных трубок мозгового вещества почки (Селивёрстова и др., 2009), эпителии тонкого кишечника (Chiu et al., 2002). Рецепторы V2 выявлены в эндотелии кровеносных сосудов и гепатоцитах (Koshimizu et al., 2012). Установлено, что опухоли эпителиального происхождения экспрессируют все известные типы рецепторов вазопрессина, включая специфические для головного мозга рецепторы V1b (Requeux et al., 2004; MacKinnon et al., 2009). В ряде случаев развивающиеся в течение нескольких минут физиологические реакции сопровождаются более продолжительными по времени молекулярными эффектами вазопрес-

сина, связанными с активацией генов и синтезом белков, транспортом и секрецией цитоплазматических везикул и пролиферацией клеток (Tahara et al., 2008). Ткани, экспрессирующие рецепторы типа V1a, способны транслировать гормональный сигнал на внутриклеточные митогенные каскады MAPK/ERK, контролирующие деление клеток. Так, вазопрессин индуцирует синтез ДНК и митозы в клетках тонкого кишечника (Chiu et al., 2002). В экспериментах с частичной гепатэктомией антагонисты рецепторов V1a блокируют восстановительную регенерацию печени, а введение синтетических препаратов вазопрессина, наоборот, усиливает скорость регенерации. Аналогичным образом вазопрессин стимулирует пролиферацию гломерулярных мезангиальных клеток почки, а антагонисты рецепторов V1a оказывают противоположный цитостатический эффект (Koshimizu et al., 2012).

Вазопрессин участвует в дифференцировке кардиомиоцитов и гипертрофии кардиоваскулярной ткани (Gutkowska et al., 2007). Показано, что рецепция вазопрессина инициирует агрегацию и активацию рецепторов эпидермального фактора роста (Ghosh et al., 2001). В отсутствие эндогенного вазопрессина изменяется экспрессия белков протеасом и угнетается рост карциносаркомы Walker 256 (Sharova et al., 2008). В безъядерных тромбоцитах и эндотелиоцитах, экспрессирующих рецепторы V2, пролиферативные эффекты вазопрессина проявляются опосредованно и требуют предварительной стадии активации клеток.

Активация тромбоцитов и эндотелиоцитов

Наряду с опухолями, тромбоциты и эндотелиоциты относятся к наименее исследованным клеточным мишеням вазопрессина. Можно предположить, что в ряде случаев, в том числе в опухолях, центральная нейрогормональная регуляция синхронизирует в тканях пролиферативный ответ на вазопрессин (Gagona et al., 2015). Активная пролиферация тканей всегда сопровождается реорганизацией локальной системы кровоснабжения (Maharaj et al., 2006). Пролиферативные эффекты вазопрессина тесно сопряжены с механизмом действия васкулоэндотелиального ростового фактора VEGF – основного индуктора ангиогенеза. В норме ангиогенез в тканях начинается с повреждения по тем или иным причинам стенок существующих кровеносных сосудов и активации митозов в эндотелиоцитах. В частности, это происходит при заживлении ран. Аналогичный процесс наблюдается в прогрессирующих опухолях, при этом в роли повреждающего и активирующего агента выступает инвазия солидного новообразования. Непосредственным событием, инициирующим ангиогенез, является взаимодействие тромбоцитов и эндотелиоцитов, экспрессирующих, соответственно, рецепторы V1a и V2 вазопрессина.

Как в физиологических, так и при патологических состояниях локальный дисбаланс кровоснабжения тканей и нехватка кислорода активируют транскрипционный индуцируемый гипоксией фактор HIF-1 (hypoxia inducible Factor-1). Стабилизированный белок HIF-1 транслоцируется в ядро и усиливает экспрессию генов ферментов гликолиза и ростовых факторов (Klock et al., 2011). Ключевым ростовым фактором для ангиогенеза является VEGF. Тромбоциты – основной источник VEGF на начальной стадии пролиферации эндотелия (Battinelli et al., 2011). Проллиферирующие клетки эндотелия секретируют адгезивные белки, необходимые для дальнейшей миграции клеток (Shibuya, 2013). Несмотря на значительный прогресс в функциональном анализе тромбоцитов, регуляция секреторной активности и внеклеточные регуляторные эффекты тромбоцитов остаются областью активного поиска (Wojtukiewicz et al., 2017; van der Meijden, Heemskerk, 2019).

Известно, что форменные элементы крови и эндотелиоциты происходят от общего предшественника – первичных гемангиобластов – и сохраняют сходный механизм реакции на локальные деструктивные факторы (Sequeira Lopez et al., 2003; Lu et al., 2007). Контакт с внешними белковыми субстратами, цитокинами и гормонами активирует секреторную функцию клеток. Низкомолекулярные растворимые лиганды кинетически более эффективны и быстрее связываются с рецепторами. Вазопрессин относится к гуморальным стимуляторам начальной обратимой стадии гормон-рецепторной активации тромбоцитов. Действуя через V1a-фосфатидилинозитол-кальциевый сигнальный путь, вазопрессин активирует мембранные Na^+/H^+ ко-транспортёры, вызывая дозозависимое увеличение объема цитоплазмы и экспозицию наружу молекул фосфатидилсерина (Tomasiak et al., 2008). Реорганизация клеточной структуры тромбоцитов сопровождается образованием микровезикул, дегрануляцией и способствует активации мембранных интегриновых рецепторов (Heemskerk et al., 2000). Активированные интегрины формируют более стабильные связи клеток с фибриллярными белками соединительной ткани, преимущественно с коллагеном, и переводят тромбоциты в состояние дальнейшей необратимой активации. Адгезия таких тромбоцитов на базальную мембрану поврежденных кровеносных сосудов в свою очередь активирует эндотелиоциты.

Миграция, адгезия, активация тромбоцитов и эндотелиоцитов осуществляются при непосредственном участии соединительной ткани. Интерстиций функционирует одновременно как биомеханическая поддерживающая среда и как активное звено в сигнальной коммуникации гетеро- и гомогенных специализированных клеток. Основным пластический компонент интерстициальной ткани – гиалуроновая кислота, играющая важную роль в межклеточном взаимодействии. Строма в опухолях характеризуется повышенным содержанием и активным метаболизмом гиалуроновой кислоты (Nguyen et al., 2017; McCarthy et al., 2018). Гиалуроновая кислота может поразному влиять на внутриклеточные процессы в зависимости от степени полимеризации молекул гиалуронана (Vigetti et al., 2014). В настоящее время рассматривается несколько вариантов вовлечения гиалуроновых кислот в

клеточную пролиферацию. Роль гиалуроновой кислоты в малигнизации опухолей подробно рассмотрена в обзорах (Sironen et al., 2011; Bohaumilutzky et al., 2017). Авторы показали неоднозначную многоплановую роль гиалуронана.

Пролиферация и метастазирование опухолей всегда сопровождаются накоплением гиалуроновой кислоты. Агрессивность аденокарциномы человека прямо ассоциирует с высоким уровнем гиалуронана в строме. Гиалуроновая кислота активирует рецептор ErbB2 эпидермального фактора роста и модулирует внутриклеточную трансдукцию пролиферативного сигнала в эпителиальных клетках и эндотелиоцитах. Важный фактор, влияющий на малигнизацию опухоли молочной железы, – соотношение активности гиалуронансинтаз и гиалуронидаз (Siiskonen et al., 2013; Auvinen et al., 2014). Локализованные на клеточной мембране гиалуронансинтазы 1, 2 и 3 последовательно присоединяют гликозаминогликановые мономеры и нарабатывают высокомолекулярный субстрат, экструдируемый во внеклеточный матрикс. В нормальных физиологических условиях гиалуроновая кислота представлена преимущественно в макромолекулярной форме с общим весом в несколько мегадальтон. Интактные гиалуроновые мегаполимеры ингибируют пролиферацию и миграцию и обладают антиангиогенными свойствами (Tian et al., 2013).

Обязательное условие реализации пролиферативных и васкулогенных эффектов гиалуроновых кислот – их взаимодействие со своим специфическим мембранным рецептором CD44. Он относится к классическим интегральным гликопротеинам с несколькими сайтами гликозилирования во внеклеточном эктодомене. Размеры белка варьируют в зависимости от количества и длины лигированных олигосахаридных гликозидов (Sackstein, 2011). Свободные интерстициальные гиалуроновые фрагменты аффинно связываются с локализованными на плазматических мембранах белками CD44, иммобилизуются и переводят рецепторный комплекс в конформацию с дополнительными сайтами распознавания белков. Образующаяся надмембранная структура интегрирует трансдукцию пролиферативных сигналов всех локальных ростовых факторов (Senbanjo, Chellaiah, 2017). Связующий домен CD44 содержит сайты распознавания и связывания коллагена, ламинина, фибронектина, селектина в структуре фибриллярной соединительной ткани (Goodison et al., 1999). Белки фибриллярной соединительной ткани выполняют опорную функцию, а также участвуют в трансляции пролиферативных сигналов. В частности, в аминокислотной последовательности ламинина зафиксированы десятки полноразмерных доменов эпидермального фактора роста EGF, основного регулятора пролиферации эпителиальных тканей (Mayer et al., 1995). Суммарный регуляторный эффект гиалуронан-CD44 рецепторных комплексов зависит от линейных размеров полимерного лиганда (Bourguignon et al., 2017). Показано, что нативные полноразмерные цепи гиалуронана за счет своей длины способствуют локальной концентрации и кластеризации CD44 и в целом ингибируют ангиогенез (Yang et al., 2012). Под действием гиалуронидазы HYAL2 из мегаполимеров гиалуроновой кислоты образуются более короткие фрагменты весом около 200 кДа. У фрагментированной гиалуроновой кислоты

сохраняется способность взаимодействовать с рецепторами CD44, но такие комплексы в отличие от мегамолекулярного гиалуронана оказывают противоположный эффект и стимулируют трансляцию пролиферативных сигналов на внутриклеточные каскады с выходом на протеинкиназу С и митоген-активируемые протеинкиназы (Slevin et al., 2002). Ангиогенный эффект у промежуточных фракций гиалуронана сохраняется вплоть до фрагментов размером около 2.5 кДа (Stern et al., 2006). Гомологичный фермент гиалуронидаза HYAL1 выполняет более глубокий дальнейший гидролиз гиалуронана до олиго- и мономеров. Олигомерные фрагменты вызывают провоспалительную реакцию, индуцируя сигнальный путь NF κ B/Stat-3 и синтез малых интерферирующих РНК miR-2 (Bourguignon et al., 2009). Низкомолекулярные гиалуронаны также усиливают экспрессию матриксных металлопротеиназ и провоспалительных цитокинов (Voelcker et al., 2008).

В работе (Albeiroti et al., 2015) показано, что активированные тромбоциты транслоцируют на внешнюю поверхность гиалуронидазу HYAL2. Угнетение циркуляции тромбоцитов в крови фармакологической блокадой интегриновых рецепторов адгезии тромбоцитов оказывает выраженное противоопухолевое действие (Erpenbeck et al., 2010). Большинство клеток, рецептирующих гиалуронан, экспрессируют полноразмерный CD44, собранный из правильно ориентированных экзонов. Особенностью некоторых типов эпителиальных клеток и злокачественных опухолей эпителиального происхождения является преимущественная экспрессия альтернативно сплайсированного варианта белка изоформы CD44v (Williams et al., 2013). Локализация CD44v на клеточной мембране – важнейший фактор миграции и инвазии опухолевых клеток (Wang et al., 2018).

Молекулярные механизмы межклеточной адгезии тромбоцит-эндотелиоцит

Непосредственные межклеточные контакты тромбоцитов и эндотелиоцитов стабилизируются за счет взаимодействия трансмембранных гликопротеинов PECAM-1 (platelet/endothelial cell adhesion molecule-1). Этот тип адгезивных белков активно экспрессируется в эндотелиоцитах, эпителиальных гемангиоэндотелиомах, лейкоцитах и тромбоцитах (Gratzinger et al., 2003). PECAM-1 имеет типичную структуру интегрального мембранного белка с tandemными внеклеточными рецепторными доменами и протяженным цитоплазматическим отделом, включающим несколько фосфорилируемых сайтов ITIM (immunoreceptor tyrosine-based inhibitory motif), регулирующих активность фосфатазы SHP-2 (Src-homology 2 domain (SH2)-containing phosphatase-2) (Zhang et al., 2015). Нерепцепторная протеин-тирозин фосфатаза SHP-2 дефосфорилирует киназу фокальной адгезии (PTK2 – focal adhesion kinase 1) – ключевой фермент, ассоциированный с цитоскелетом и контролирующий перемещение клеток (Zachary, Rozengurt, 1992; Zhu et al., 2010). На уровне киназы фокальной адгезии интегрируются сигнальные пути нейротрофинов, интегринов и васкулогенных ростовых факторов (Eliceiri et al., 2002). Мономеры PECAM-1, локализованные на соседних контактирующих клетках, собираются попарно в связующие димеры. Гомогенная

димеризация PECAM-1 эндотелиоцитов формирует трехмерную сеть межклеточных контактов в процессе формирования трубчатой структуры эндотелиального пласта в растущем капилляре (Lertkietmongkol et al., 2016). Гетерогенные димеры PECAM-1 образуются при контакте тромбоцитов и эндотелиоцитов. Введение экзогенных препаратов вазопрессина усиливает адгезию тромбоцитов на эндотелий кровеносных сосудов и не влияет на собственную агрегацию тромбоцитов (Mannucci, 1997). Показано, что гетерогенные димеры PECAM-1, фиксирующие связь тромбоцитов и эндотелиоцитов, ингибируют образование гомогенных димеров PECAM-1 между тромбоцитами и угнетают агрегацию и образование тромбов (Falati et al., 2006).

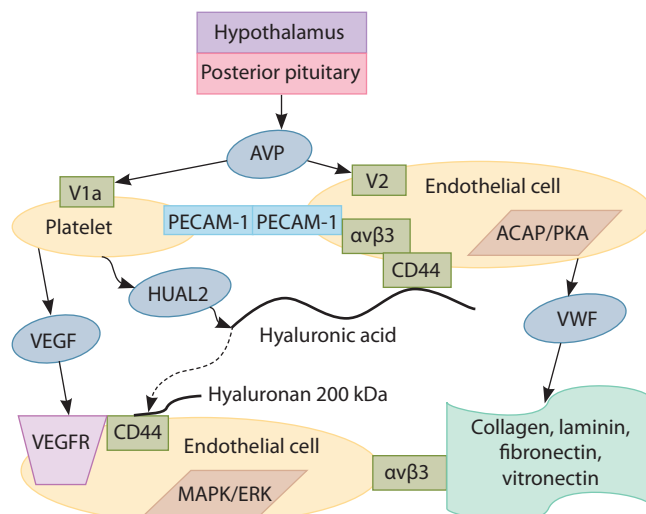
В димерной конформации молекула PECAM-1 приобретает дополнительные свойства, проявляющиеся в усилении способности рецептировать и активировать интегрины, расширяя таким образом спектр межклеточной сигнальной коммуникации (Zhao, Newman, 2001). Интегрины выполняют функцию основных рецепторных белков на стадии формирования связей клеток с интерстициальным матриксом. Лигандами служат фибронектин, витронектин, коллаген, ламинин и другие белки, экспрессирующие пептидный мотив RGD (Arg-Gly-Asp) в сайтах связывания интегринов (Kunicki et al., 1997; Mahalingam et al., 2014). Структурно интегрины представляют облигатные гетеродимеры α - и β -субъединиц. Активация интегринов заключается в переходе α -субъединиц в состояние с более высокой аффинностью к пептидным мотивам RGD. Гликозилированная β -субъединица выполняет функцию кофактора, усиливающего адгезию. Специфичность адгезии белков реализуется за счет внеклеточных доменов α -субъединиц, а цитоплазматический отдел β -субъединиц участвует в передаче сигналов на внутриклеточные регуляторные каскады (Eliceiri et al., 2002). В активированных ангиогенными факторами эндотелиоцитах резко усиливается экспрессия интегрин α v β 3 (Liu et al., 2008). Интегрины α v β 3 также входят в состав белков, синтезируемых и секретируемых тромбоцитами, и в норме необходимы для образования связей с фактором фон Виллебранда и фибриллярными белками матрикса (Liu et al., 2009). В малигнизующих опухолях наблюдается усиление экспрессии интегрин α v β 3. Этот процесс коррелирует с нарастающим объемом контактов опухолевой паренхимы и интерстиция (Felding-Habermann, 2003; Rolli et al., 2003). Установлено, что при адгезии опухолевых клеток и тромбоцитов интегрины α v β 3 активируются, колокализуются и далее каким-то образом участвуют в регуляции состава секретируемых тромбоцитами альфа-гранул (Weber et al., 2016).

Альфа-гранулы представляют наиболее обширную группу везикул в цитоплазме тромбоцитов (Blair, Flaumenhaft, 2009). В альфа-гранулах депонированы сотни белков, транспортируемых экзоцитозом на плазматическую мембрану или секретируемых во внеклеточное пространство. Список идентифицированных белков включает белки адгезии, комплексы SNARE распознавания мембран везикул (soluble N-ethylmaleimide-sensitive factor attachment protein receptors), ростовые факторы, ферменты (Rendu, Brohard-Bohn, 2001). Альфа-гранулы гетерогенны

по составу секретируемых белков. В тромбоцитах дифференцированно локализованы субпопуляции везикул с ангиогенными ростовыми факторами VEGF и PDGF (platelet-derived growth factor) и отдельно от них альфа-гранулы с антиангиогенным пептидом эндостатином (Battinelli et al., 2011). Выявлено, что в составе альфа-гранул также депонируется гиалуронидаза HYAL2, один из ключевых ферментов метаболизма гиалуроновой кислоты (Albeiroti et al., 2015). Показано, что адгезия тромбоцитов инициирует секрецию гиалуронидазы HYAL2.

Вазопрессин относится к быстродействующим гуморальным активаторам адгезии тромбоцитов (Tomasiak et al., 2008). Следовательно, секреция HYAL2, по крайней мере, частично находится под гормональным контролем вазопрессина. Фермент HYAL2 проявляет свою активность на начальных стадиях фрагментирования мегаполимерных цепей интактных гиалуроновых кислот и, по сути, функционирует как молекулярный переключатель регуляторных эффектов гиалуронана с угнетения ангиогенеза на стимулирование (рисунок). Для эффективной работы гиалуронидазы HYAL2 необходимо, чтобы гиалуроновая кислота была в связанном с мембранным рецептором CD44 состоянии и иммобилизирована на базолатеральной поверхности эндотелиоцитов.

Вазопрессин инициирует адгезию тромбоцитов на эндотелий и индуцирует в клетках секреторные процессы. Этот гормональный эффект имеет практическое значение для медицины. Препараты вазопрессина широко используются в лечении наследственных дефектов кровеносной системы. Синтетический DDAVP (1-деамино-8-D-аргинин вазопрессин), специфический агонист рецепторов V2, применяется для быстрого повышения свертываемости крови. Непосредственно под действием вазопрессина эндотелиоциты секретируют фактор фон Виллебранда, стабилизирующий антигемофильный фактор свертывания крови VIII, и экспонируют на мембрану Р-селектин (Kanwar et al., 1995). Эти белки также относятся к семейству адгезивных молекул и участвуют во взаимодействии с внеклеточным матриксом и трансмиграции клеток. Фактор фон Виллебранда и Р-селектин – основные секретруемые компоненты в составе телец Вайбеля–Паладе эндотелиоцитов. В опытах *in vitro* с использованием ПЦР с обратной транскрипцией показано, что эндотелиоциты экспрессируют полноразмерные функционально активные рецепторы V2 вазопрессина (Kaufmann et al., 2000). Экспозиция с DDAVP стимулировала секрецию фактора фон Виллебранда, опосредованную цАМФ. Рецепция вазопрессина на рецепторы типа V2 активирует в эндотелиоцитах цАМФ-зависимую протеинкиназу А и переключает свободные каталитические субъединицы на взаимодействие с регуляторным белком ACAP (A kinase anchoring protein). Белки ACAP фиксируют протеинкиназу А на локальных клеточных органеллах, вовлеченных в транспорт и экзоцитоз секреторных гранул Вайбеля–Паладе (Nedvetsky et al., 2009; Biesemann et al., 2017). Многоуровневая регуляция адгезии эндотелиоцитов на фибриллярный коллаген и мембранные гликопротеины в кооперации с PECAM-1/ α v β 3-интегриновыми комплексами ориентирует миграцию пролиферирующих клеток в окружающем интерстиции в направлении локальных



Neurohormonal regulation of angiogenesis.

AVP, arginin-vasopressin; V1a, vasopressin receptor type V1a; V2, vasopressin receptor type V2; PECAM-1, platelet/endothelial cell adhesion molecule-1; α v β 3, integrin; CD44, hyaluronan receptor; HUAL2, hyaluronidase-2; VEGF, vascular endothelial growth factor; VEGFR, receptor of vascular endothelial growth factor; MAPK/ERK, cascade of mitogen-activated protein kinases; ACAP, A kinase anchoring protein; PKA, protein kinase A; VWF, von Willebrand factor.

градиентов ростовых факторов в процессе ангиогенеза (Zhu et al., 2010; Privratsky et al., 2011; Williams et al., 2013).

Заключение

Нейрогормональная регуляция секреторной активности тромбоцитов и эндотелиоцитов функционально корректирует динамику отдельных стадий локальных пролиферативных процессов и интегрирует их с адаптивными возможностями всего организма.

Список литературы / References

- Селиверстова Е.В., Соловьёв А.А., Насыров Р.А., Наточин Ю.В.
Изучение локализации рецепторов вазопрессина в почке детей при гломерулонефрите. Нефрология. 2009;13(4):51-58.
[Seliverstova E.V., Solov'yov A.A., Nasyrov R.A., Natochin Y.V. Study of localization of vasopressin receptors in kidney of children with glomerulonephritis. Nefrologiya = Nephrology. 2009;13(4): 51-58. (in Russian)]
- Aisenbrey G.A., Handelsman W.A., Arnold P., Manning M., Schrier R.W. Vascular effects of arginine vasopressin during fluid deprivation in the rat. J. Clin. Invest. 1981;67(4):961-996.
- Albeiroti S., Ayasoufi K., Hill D.R., Shen B., de la Motte C.A. Platelet hyaluronidase-2: an enzyme that translocates to the surface upon activation to function in extracellular matrix degradation. Blood. 2015; 125(9):1460-1469. DOI 10.1182/blood-2014-07-590513.
- Auvinen P., Rilla K., Tumelius R., Tammi M., Sironen R., Soini Y., Kosma V.M., Mannermaa A., Viikari J., Tammi R. Hyaluronan synthases (HAS1-3) in stromal and malignant cells correlate with breast cancer grade and predict patient survival. Breast Cancer Res. Treat. 2014;143(2):277-286. DOI 10.1007/s10549-013-2804-7.
- Bankir L., Bichet D.G., Morgenthaler N.G. Vasopressin: physiology, assessment and osmosensation. J. Intern. Med. 2017;282(4):284-297. DOI 10.1111/joim.12645.
- Battinelli E.M., Markens B.A., Italiano J.E. Release of angiogenesis regulatory proteins from platelet alpha granules: modulation of physiologic and pathologic angiogenesis. Blood. 2011;118(5):1359-1369. DOI 10.1182/blood-2011-02-334524.

- Biesemann A., Gorontzi A., Barr F., Gerke V. Rab35 regulates evoked exocytosis of endothelial Weibel-Palade bodies. *J. Biol. Chem.* 2017;292(28):11631-11640. DOI 10.1074/jbc.M116.773333.
- Blair P., Flaumenhaft R. Platelet α -granules: Basic biology and clinical correlates. *Blood. Rev.* 2009;23(4):177-189. DOI 10.1016/j.blr.2009.04.001.
- Bohaumilitsky L., Huber A.K., Stork E.M., Wengert S., Woelfl F., Boehm H. A. trickster in disguise: Hyaluronan's ambivalent roles in the matrix. *Front. Oncol.* 2017;7:242. DOI 10.3389/fonc.2017.00242.
- Bourguignon L.Y.W., Earle C., Shiina M. Activation of matrix hyaluronan-mediated CD44 signaling, epigenetic regulation and chemoresistance in head and neck cancer stem cells. *Int. J. Mol. Sci.* 2017;18(9):1849-1863. DOI 10.3390/ijms18091849.
- Bourguignon L.Y.W., Spevak C., Wong G., Xia W., Gilad E. Hyaluronan-CD44 interaction with PKC ϵ promotes oncogenic signaling by the stem cell marker, Nanog and the production of microRNA-21 leading to down-regulation of the tumor suppressor protein, PDCD4, anti-apoptosis and chemotherapy resistance in breast tumor cells. *J. Biol. Chem.* 2009;284(39):26533-26546. DOI 10.1074/jbc.M109.027466.
- Chiu T., Wu S.S., Santiskulvong C., Tangkijvanich P., Yee H.F., Rozen-gurt E. Vasopressin-mediated mitogenic signaling in intestinal epithelial cells. *Am. J. Physiol. Cell. Physiol.* 2002;282(3):C434-C450. DOI 10.1152/ajpcell.00240.2001.
- Eliceiri B.P., Puente X.S., Hood J.D., Stupack D.G., Schlaepfer D.D., Huang X.Z., Sheppard D., Cheres D.A. Src-mediated coupling of focal adhesion kinase to integrin $\alpha\beta 5$ in vascular endothelial growth factor signaling. *J. Cell. Biol.* 2002;157(1):149-160. DOI 10.1083/jcb.200109079.
- Erpenbeck L., Nieswandt B., Schon M., Pozgajova M., Schon M.P. Inhibition of platelet GPIb α and promotion of melanoma metastasis. *J. Invest. Dermatol.* 2010;130(2):576-586. DOI 10.1038/jid.2009.278.
- Falati S., Patil S., Gross P.L., Stapleton M., Merrill-Skoloff G., Barrett N.E., Pixton K.L., Weiler H., Cooley B., Newman D.K., Newman P.J., Furie B.C., Furie B., Gibbins J.M. Platelet PECAM-1 inhibits thrombus formation *in vivo*. *Blood.* 2006;107(2):535-541. DOI 10.1182/blood-2005-04-1512.
- Felding-Habermann B. Integrin adhesion receptors in tumor metastasis. *Clin. Exp. Metastasis.* 2003;20(3):203-213.
- Garona J., Pifano M., Orlando U.D., Pastrian M.B., Iannucci N.B., Ortega H.H., Podesta E.J., Gomez D.E., Ripoll G.V., Alonso D.F. The novel desmopressin analogue [V4Q5]dDAVP inhibits angiogenesis, tumour growth and metastases in vasopressin type 2 receptor-expressing breast cancer models. *Int. J. Oncol.* 2015;46(6):2335-2245. DOI 10.3892/ijo.2015.2952.
- Ghosh P.M., Mikhailova M., Bedolla R., Kreisberg J.I. Arginine vasopressin stimulates mesangial cell proliferation by activating the epidermal growth factor receptor. *Am. J. Physiol. Renal. Physiol.* 2001;280(6):F972-F979. DOI 10.1152/ajprenal.2001.280.6.F972.
- Goodison S., Urquidí V., Tarin D. CD44 cell adhesion molecules. *Mol. Pathol.* 1999;52(4):189-196.
- Gratzinger D., Canosa S., Engelhardt B., Madri J.A. Platelet endothelial cell adhesion molecule-1 modulates endothelial cell motility through the small G-protein Rho. *FASEB J.* 2003;17(11):1458-1469. DOI 10.1096/fj.02-1040.com.
- Gutkowska J., Miszkurka M., Danalache B., Gassanov N., Wang D., Jankowski M. Functional arginine vasopressin system in early heart maturation. *Am. J. Physiol. Heart Circ. Physiol.* 2007;293(4):H2262-H2270. DOI 10.1152/ajpheart.01320.2006.
- Heemskerk J.W., Siljander P.R., Bevers E.M., Farndale R.W., Lindhout T. Receptors and signalling mechanisms in the procoagulant response of platelets. *Platelets.* 2000;11(6):301-306.
- Innamorati G., Le Gouill C., Balamotis M., Birnbaumer M. The long and the short cycle. Alternative intracellular routes for trafficking of G-protein-coupled receptors. *J. Biol. Chem.* 2001;276(16):13096-13103. DOI 10.1074/jbc.M009780200.
- Kanwar S., Woodman R.C., Poon M.C., Murohara T., Lefer A.M., Davenpeck K.L., Kubes P. Desmopressin induces endothelial P-selectin expression and leukocyte rolling in postcapillary venules. *Blood.* 1995;86(7):2760-2766.
- Kaufmann J.E., Oksche A., Wollheim C.B., Gonther G., Rosenthal W., Vischer U.M. Vasopressin-induced von Willebrand factor secretion from endothelial cells involves V2 receptors and cAMP. *J. Clin. Invest.* 2000;106(1):107-116. DOI 10.1172/JCI9516.
- Koshimizu T.A., Nakamura K., Egashira N., Hiroshima M., Nonoguchi H., Tanoue A. Vasopressin V1a and V1b receptors: from molecules to physiological systems. *Physiol. Rev.* 2012;92(4):1813-1864. DOI 10.1152/physrev.00035.2011.
- Krock B.L., Skuli N., Simon M.C. Hypoxia-induced angiogenesis: good and evil. *Genes Cancer.* 2011;2(12):1117-1133. DOI 10.1177/1947601911423654.
- Kunicki T.J., Annis D.S., Felding-Habermann B. Molecular determinants of arg-gly-asp ligand specificity for beta3 integrins. *J. Biol. Chem.* 1997;272(7):4103-4107.
- Landry D.W., Levin H.R., Gallant E.M., Ashton R.C., Seo S., D'Alessandro D., Oz M.C., Oliver J.A. Vasopressin deficiency contributes to the vasodilation of septic shock. *Circulation.* 1997;95(5):1122-1125.
- Lertkiatmongkol P., Liao D., Mei H., Hu Y., Newman P.J. Endothelial functions of PECAM-1 (CD31). *Curr. Opin. Hematol.* 2016;23(3):253-259. DOI 10.1097/MOH.0000000000000239.
- Liu Y., Zhao F., Gu W., Yang H., Meng Q., Zhang Y., Yang H., Duan Q. The roles of platelet GPIIb/IIIa and $\alpha\beta 3$ integrins during hela cells adhesion, migration, and invasion to monolayer endothelium under static and dynamic shear flow. *J. Biomed. Biotechnol.* 2009;2009(829243):1-9. DOI 10.1155/2009/829243.
- Liu Z., Wang F., Chen X. Integrin $\alpha\beta 3$ -targeted cancer therapy. *Drug Dev. Res.* 2008;69(6):329-339. DOI 10.1002/ddr.20265.
- Lolait S.J., O'Carroll A.M., McBride O.W., König M., Morel A., Brownstein M.J. Cloning and characterization of a vasopressin V2 receptor and possible link to nephrogenic diabetes insipidus. *Nature.* 1992;357(6376):336-339.
- Lu S.J., Feng Q., Ivanova Y., Luo C., Li T., Li F., Honig G.R., Lanza R. Recombinant HoxB4 fusion proteins enhance hematopoietic differentiation of human embryonic stem cells. *Stem Cells Dev.* 2007;16(4):547-559. DOI 10.1089/scd.2007.0002.
- MacKinnon A.C., Tufail-Hanif U., Wheatley M., Rossi A.G., Haslett C., Seckl M., Sethi T. Targeting V1A-vasopressin receptors with [Arg6, D-Trp7,9, NmePhe8]-substance P (6-11) identifies a strategy to develop novel anti-cancer therapies. *Br. J. Pharmacol.* 2009;156(1):36-47. DOI 10.1111/j.1476-5381.2008.00003.x.
- Mahalingam B., Van Agthoven J.F., Xiong J.-P., Alonso J.L., Adair B.D., Rui X., Anand S., Mehrbod M., Mofrad M.R.K., Burger C., Goodman S.L., Arnaout M.A. Atomic basis for the species-specific inhibition of αV integrins by monoclonal antibody 17E6 is revealed by the crystal structure of $\alpha V\beta 3$ ectodomain-17E6 fab complex. *J. Biol. Chem.* 2014;289(20):13801-13809. DOI 10.1074/jbc.M113.546929.
- Maharaj A.S., Saint-Geniez M., Maldonado A.E., D'Amore P.A. Vascular endothelial growth factor localization in the adult. *Am. J. Pathol.* 2006;168(2):639-648. DOI 10.2353/ajpath.2006.050834.
- Mannucci P.M. Desmopressin (DDAVP) in the treatment of bleeding disorders: the first 20 years. *Blood.* 1997;90(7):2515-2521.
- Mayer U., Poschl E., Gerecke D.R., Wagman D.W., Burgeson R.E., Timpl R. Low nidogen affinity of laminin-5 can be attributed to two serine residues in EGF-like motif gamma 2III4. *FEBS Lett.* 1995;365(2-3):129-132.
- McCarthy J.B., El-Ashry D., Turley E.A. Hyaluronan, cancer-associated fibroblasts and the tumor microenvironment in malignant progression. *Front. Cell Dev. Biol.* 2018;6:48. DOI 10.3389/fcell.2018.00048.
- Nedvetsky P.I., Tamma G., Beulshausen S., Valenti G., Rosenthal W., Klusmann E. Regulation of aquaporin-2 trafficking. Ed. E. Beitz. *Aquaporins. Handbook of Experimental Pharmacology.* Berlin; Heidelberg: Springer-Verlag, 2009;133-157.
- Nguyen N., Kumar A., Chacko S., Ouellette R.J., Ghosh A. Human hyaluronan synthase-1 promotes malignant transformation via epithelial-to-mesenchymal transition, micronucleation and centrosome abnormalities. *Cell Commun. Signal.* 2017;15:48. DOI 10.1186/s12964-017-0204-z.

- Ostrowski N.L., Young W.S. Expression of vasopressin V1a and V2 receptor messenger ribonucleic acid in the liver and kidney of embryonic, developing, and adult rats. *Endocrinology*. 1993;133(4):1849-1859. DOI 10.1210/endo.133.4.8404628.
- Pequeux C., Keegan B.P., Hagelstein M.T., Geenen V., Legros J.J., North W.G. Oxytocin- and vasopressin-induced growth of human small-cell lung cancer is mediated by the mitogen-activated protein kinase pathway. *Endocr. Relat. Cancer*. 2004;11(4):871-885. DOI 10.1677/erc.1.00803.
- Privratsky J.R., Paddock C.M., Florey O., Newman D.K., Muller W.A., Newman P.J. Relative contribution of PECAM-1 adhesion and signaling to the maintenance of vascular integrity. *J. Cell Sci*. 2011; 124(9):1477-1485. DOI 10.1242/jcs.082271.
- Rendu F., Brohard-Bohn B. The platelet release reaction: granules' constituents, secretion and functions. *Platelets*. 2001;12(5):261-273. DOI 10.1080/09537100120068170.
- Rolli M., Fransvea E., Pilch J., Saven A., Felding-Habermann B. Activated integrin $\alpha v \beta 3$ cooperates with metalloproteinase MMP-9 in regulating migration of metastatic breast cancer cells. *Proc. Natl. Acad. Sci. USA*. 2003;100(16):9482-9487. DOI 10.1073/pnas.1633689100.
- Sackstein R. The biology of CD44 and HCELL in hematopoiesis: the 'step 2-bypass pathway' and other emerging perspectives. *Curr. Opin. Hematol*. 2011;18(4):239-248. DOI 10.1097/MOH.0b013e3283476140.
- Senbanjo L.T., Chellaiah M.A. CD44: A multifunctional cell surface adhesion receptor is a regulator of progression and metastasis of cancer cells. *Front. Cell Dev. Biol*. 2017;5:18. DOI 10.3389/fcell.2017.00018.
- Sequeira Lopez M.L., Chernavsky D.R., Nomasa T., Wall L., Yanagisawa M., Gomez R.A. The embryo makes red blood cell progenitors in every tissue simultaneously with blood vessel morphogenesis. *Am. J. Physiol. Regul. Integr. Comp. Physiol*. 2003;284(4):R1126-R1137. DOI 10.1152/ajpregu.00543.2002.
- Sharova N.P., Melnikova V.I., Khagai I.I., Karpova Y.D., Dmitrieva S.V., Astakhova T.M., Afanas'eva M.A., Popova N.A., Ivanova L.N., Zakharova L.A. Pattern of proteasome expression in Walker 256 tumor cells after their transplantation into the Brattleboro rats with genetic defect of vasopressin synthesis. *Doklady Biochemistry and Biophysics*. 2008;419(1):93-97. DOI 10.1134/S1607672908020129.
- Shibuya M. Vascular endothelial growth factor and its receptor system: physiological functions in angiogenesis and pathological roles in various diseases. *J. Biochem*. 2013;153(1):13-19. DOI 10.1093/jb/mvs136.
- Siiskonen H., Poukka M., Tyynela-Korhonen K., Sironen R., Pasonen-Seppanen S. Inverse expression of hyaluronidase 2 and hyaluronan synthases 1-3 is associated with reduced hyaluronan content in malignant cutaneous melanoma. *BMC Cancer*. 2013;13:181. DOI 10.1186/1471-2407-13-181.
- Sironen R.K., Tammi M., Tammi R., Auvinen P.K., Anttila M., Kosma V.M. Hyaluronan in human malignancies. *Exp. Cell Res*. 2011; 317(4):383-391. DOI 10.1016/j.yexcr.2010.11.017.
- Slevin M., Gaffney J., Kumar S. Angiogenic oligosaccharides of hyaluronan induce multiple signaling pathways impacting vascular endothelial cell mitogenesis and wound healing. *J. Biol. Chem*. 2002;277(43):41046-41059. DOI 10.1074/jbc.M109443200.
- Stern R., Asari A.A., Sugahara K.N. Hyaluronan fragments: an information-rich system. *Eur. J. Cell Biol*. 2006;85(8):699-715. DOI 10.1016/j.ejcb.2006.05.009.
- Tahara A., Tsukada J., Tomura Y., Yatsu T., Shibasaki M. Vasopressin increases type IV collagen production through the induction of transforming growth factor-beta secretion in rat mesangial cells. *Pharmacol. Res*. 2008;57(2):142-150. DOI 10.1016/j.npep.2010.12.001.
- Thibonnier M., Coles P., Thibonnier A., Shoham M. The basic and clinical pharmacology of nonpeptide vasopressin receptor antagonists. *Annu. Rev. Pharmacol. Toxicol*. 2001;41:175-202. DOI 10.1146/annurev.pharmtox.41.1.175.
- Thibonnier M., Goraya T., Berti-Mattera L. G protein coupling of human platelet V1 vascular vasopressin receptors. *Am. J. Physiol. Cell Physiol*. 1993;264(5):1336-1344. DOI 10.1152/ajpcell.1993.264.5.C1336.
- Tian X., Azpuru J., Hine C., Vaidya A., Myakishev-Rempel M., Ablauva J., Mao Z., Nevo E., Gorbunova V., Seluanov A. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature*. 2013;499(7458):346-349. DOI 10.1038/nature12234.
- Tomasiak M., Stelmach H., Rusak T., Ciborowski M., Radziwon P. Vasopressin acts on platelets to generate procoagulant activity. *Blood Coagul. Fibrinolysis*. 2008;19(7):615-624. DOI 10.1097/MBC.0b013e328309905d.
- van der Meijden P.E.J., Heemskerk J.W.M. Platelet biology and functions: new concepts and clinical perspectives. *Nat. Rev. Cardiol*. 2019;16(3):166-179. DOI 10.1038/s41569-018-0110-0.
- Vigetti D., Deleonibus S., Moretto P., Bowen T., Fischer J.W., Grandoch M., Oberhuber A., Love D.C., Hanover J.A., Cinquetti R., Karousou E., Viola M., D'Angelo M.L., Hascall V.C., De Luca G., Passi A. Natural antisense transcript for hyaluronan synthase 2 (HAS2-AS1) induces transcription of HAS2 via protein O-GlcNAcylation. *J. Biol. Chem*. 2014;289(42):28816-28826. DOI 10.1074/jbc.M114.597401.
- Voelcker V., Gebhardt C., Averbeck M., Saalbach A., Wolf V., Weih F., Sleeman J., Anderegg U., Simon J. Hyaluronan fragments induce cytokine and metalloprotease upregulation in human melanoma cells in part by signalling via TLR4. *Exp. Dermatol*. 2008;17(2):100-107. DOI 10.1111/j.1600-0625.2007.00638.x.
- Wang Z., Zhao K., Hackert T., Zuller M. CD44/CD44v6 a reliable companion in cancer-initiating cell maintenance and tumor progression. *Front. Cell Dev. Biol*. 2018;6:97. DOI 10.3389/fcell.2018.00097.
- Weber M.R., Zuka M., Lorgier M., Tschan M., Torbett B.E., Zijlstra A., Quigley J.P., Staffin K., Eliceiri B.P., Krueger J.S., Marchese P., Ruggeri Z.M., Brunhilde H., Felding B.H. Activated tumor cell integrin $\alpha v \beta 3$ cooperates with platelets to promote extravasation and metastasis from the blood stream. *Thromb. Res*. 2016;140(Suppl. 1):S27-S36. DOI 10.1016/S0049-3848(16)30095-0.
- Williams K., Motiani K., Giridhar P.V., Kasper S. CD44 integrates signaling in normal stem cell, cancer stem cell and (pre)metastatic niches. *Exp. Biol. Med. (Maywood)*. 2013;238(3):324-338. DOI 10.1177/1535370213480714.
- Wojtukiewicz M.Z., Sierko E., Hempel D., Tucker S.C., Honn K.V. Platelets and cancer angiogenesis nexus. *Cancer Metastasis Rev*. 2017;36(2):249-262. DOI 10.1007/s10555-017-9673-1.
- Yang C., Cao M., Liu H., He Y., Xu J., Du Y., Liu Y., Wang W., Cui L., Hu J., Gao F. The high and low molecular weight forms of hyaluronan have distinct effects on CD44 clustering. *J. Biol. Chem*. 2012; 287(51):43094-43107. DOI 10.1074/jbc.M112.349209.
- Zachary I., Rozengurt E. Focal adhesion kinase (p125FAK): a point of convergence in the action of neuropeptides, integrins, and oncogenes. *Cell*. 1992;71(6):891-894.
- Zhang J., Zhang F., Niu R. Functions of Shp2 in cancer. *J. Cell Mol. Med*. 2015;19(9):2075-2083. DOI 10.1111/jcmm.12618.
- Zhao T.M., Newman P.J. Integrin activation by regulated dimerization and oligomerization of platelet endothelial cell adhesion molecule (PECAM)-1 from within the cell. *J. Cell Biol*. 2001;152(1):65-73.
- Zhu J.X., Cao G., Williams J.T., Delisser H.M. SHP-2 phosphatase activity is required for PECAM-1-dependent cell motility. *Am. J. Physiol. Cell Physiol*. 2010;299(4):C854-C865. DOI 10.1152/ajpcell.00436.2009.

ORCID ID

I.I. Khgaya orcid.org/0000-0003-0545-4311

Acknowledgements. This work was supported by State Budgeted Project 0324-2019-0041.

Conflict of interest. The author declares no conflict of interest.

Received March 11, 2019. Revised April 4, 2019. Accepted April 5, 2019.

Anxiety and neurometabolite levels in the hippocampus and amygdala after prolonged exposure to predator-scent stress

O.B. Shevelev¹✉, V.E. Tseilikman², N.V. Khotskin¹, A.S. Khotskina¹, G.V. Kontsevaya¹, M.S. Lapshin², M.P. Moshkin¹, M.V. Komelkova², I.V. Feklicheva², O.B. Tseilikman², E.B. Manukhina^{2, 3, 4}, H.F. Downey⁴, E.L. Zavjalov¹

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

² South Ural State University (National Research University), Chelyabinsk, Russia

³ Institute of General Pathology and Pathophysiology, Moscow, Russia

⁴ Department of Physiology and Anatomy, University of North Texas Health Science Center, Fort Worth, Texas, USA

✉ e-mail: shevelev.oleg.nsk@gmail.com

Here, to study the relationship between anxiety levels with changes in the neurometabolic profile in the hippocampus and amygdala, an experimental predator stress model was reproduced in which Sprague-Dawley rats were exposed to cat urine for 10 minutes on a daily basis for 10 days. At the time of presentation of the stimulus, an on-line survey of behavioral reactions was conducted. Fear, aggressiveness, avoidance of stimulus and grooming were recorded. Fourteen days after the completion of the last stress exposure, the total level of anxiety was determined in the test of the "cross maze". Using the method of *in vivo* NMR spectroscopy, the content of neurometabolites was determined in the hippocampus and in the amygdala. According to the peculiarities of behavioral reactions to a stressor, animals were retrospectively divided into two phenotypes. The first phenotype used a passive behavioral strategy, and the second phenotype was active. In animals of the first phenotype, the indicators of anxiety behavior remained at the control level. In animals of the second phenotype, a decrease in anxiety was observed. Animals of the second phenotype showed elevated levels of lactate in the hippocampus compared to animals of the first phenotype, and the lowest N-acetylaspartate levels significantly differed from those in the control and the first phenotype animals. In the amygdala, in animals of the second phenotype, the content of taurine is sharply reduced in comparison with those in the control and the animals of the first phenotype. Thus, the results obtained indicate a relationship of post-stress changes in anxiety, with the peculiarities of the behavioral reactions presented at the moment of the immediate action of the stressor. Among the hippocampal and amygdala neurometabolites, the most informative for the characterization of the anxiolytic action of the predator stress are identified.

Key words: predator scent stress; neurometabolites; amygdala; hippocampus; anxiety index.

For citation: Shevelev O.B., Tseilikman V.E., Khotskin N.V., Khotskina A.S., Kontsevaya G.V., Lapshin M.S., Moshkin M.P., Komelkova M.V., Feklicheva I.V., Tseilikman O.B., Manukhina E.B., Downey H.F., Zavjalov E.L. Anxiety and neurometabolite levels in the hippocampus and amygdala after prolonged exposure to predator-scent stress. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5):582-587. DOI 10.18699/VJ19.528

Уровень тревожности и содержание нейрометаболитов в гиппокампе и амигдале крыс после завершения хронического предаторного стресса

О.Б. Шевелев¹✉, В.Э. Цейликман², Н.В. Хоцкин¹, А.С. Хоцкина¹, Г.В. Концевая¹, М.С. Лапшин², М.П. Мошкин¹, М.В. Комелькова², И.В. Фекличева², О.Б. Цейликман², Е.Б. Манухина^{2, 3, 4}, Г.Ф. Дауни⁴, Е.Л. Завьялов¹

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

² Южно-Уральский государственный университет (национальный исследовательский университет), Челябинск, Россия

³ Научно-исследовательский институт общей патологии и патофизиологии, Москва, Россия

⁴ Центр медицинских наук Университета Северного Техаса, Форт-Уэрт, Техас, США

✉ e-mail: shevelev.oleg.nsk@gmail.com

В этом исследовании для изучения соотношения между уровнем тревожности с изменениями нейрометаболического профиля в гиппокампе и амигдале воспроизводилась экспериментальная модель предаторного стресса, в которой крысы линии Sprague-Dawley в течение 10 минут подвергались экспозиции кошачьей мочи на протяжении 10 дней ежедневно. В момент предъявления стимула велась съемка поведенческих реакций. Регистрировались реакции испуга, агрессивности, избегания стимула и груминга. Через 14 дней после завершения последнего стрессорного воздействия в тесте «крестообразный лабиринт» определялся общий уровень тревожности. С помощью метода прижизненной ЯМР-спектроскопии определяли содержание нейрометаболитов в гиппокампе и амигдале. По особенностям поведенческих реакций на стрессор животные были ретроспективно разделены на два фенотипа. Первый фенотип использовал пассивную поведенческую стратегию, а второй – активную. У животных первого фенотипа показатели тревожного поведения сохранялись на контрольном уровне, в то время как у животных второго фенотипа наблюдалось снижение

этого показателя. В гиппокампе у животных второго фенотипа отмечалось повышенное содержание лактата по сравнению с животными первого фенотипа, тогда как уровень N-ацетиласпартата имел самые низкие значения, отличные от таковых у животных двух других групп. В амигдале у животных второго фенотипа было снижено содержание таурина по сравнению с животными первого фенотипа и контрольной группы. Таким образом, полученные результаты свидетельствуют о связи постстрессорных изменений тревожности с особенностями поведенческих реакций, выявленных в момент непосредственного действия стрессора. Среди нейрометаболитов гиппокампа и амигдалы определены наиболее информативные для характеристики анксиолитического действия предаторного стресса.

Ключевые слова: предаторный стресс; нейрометаболиты; амигдала; гиппокамп; индекс тревожности.

Introduction

Different kinds of predator stress such as visual contact with a predator and exposure to its scent are considered as experimental models to study post-war stress syndrome and post-traumatic stress disorder (PTSD) (Cohen et al., 2008). What is interesting about this disorder is that PTSD develops electively and a lot of people who have been exposed to severe psychological stress are resistant to PTSD (Pitman et al., 2012). Another interesting neurobiological phenomenon that has been discovered lately, is called post-traumatic growth (PTG) syndrome. This syndrome also develops after a psychological trauma in a relatively small group of patients (Wong et al., 2018). So far, the biological mechanisms making people resistant to both PTSD and PTG have remained unknown.

Developing a proper experimental model remains another crucial issue for studying PTSD. Such models are required to be relevant to the diagnostic criteria (Cohen, Zohar, 2004). According to the fifth edition of Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013) PTSD can be diagnosed after identification of a stressor that triggers the disorder. Since PTSD is a long-term consequence of the stressor's effect, researchers have not paid much attention to a patient's behavioral response to emotional distress when resistance to PTSD is being formed.

Despite the long-term history of investigation into the role of behavioral adaptation strategies in forming response to stressors, PTSD has rarely been studied in this respect. Neurobiology knows two adaptation strategies. The first one is aimed at active overcoming of a stressing situation and is called a stimulus-response or S-R strategy (Schwabe et al., 2008; De Kloet, 2012). The other, adaptive strategy aimed at getting accustomed to a stressor is often called a spatial or S strategy (Schwabe et al., 2008). It is apparent that the S-R strategy is necessary in a situation when a stressor can be surmounted by fight or run, while the S strategy is useful in the presence of a continuous stressor. Following the S-R strategy activates the hippocampus-striatum axis, while transmitting neural signals from the hippocampus to amygdala launches the S strategy (Schwabe et al., 2008).

Most of the studies dealing with the metabolism of these key limbic structures under chronic stress have been conducted *in vitro*, while modern techniques of NMR spectroscopy enable one to obtain *in vivo* estimations of the hippocampus and amygdala's metabolic profile. Unfortunately, these have been just a few studies based on the NMR spectroscopy data obtained under chronic stress. Earlier it has been demonstrated (Hemanth Kumar et al., 2012) that chronic stress increases

the proportion of some metabolites (myoinositol, taurine) and decreases the proportion of others (N-acetylaspertate, glutamate, glutamine, gamma-aminobutyric acid (GABA)) in the hippocampus. In this study, the changes were measured right after the end of a stressor effect. Unfortunately, no studies investigating the long-term effects of chronic stress on hippocampus and amygdala NMR spectroscopy data have been available so far.

In the presented study we investigated the immediate reaction to a stressor in animals under predator stress. During the investigation, we, first hand, tried to understand whether it was possible to distinguish different behavioral phenotypes with different behavioral strategies in immediate response to a stressor based on the animals' level of anxiety. For that purpose, the determined phenotypes were compared to match their anxiety levels to changes in their hippocampus and amygdala metabolic profiles. These limbic structures had been selected due to their crucial role in PTSD neurobiology in general and in PTSD anxiety symptomatics in particular.

Materials and methods

Animal model. The presented study was performed during a spring period using the equipment provided by Center for Genetic Resources of Laboratory Animals, Institute of Cytology and Genetics, SB RAS (RFMEFI61914X0005, RFMEFI62114X0010). The cohort was 26 Sprague-Dawley male rats in the age of 8–9 weeks by the time the experiments began. The animals were kept in pairs in the OptiRAT ventilated cages (Animal Care, USA). They had free access to water and granulated food (Ssniff, Germany). The cage's photoperiod was 14 light:10 dark, temperature – 24 ± 2 °C and humidity – 40–50 %. As the nesting material, the cages were inlaid with dedusted wood shavings. Both the food and shavings were autoclaved before use. To water the animals, one used the water cleansed in a Millipore purifier and enriched with the Severyanka mineral additive (ECOPROJECT, Saint-Petersburg). All animal experiments were performed in compliance with the protocols and recommendations for the proper use and care of laboratory animals (ECC Directive 86/609/EEC). All efforts were made to minimize the number of animals used and their pain or discomfort.

Research protocol. The animals were divided into 2 subgroups: control ($n = 7$) and the rats undergoing predator stress to model PTSD ($n = 19$). To model the stress the modified Cohen & Zohar model (2004) that was previously described in (Tseilikman et al., 2017) was applied. The experimental PTSD model was reproduced through daily exposition of the rats to the urinary marks of a mature domestic cat, deposited on

softwood cat litter. To do so, Petri dishes with 20 g of marked litter, covered with nylon fiber were placed in the cages for 15 minutes a day during 10 days. The rats from the control group also had the litter in their cages but without urinary marks. During the next 10 days, the animals experienced no additional exposure, and on day 23 their emotional reactions were estimated using an elevated cross maze (ECM). On day 26 neurometabolite content in their hippocampus and amygdala was estimated. On day 29 the animals were sacrificed.

The rats' behavioral activity. The number of behavioral acts as a response to the stimulus was counted during being exposed to predator stress. Such reactions as fear (a rat stops when passing on feline scent), grooming, explorative reactions (a rat sniffs a Petri dish with urine-marked litter). Other reactions taken into account included evading (a rat digs the dish into the nesting material); fearlessness (a rat jumps onto the dish); aggression (a rat attacks the dish trying to tear the nylon fiber covering it).

To verify their anxiety, the rats were given an ECM test. The cross maze is a central area of 10 by 10 cm in size with 4 arms of 50 × 10 cm cross-stretching from it, two are placed opposite one another and have no closing walls (open arms) and two other – equipped with closing walls of 50 cm in height (closed arms). The maze was set at 1 m above the floor. The individual behavior of each rat was studied in the maze for 10 min. As an integral value characterizing the presence of anxious disorders in the rats, anxiety index (AI) was used that was calculated based on the following formula (Cohen et al., 2008):

$$AI = 1 - \left[\frac{\text{time in open arms}}{\sum \text{time in maze}} + \frac{\text{enters into open arms}}{\sum \text{number of enters}} \right] / 2.$$

NMR spectroscopy of the hippocampus and amygdala.

The content of neurometabolites in the rats' hippocampus and amygdala was studied using a Bruker BioSpec 117/16 USR horizontal tomograph with a magnetic field of 11.7 T. Five minutes prior to the study, the rats were immobilized with gas anesthesia (Isofluran; Baxter Healthcare Corp., Deerfield, IL) using a Univentor 400 Anesthesia Unit (Univentor, Zejtun, Malta). The body temperature of the animals was supported using the water loop of the tomograph's table that warmed its surface to 30 °C. A pneumatic respiration sensor to control the depth of anaesthesia (SA Instruments, Stony Brook, N.Y., USA) was placed underneath each rat.

¹H MRS. All proton spectra of the rats' brains were obtained using transmitting volume (T11232V3) and receiving surface (T11425V3) ¹H RF coils. For proper positioning of the spectroscopic voxels, whose size was 3.0 × 1.5 × 3.0 mm for the amygdala and 1.5 × 3.0 × 3.0 mm – for the hippocampus, TurboRARE, a spin echo technique, was applied (TE = 11 ms, TR = 2.5 ms) to obtain T2-weighted high-resolution images (slice thickness – 0.5 mm, filed of view – 2.5×2.5 cm, matrix size – 256×256 pixels) in three projections. All the voxels were positioned manually based on the images obtained. All the proton spectra were obtained using STEAM (stimulated echo acquisition mode spectroscopy; TE = 3 ms, TR = 5 s, and

120 accumulations). Before each spectroscopic measurement, the magnetic field heterogeneity was adjusted for a selected voxel using FastMap (Gruetter, 1993). The water signal in the spectra was suppressed using variable pulse power and optimized relaxation delays (VAPOR) (Tkáč et al., 1999).

¹H spectra processing. To process the experimental spectra and determine the qualitative content of metabolites, a specially designed software solution was used. As LC Model (Provencher, 1993) the solution assumes that the spectrum of a mixture of known chemical compounds is a linear combination of the spectra of analyzed components. More details about the package can be found in (Moshkin et al., 2014). The percentage content of the studied metabolites was estimated in relation to their total amount. In total, 12 metabolites were studied (N-acetylaspartate, GABA, alanine, aspartate, total choline-containing compounds, total creatine + creatine P, total glutamine + glutamate, myoinositol, taurine, glycine, lactate, phosphorylethanolamine), which allowed for integral estimation of neuronal activity in the rats (Hemanth Kumar et al., 2012; Gulevich et al., 2015; Shevelev et al., 2018) for this set reflects the balance between inhibitory (GABA) and excitatory (glutamine + glutamate) neurotransmitters; contains a neural viability marker (N-acetylaspartate), key amino acids (alanine and aspartate), moderators of neurotransmission (taurine) and phosphate metabolism (creatine + creatine P), and reflects membrane synthesis and cell proliferation (choline, myoinositol and phosphorylethanolamine).

Statistical methods. To process the obtained data, a one-way ANOVA was applied. The significance of the differences between the phenotypes of the studied animals was estimated with a post hoc test (LSD Fisher). The values of the studied parameters were presented as mean ± standard error of mean (M ± SEM).

Results

Based on their behavioral reactions to the chronic effect of predator stress, the studied rats were separated into two phenotypes. Phenotype 1 (*n* = 13) exposed a passive defense reaction in which fear and grooming dominated. Phenotype 2 was characterized by active behavioral reactions as absence of fear, aggression towards the stimulus, and searching. On day 13 after the stress exposure was ended (Table 1), statistically significant differences were observed for the time of being in ECM center ($F_{2,24} = 5.89, p = 0.008$). The post hoc analysis demonstrated statistically significant differences between Phenotype 2 and the control group ($p = 0.003$) as well as between the animals belonging to different phenotypes ($p = 0.011$).

In Phenotype 2 the time of being in ECM center was higher than that in the control rats and Phenotype 1. Moreover, statistically significant differences were observed for the times of being in the open arms ($F_{2,24} = 22.32, p < 0.001$), which was higher for Phenotype 2 than for the control group ($p < 0.001$) and Phenotype 1 ($p < 0.001$). In addition, statistically significant differences were observed in the number of the times the animals entered open arms ($F_{2,24} = 11.56, p = 0.003$). The rats of Phenotype 2 entered them more often than the control rats ($p < 0.001$) and those belonging to Phenotype 1 ($p < 0.001$).

When studying the metabolite level in the hippocampus (Table 2), significant differences were observed in N-acetyl-

Table 1. Anxiety indexes in cross-maze test in the rats with different behavioral response to stressor (M ± SEM)

Anxiety indexes	Control (n = 7)	Phenotype 1 (n = 13)	Phenotype 2 (n = 6)
ECM center, % of test time	0.10 ± 0.02	0.12 ± 0.02	0.19 ± 0.01*#
Closed arms, % of test time	0.86 ± 0.03	0.83 ± 0.03	0.62 ± 0.03*#
Open arms, % of test time	0.04 ± 0.01	0.04 ± 0.01	0.19 ± 0.03*#
Searching in the closed arms, min	0.92 ± 0.06	0.94 ± 0.04	0.98 ± 0.01
Searching in the open arms, min	0.35 ± 0.06	0.37 ± 0.03	0.70 ± 0.06*#
Number of enters into the open arms	0.63 ± 0.50	0.69 ± 0.21	4.43 ± 1.19*#
Anxiety index	0.92 ± 0.01	0.91 ± 0.01	0.88 ± 0.01*#

* $p < 0.05$ mark the differences when compared to the control; # $p < 0.05$ mark the differences between Phenotypes 1 and 2.

Table 2. Content of neurometabolites in hippocampus and amygdala of the rats with different behavioral response to stressor (M ± SEM)

Neurometabolites	Hippocampus			Amygdala		
	Control (n = 7)	Phenotype 1 (n = 13)	Phenotype 2 (n = 6)	Control (n = 7)	Phenotype 1 (n = 13)	Phenotype 2 (n = 6)
N-acetylaspartate	17.85 ± 0.67	18.22 ± 0.68	15.35 ± 0.53*#	9.62 ± 0.83	9.98 ± 1.13	9.97 ± 3.02
GABA	4.99 ± 0.68	5.98 ± 0.96	6.09 ± 1.17	2.29 ± 1.09	1.68 ± 0.51	0.92 ± 0.66
Alanine	4.74 ± 1.44	3.53 ± 0.77	5.35 ± 1.85	0.49 ± 0.26	2.28 ± 0.77	1.89 ± 1.04
Aspartate	0.09 ± 0.01	0.13 ± 0.03	0.08 ± 0.01	1.78 ± 0.93	1.89 ± 0.89	2.67 ± 1.47
Choline-containing compounds	1.78 ± 0.42	1.93 ± 0.26	1.15 ± 0.26	1.32 ± 0.48	1.45 ± 0.49	0.83 ± 0.49
Creatine + phosphocreatine	11.96 ± 1.19	12.36 ± 0.84	9.58 ± 0.84	11.63 ± 0.90	10.81 ± 0.83	7.98 ± 1.81
Glutamine + glutamate	15.07 ± 1.66	16.90 ± 1.47	13.43 ± 1.83	21.27 ± 2.39	19.81 ± 1.43	15.37 ± 4.70
Myoinositol	5.46 ± 1.64	3.83 ± 1.09	3.48 ± 1.13	11.14 ± 3.41	13.74 ± 2.44	14.11 ± 4.24
Taurine	4.92 ± 0.55	5.45 ± 0.48	4.52 ± 0.42	3.44 ± 0.52	2.96 ± 0.84	0.19 ± 0.13*#
Glycine	19.10 ± 3.63	21.31 ± 2.40	20.85 ± 1.48	12.86 ± 4.15	8.30 ± 3.57	9.84 ± 5.77
Lactate	4.44 ± 1.01	1.90 ± 0.83	6.26 ± 1.53#	1.08 ± 1.00	5.32 ± 3.35	2.35 ± 2.06
Phosphorylethanolamine	9.59 ± 3.12	8.63 ± 2.14	13.30 ± 1.74	23.09 ± 3.39	22.00 ± 2.80	33.18 ± 7.17

* $p < 0.05$ mark the differences when compared to the control; # $p < 0.05$ mark the differences between Phenotypes 1 and 2.

aspartate ($F_{2,24} = 3.74$, $p = 0.041$) and lactate ($F_{2,24} = 4.42$, $p = 0.023$) contents. It turned out that N-acetylaspartate content in Phenotype 1 was higher than that in Phenotype 2 ($p = 0.014$). However, when compared to the control, no significant differences were seen. On the other hand, the lactate content in Phenotype 2 was higher than that of Phenotype 1 ($p < 0.01$) with no statistical differences registered against the control.

When studying the metabolite level in the amygdala (see Table 2), statistically significant differences were observed only for taurine content ($F_{2,24} = 3.51$, $p = 0.049$). It was lowest in Phenotype 2 and statistically differed from both the control ($p = 0.022$) and Phenotype 1 ($p = 0.031$).

Discussion

First hand, the obtained results attest there is a link between a behavioral reaction to stress and the anxiety level registered long after the stress. It is noteworthy that in the experiment, the anxiety level in the control rats was initially high. According to other studies (Cohen, Zohar, 2004; Kondashevskaya et al., 2017) using this index, its mean value usually varies from 0.5 to 0.75, although there were those whose anxiety index exceeded 0.8. In our design, the initial level of anxiety was so high that even daily predator stress could not worsen it.

So far, the rats included in Phenotype 2 that demonstrated active behavior (S-R strategy) in response to the predator

stress had a lower level of anxiety if compared to the control group, while the rats belonging to Phenotype 1 with passive behavior (S strategy) preserved the high level of anxiety even when being no longer exposed to the stress. In other words, the anxiolytic effect of the stress was observed only in the rats using S-R strategy. Thus, the experiment has demonstrated the role of an animal's initial condition in the formation of the long-term effects of chronic predator stress.

For the time being, there have only been a few studies investigating changes in the level of brain metabolites in animal models in the presence of chronic stress. The discovered difference in the behavioral reactions of rats to predator stress has found their reflection in metabolic changes in certain structures of the brain. Our results obtained for certain hippocampal metabolites have demonstrated that N-acetylaspartate responsible for the general viability of neurons (Moffett et al., 2007; Shevelev et al., 2018) reduces in Phenotype 2 (S-R strategy) to lower extend if compared both to Phenotype 1 (S strategy) and the control rats. These results are comparable to the ones obtained by Hemanth Kumar et al. (2012), who studied chronic stress in a rat model and observed reduction of the same metabolite in the hippocampus. Similar reduction of N-acetylaspartate level was reported for the hippocampus of stressed macaques (Coplan et al., 2010). Hemanth Kumar et al. (2012) also demonstrated that in presence of chronic stress, rats had increased hippocampal taurine levels. In our study, a similar tendency has been demonstrated by the rats belonging to Phenotype 1 if compared to the control group, but these differences have not been statistically significant.

Another interesting peculiarity of the rats using S-R strategy has been a higher level of lactate if compared to the other experimental group. Meanwhile, being exposed to the stimulus only active (S-R) rats have shown signs of aggression. Opposite data were obtained for the grey rats selected for their aggressive defense reaction to humans. As it turned out, domestic rats had a higher level of hippocampal lactate if compared to these aggressive rats (Gulevich et al., 2015).

Among the studied neurometabolites of the amygdala, only taurine has been sensitive to predator stress. Its content has reduced in the active rats of Phenotype 2 if compared both to Phenotype 1 and the control group. In this respect, it is important to note that taurine modulates neurosynaptic transmission (McCool, Chappell, 2007), so its lower level may be indirect evidence of reduced amygdala activity in Phenotype 2. And, possibly, this reduction is one of the chains of the mechanism producing the anxiolytic effect in this predator stress model.

Conclusion

Finding this effect in the rats using S-R strategy while stressed has been the main scientific novelty of the performed study, in which we have been able to detect informative neurometabolites in the rats demonstrating different behavioral reactions under stress. If one assumes there is a link between this experimental reduction of anxiety in rats and post-traumatic stress disorder in humans, our experimental protocol can be used for future development and validation of a model that characterizes this underinvestigated consequence of psychological trauma.

References

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington, V.A.: Amer. Psychiatric Publ., 2013.
- Cohen H., Matar M.A., Buskila D., Kaplan Z., Zohar J. Early poststressor intervention with high-dose corticosterone attenuates posttraumatic stress response in an animal model of posttraumatic stress disorder. *Biol. Psychiatry*. 2008;64:708-717. DOI 10.1016/j.biopsych.2008.05.025.
- Cohen H., Zohar J. An animal model of posttraumatic stress disorder: the use of cut-off behavioral criteria. *Ann. NY Acad. Sci.* 2004;1032:167-178. DOI 10.1196/annals.1314.014.
- Coplan J.D., Mathew S.J., Abdallah C.G., Mao X., Kral J.G., Smith E.L., Rosenblum L.A., Perera T.D., Dwork A.J., Hof P.R., Gorman J.M., Shungu D.C. Early-life stress and neurometabolites of the hippocampus. *Brain Res.* 2010;1358:191-199. DOI 10.1016/j.brainres.2010.08.021.
- De Kloet E.R. From receptor balance to rational glucocorticoid therapy. *Endocrinology*. 2014;155(8):2754-2769. DOI 10.1210/en.2014-1048.
- Gruetter R. Automatic, localized *in vivo* adjustment of all first- and second-order shim coils. *Magn. Reson. Med.* 1993;29:804-811.
- Gulevich R.G., Akulov A.E., Shikhevich S.G., Kozhemyakina R.V., Plyusnina I.Z. Proton magnetic resonance spectroscopy of neurometabolites in the hippocampi of aggressive and tame male rats. *Russ. J. Genet.: Appl. Res.* 2016;6(4):430-436. DOI 10.1134/S2079059716040079.
- Hemanth Kumar B.S., Mishra S.K., Rana P., Singh S., Khushu S. Neurodegenerative evidences during early onset of depression in CMS rats as detected by proton magnetic resonance spectroscopy at 7T. *Behav. Brain Res.* 2012;232(1):53-59. DOI 10.1016/j.bbr.2012.03.011.
- Kondashevskaya M.V., Tseilikman V.E., Manukhina E.B., Downey H.F., Komelkova M.V., Lapshin M.S., Samoylov E.A., Popkov P.N., Aliluev A.V., Vasileva M.V., Kurganov A.S., Maltseva N.V., Tseilikman O.B. Disorder in the morphology and function of adrenal glands in experimental post-traumatic stress disorder in rats: correlation with behavioral markers. *Rossiyskiy Fiziologicheskii Zhurnal im. I.M. Sechenova = I.M. Sechenov Physiological Journal*. 2017;103(7):808-818. (in Russian)
- McCool B.A., Chappell A. Strychnine and taurine modulation of amygdala-associated anxiety-like behavior is 'state' dependent. *Behav. Brain Res.* 2007;178(1):70-81. DOI 10.1016/j.bbr.2006.12.002.
- Moffett J.R., Ross B., Arun P., Madhavarao C.N., Nambodiri A. N-Acetylaspartate in the CNS: from neurodiagnostics to neurobiology. *Prog. Neurobiol.* 2007;81:89-131. DOI 10.1016/j.pneurobio.2006.12.003.
- Moshkin M.P., Akulov A.E., Petrovski D.V., Saik O.V., Petrovsky E.D., Savelov A.A., Koptuyug I.V. Proton magnetic resonance spectroscopy of brain metabolic shifts induced by acute administration of 2-deoxy-D-glucose and lipopolisaccharides. *NMR Biomed.* 2014;27(4):399-405. DOI 10.1002/nbm.3074.
- Pitman R., Rasmusson A., Koenen K., Shin L., Orr S., Gilbertson M., Milad M., Liberzon I. Biological studies of post-traumatic stress disorder. *Nat. Rev. Neurosci.* 2012;13:769-787. DOI 10.1038/nrn3339.
- Provencher S.W. Estimation of metabolite concentrations from localized *in vivo* proton NMR spectra. *Magn. Reson. Med.* 1993;30(6):672-679.
- Schwabe L., Dalm S., Schachinger H., Oitzl M.S. Chronic stress modulates the use of spatial and stimulus-response learning strategies in mice and man. *Neurobiol. Learn. Mem.* 2008;90(3):495-503. DOI 10.1016/j.nlm.2008.07.015.
- Shevelev O.B., Seryapina A.A., Zavjalov E.L., Gerlinskaya L.A., Goryachkovskaya T.N., Slynko N.M., Kuibida L.V., Peltek S.E., Markel A.L., Moshkin M.P. Hypotensive and neurometabolic effects of intragastric Reishi (*Ganoderma lucidum*) administration in hypertensive ISIAH rat strain. *Phytomedicine*. 2018;41:1-6. DOI 10.1016/j.phymed.2018.01.013.

Tkáč I., Starčuk Z., Choi I.-Y., Gruetter R. *In vivo* ^1H NMR spectroscopy of rat brain at 1 ms echo time. *Magn. Reson. Med.* 1999;41:649-656.

Tseilikman V.E., Shevelev O.B., Khotskin N.V., Dotsenko A.S., Kontsevaya G.V., Lapshin M.S., Moshkin M.P., Komelkova M.V., Feklicheva I.V., Tseilikman O.B., Dremencov E., Zavjalov E.L. Magnetic resonance spectroscopy of hippocampal and striatal neuro-

metabolites in experimental PTSD rat modeling. *Vavilovskii Zhurnal Genetiki i Selekcii* = Vavilov Journal of Genetics and Breeding. 2017;21(7):783-787. DOI 10.18699/VJ17.293. (in Russian)

Wong A., Lee H.S., Lee H.P., Choi Y.K., Lee J.H. Posttraumatic stress disorder symptoms and posttraumatic growth following indirect trauma from the Sewol ferry disaster, 2014. *Psychiatry Investig.* 2018;15(6):613-619. DOI 10.30773/pi.2017.12.03.

ORCID ID

O.B. Shevelev orcid.org/0000-0003-3200-958X

E.L. Zavjalov orcid.org/0000-0002-9412-3874

Acknowledgements. The animals and behavioral testing are supported by the budget project (No. 0324-2019-0041). The MRI study is supported by the budget project (No. 0259-2019-0004). All studies are implemented using the equipment of Center for Genetic Resources of Laboratory Animals at ICG SB RAS, supported by the Ministry of Education and Science of Russia (Unique ID# of the project: RFMEFI62117X0015).

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

Received March 4, 2019. Revised May 8, 2019. Accepted May 12, 2019.

Истоки митохондриального генофонда русских по результатам анализа современных и палеогеномных данных

Б.А. Мальярчук

Институт биологических проблем Севера Дальневосточного отделения Российской академии наук, Магадан, Россия
✉ e-mail: malyarchuk@ibpn.ru

Палеогеномные исследования последних лет показали, что на формирование генетического облика современных европейцев большое влияние оказали миграции населения Понто-Каспийских степей эпохи бронзового века с востока на запад Европы. Результаты исследований изменчивости митохондриальных геномов у современного русского населения Восточной Европы также позволили выявить рост эффективной численности популяций в эпоху бронзового века, что, по всей видимости, могло быть связано с миграционными процессами этого времени. В настоящей работе проанализированы данные об изменчивости целых митохондриальных геномов у современного русского населения в сравнении с распределением гаплогрупп мтДНК у древнего населения Европы и Кавказа эпохи неолита и бронзового века. Установлено, что формирование современной структуры митохондриального генофонда русских началось примерно 4 тыс. лет до н. э. в связи с притоком на восток Европы гаплотипов мтДНК, характерных для населения Центральной и Западной Европы. Предполагается, что миграции древнего населения Понто-Каспийских степей в западном направлении привели к формированию в Центральной Европе смешанных популяций, характеризующихся свойственными для западных и центральных европейцев митохондриальными гаплогруппами H, J, T, K, W. Дальнейшая экспансия этих популяций на восток Европы и далее в Азию объясняет появление у восточных европейцев новых черт митохондриального генофонда. Приведены результаты филогеографического анализа, показывающего, что особенности географического распределения подгрупп митохондриальной гаплогруппы R1a в Европе – это отражение «кавказского» компонента, появившегося в генофондах различных групп европейцев в ходе миграций бронзового века. О миграциях древних восточных европейцев в Азию (юг Сибири и Индийский субконтинент) свидетельствуют результаты филогеографического анализа митохондриальных гаплогрупп U2e2a1d, U4d2, N1a1a1a1, H2b, H8b1.

Ключевые слова: митохондриальный геном; популяции человека; Восточная Европа; палеогеномика; формирование генофонда.

Для цитирования: Мальярчук Б.А. Истоки митохондриального генофонда русских по результатам анализа современных и палеогеномных данных. Вавиловский журнал генетики и селекции. 2019;23(5):588-593. DOI 10.18699/VJ19.529

Sources of the mitochondrial gene pool of Russians by the results of analysis of modern and paleogenomic data

B.A. Malyarchuk

Institute of Biological Problems of the North, FEB RAS, Magadan, Russia
✉ e-mail: malyarchuk@ibpn.ru

Paleogenomic studies of recent years have shown that the Bronze Age migrations of populations of the Ponto-Caspian steppes from the east to the west of Europe had a great influence on the formation of the genetic makeup of modern Europeans. The results of studies of the variability of mitochondrial genomes in the modern Russian populations of Eastern Europe also made it possible to identify an increase in the effective population size during the Bronze Age, which, apparently, could be related to the migration processes of this time. This paper presents the results of analysis of data on the variability of entire mitochondrial genomes in the modern Russian populations in comparison with the distribution of mtDNA haplogroups in the ancient populations of Europe and the Caucasus of the Neolithic and Bronze Age. It was shown that the formation of the modern appearance of the Russian mitochondrial gene pool began approximately 4 thousand years B.C. due to the influx of mtDNA haplotypes characteristic of the population of Central and Western Europe to the east of Europe. It is assumed that the migrations of the ancient populations of the Ponto-Caspian steppes in the western direction led to the formation of mixed populations in Central Europe, bearing mitochondrial haplogroups H, J, T, K, W characteristic of Western and Central Europeans. Further expansion of these populations to the east of Europe and further to Asia explains the emergence of new features of the mitochondrial gene pool in Eastern Europeans. The results of a phylogeographic analysis are also presented, showing that the features of the geographical distribution of the subgroups of the mitochondrial haplogroup R1a in Europe are a reflection of the "Caucasian" component that appeared in the gene pools of various

groups of Europeans during the migration of the Bronze Age. The results of phylogeographic analysis of mitochondrial haplogroups U2e2a1d, U4d2, N1a1a1a1, H2b, and H8b1 testify to the migrations of ancient Eastern Europeans to Asia – the south of Siberia and the Indian subcontinent.

Key words: mitochondrial gene pool; human populations; Eastern Europe; paleogenomics; gene pool formation.

For citation: Malyarchuk B.A. Sources of the mitochondrial gene pool of Russians by the results of analysis of modern and paleogenomic data. *Vavilovskii Zhurnal Genetiki i Selekcii* = Vavilov Journal of Genetics and Breeding. 2019; 23(5):588-593. DOI 10.18699/VJ19.529 (in Russian)

Введение

Палеогеномные исследования нескольких последних лет показали, насколько велика роль населения степной части Восточной Европы эпохи бронзового века в формировании генетического облика современных европейцев (Allentoft et al., 2015; Haak et al., 2015; Mathieson et al., 2015). Анализ полиморфизма аутосомных локусов у древнего населения Европы продемонстрировал, что так называемый степной генетический компонент распространялся в позднем неолите/раннем бронзовом веке (4–3.2 тыс. лет до н. э.) из одного источника – понто-каспийской степной зоны – в различных направлениях: на запад – в Европу, на восток – в Центральную Азию и юг Сибири и на юг – в Индию. Предполагается, что эти события были связаны с миграциями носителей курганной культуры (Allentoft et al., 2015; Haak et al., 2015).

Между тем результаты исследований полиморфизма Y-хромосомы у древнего и современного населения Европы противоречат этой гипотезе, поскольку восточные и западные европейцы как в эпоху бронзового века, так и в настоящее время характеризуются разными подгруппами одной из самых распространенных среди западных европейцев гаплогруппы – R1b-L23 (Poznik et al., 2016). Это ставит под сомнение возможность происхождения западноевропейских R1b-L23-хромосом от вариантов, характерных для представителей ямной культуры, населявших восточноевропейские степи (Balanovsky et al., 2017). В то же время демографический анализ, основанный на данных об изменчивости протяженных участков Y-хромосом в различных европейских популяциях, показывает отчетливый сигнал популяционной экспансии в интервале 2.1–4.2 тыс. лет назад, что связано, по мнению авторов, с миграционными процессами в Европе в эпоху бронзового века (Batini et al., 2015). Подобный анализ целых митохондриальных геномов (~20 митогеномов на популяционную выборку), однако, выявил только постепенный рост численности с послеледниковой времени (Batini et al., 2017). Авторы сделали вывод о том, что миграции населения с востока Европы на запад в эпоху бронзового века нашли свое отражение только в Y-хромосоме, а митохондриальная ДНК (мтДНК) в этом отношении неинформативна.

Тем не менее изучение больших выборок митогеномов (более 250 человек на выборку) позволило, подобно исследованиям Y-хромосомы, установить рост эффективной численности населения Восточной Европы в эпоху бронзового века (Malyarchuk et al., 2017; Översti et al., 2017). Байесовский анализ изменчивости целых митогеномов 376 русских индивидуумов показал, что в Восточной Европе имел место резкий рост эффективной численности популяций ~4.3 тыс. лет тому назад (95 % достоверный

интервал 2.9–5.8 тыс. лет) (Malyarchuk et al., 2017). Это свидетельствует, что экспансия населения Восточной Европы, ставшего основой для русских популяций, началась в бронзовом веке. Результаты молекулярного датирования возраста подгрупп мтДНК, специфичных для русских и славян в целом, демонстрируют, что формирование таких подгрупп мтДНК произошло ~2.7–3.5 тыс. лет назад (Мальярчук и др., 2019). Полученные результаты указывают, таким образом, на эпоху бронзового века как на период диверсификации митохондриального генофонда предков славян и русских в частности.

Митохондриальный генофонд русского населения Восточной Европы, как и других европейских популяций, очень разнообразен, однако изучение древней ДНК показывает, что так было не всегда. В верхнем палеолите у европейцев зарегистрированы гаплогруппы U2, U5 и U8, в мезолите – главным образом гаплогруппы U2e, U4, U5a и U5b (Bramanti et al., 2009; Malmström et al., 2009; Brandt et al., 2013; Richards et al., 2016). Между тем в раннем неолите спектр гаплогрупп мтДНК у европейцев значительно расширяется за счет митохондриальных линий мигрантов-земледельцев из Ближнего Востока (H, HV, K1, J, T, N1a1a, W, X) (Haak et al., 2010; Brandt et al., 2015; Pereira et al., 2017). Сравнительный анализ митогеномов эпохи бронзового века позволил А. Juras с коллегами (2018) сделать вывод о том, что генофонды населения западной (современная Германия) и восточной (современные Чехия и Польша) частей ареала культуры боевых топоров различаются, поскольку на востоке сильнее выражен вклад восточноевропейских гаплогрупп, характерных для населения ямной культуры (U2e, U4, U5a, W3a1). По мнению авторов, эти различия обусловлены миграциями населения ямной культуры с востока на запад Европы.

Поскольку анализ изменчивости митогеномов современного русского населения Восточной Европы позволил нам выявить эпизод резкого роста эффективной численности популяций ~4.3 тыс. лет назад, то появляется вопрос о том, чем это обусловлено. Кроме того, проведенное ранее исследование фракций митогеномов, специфичных для русских или для славян в целом, также выявило рост численности ~4–5 тыс. лет назад (Мальярчук и др., 2019). Однако спектр митохондриальных гаплогрупп, представляющих эти фракции, не ограничивается восточноевропейскими гаплогруппами U2e, U4, U5a, U5b – он значительно шире и включает гаплогруппы H1, H2a, H5, H6, H11a, J1c, T1a, T2b, V1a, V3, V7, распространенные в позднем неолите в Центральной и Западной Европе. Поэтому для определения, откуда и когда могли появиться эти гаплогруппы мтДНК в генофонде предков русских, в настоящей работе проанализированы палеомитогеномные данные для населения Европы и Кавказа.

Материалы и методы

Изучена изменчивость целых митохондриальных геномов у современного русского населения Восточной Европы (Malyarchuk et al., 2017) в сравнении с распределением гаплогрупп мтДНК у древнего населения Европы и Кавказа эпохи неолита и бронзового века (Allentoft et al., 2015; Haak et al., 2015; Mathieson et al., 2018; Wang et al., 2019). Используются также сведения из базы данных AmtDB (<https://amtdb.org>), в которой собраны древние нуклеотидные последовательности мтДНК с сопутствующими данными (идентификаторы образцов, археологических сайтов, культур, гаплогруппы мтДНК и т. д.) (Ehler et al., 2019).

Для проведения филогенетического анализа нуклеотидных последовательностей митохондриальных геномов применяли метод максимальной экономии, реализованный в пакете компьютерных программ mtPhyl v4.015 (<http://eltsov.org>). При построении филогенетических деревьев не учитывали полиморфизм длины в участках мтДНК 16180–16193, 309–315, 514–524, 573–576, а также транзицию в позиции 16519. Эволюционный возраст монофилетических кластеров мтДНК (гаплогрупп и их подгрупп) оценивали с использованием двух мутационных скоростей: 1) одна нуклеотидная замена в целом митогеноме происходит за 3624 года (Soares et al., 2009) и 2) одна нуклеотидная замена в кодирующей области мтДНК совершается за 4610 лет (Perego et al., 2009).

Результаты

Анализ палеогеномных данных показывает, что наиболее отчетливые изменения митохондриального генофонда древнего населения Восточной Европы наблюдаются ~4–3 тыс. лет до н. э. До этого времени древнее население характеризуется в основном митохондриальными гаплогруппами U2e, U4a, U4b, U4d, U5a1, U5a2, U5b2, а позже появляются гаплотипы, относящиеся к гаплогруппам H, HV, J, T, W3, W6 и др.:

U2e, U4, U5	H, J, T, W и др.
6–4 тыс. лет до н. э., население современных территорий России, Украины и Латвии	
U2e, U4a, U4b, U4d, U5a1, U5a2, U5b2	H1*, H2a1, T2
4–3 тыс. лет до н. э., население современных территорий России, Украины и Латвии	
U2e, U5a1, U5a2	H*, H1b, H2b, H5a, H6a1, H13a1a1, H15b, HV, R0a1, J2b1, T1a1, T2b, T2c, U8b1, W3a1, W6c
3–2 тыс. лет до н. э., население современных территорий России, Украины, Эстонии и Литвы	
U2e, U4*, U4a, U4c, U5a1, U5a2, U5b2	H2a1, H5a1, H6a2, H13a1, J1c, J1c1, J1c2, J1c3, J2a, J2b1, T2a1, T2b, T2c, W3a1, I3a, N1a1a, K1b, U2d, X4

Таким образом, согласно палеогеномным данным, современные черты митохондриальный генофонд восточных европейцев приобрел 4–5 тысячелетий назад. В это же время, как показали предыдущие исследования из-

менчивости мтДНК современного русского населения (Malyarchuk et al., 2017), наблюдается рост эффективной численности популяций.

Для выяснения источников митохондриальных гаплогрупп, появившихся в генофонде восточных европейцев, нами проанализировано распределение гаплогрупп мтДНК у древнего населения энеолита, бронзового и железного веков соседних регионов Европы и Кавказа в сравнении со спектром гаплогрупп, наблюдаемых у современного русского населения. Выявлено, что местное восточноевропейское происхождение, кроме многочисленных вариантов гаплогрупп U2e, U4 и U5, могут иметь лишь некоторые гаплотипы, относящиеся к гаплогруппам H, H1, H2a1a, H5a, H6a1, T2, поскольку они присутствовали в генофонде восточных европейцев еще 6–4 тыс. лет до н. э. (см. выше, Приложение 1)¹. Приток гаплогрупп в генофонд восточных европейцев от населения Кавказа может быть обусловлен небольшим числом гаплогрупп (H2a1, I1, R1a, R1a1, T2a1b, U7b), а основной приток гаплогрупп связан с населением Центральной и Западной Европы (см. Приложение 1). Это гаплогруппы H1c, H3, H5, H7, H13, H28, H44, HV0, J1c1, J1c2, J1c5, J2b1, K1a2, K1a3, K1b1, K1c, N1a1a1, T1b, T2f, U3a1, U5b1, U8a, V1a, W1, X2b, которые появились в генофонде центральных и западных европейцев еще 3 тыс. лет до н. э. и ранее в результате миграций аграриев ближневосточного происхождения (анатолийских фермеров).

Для некоторых восточноевропейских гаплогрупп мтДНК трудно определить географический источник, поскольку 3 тыс. лет до н. э. и ранее они присутствовали как в Центральной Европе, так и на Кавказе (гаплогруппы J2a1, T1a2, T2e). В отношении происхождения гаплогрупп H2a1, T2c1, U2e1 и U4a2 также нет ясности – 3 тыс. лет до н. э. и ранее они были как на востоке Европы, так и на Кавказе (см. Приложение 1). Тем не менее полученные данные демонстрируют, что формирование современного облика митохондриального генофонда восточных европейцев произошло в эпоху бронзового века за счет притока митохондриальных гаплотипов с запада Европы. Между тем результаты анализа аутосомных локусов свидетельствуют о направлении миграций с востока на запад Европы, а данные по Y-хромосоме указывают на отсутствие такого вектора движения. Вполне вероятно, что регистрируемые по аутосомным данным миграции населения с востока Европы привели к формированию на территориях Центральной Европы смешанных популяций, представленных восточным аутосомным и западным (т. е. обогащенным гаплогруппами H, J, T, K, W) митохондриальными компонентами. Эти популяции затем начали экспансию на восток Европы и далее в Азию, неся с собой генетические характеристики, отсутствующие у древнего населения собственно Восточной Европы. О возможности такого сценария свидетельствуют результаты недавних исследований древнего населения Кавказа и Юго-Восточной Европы, показавшие возможность притока «анатолийского» генетического компонента в генофонд населения ямной культуры от энеолитического населения Северного Причерноморья, представленного

¹ Приложения 1, 2 см. по адресу:
<http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx10.pdf>

трипольской культурой (Mathieson et al., 2018; Wang et al., 2019). Очевидно, что для прояснения генетической истории европейцев и согласования данных, полученных для разных генетических систем, необходимы дальнейшие более детальные исследования.

Еще один из широко обсуждаемых вопросов в последнее время – участие древнего населения Кавказа в формировании населения ямной культуры и тем самым населения Европы в целом. Результаты анализа аутосомных локусов показали, что кавказский генетический компонент присутствует в генофондах европейцев, в связи с чем вполне возможно участие кавказских популяций в миграциях эпохи бронзового века отдельно или совместно с популяциями ямной культуры (Wang et al., 2019). Выше было отмечено, что приток кавказских митохондриальных гаплогрупп в генофонд восточных европейцев ограничен несколькими гаплогруппами мтДНК и, по всей видимости, не был интенсивным. Среди них наиболее информативной является гаплогруппа R1a. Проведенный нами филогенетический анализ R1a-митогеномов выявил, что эволюционный возраст этой гаплогруппы составляет 13–19 тыс. лет (Приложение 2). Наиболее древние R1a-гаплотипы найдены у населения Ирана и Кавказа, что указывает на возможное место происхождения этой гаплогруппы. Установлено, что гаплогруппа R1a представлена двумя ветвями – одной крупной R1a1 и небольшой новой подгруппой R1a2 (см. Приложение 2). Гаплогруппа R1a2 выявлена у населения Кавказа и Египта, а также к этой же гаплогруппе принадлежит гаплотип русского индивидуума из Псковской области. Гаплогруппа R1a1, представленная тремя подгруппами, R1a1a, R1a1b, R1a1c, распространена уже не только на Кавказе, но и в Европе. В R1a1a входят три подгруппы. Первая, R1a1a1, обнаружена в Восточной Европе среди русских и мордвы; вторая, R1a1a2, – в северногерманских популяциях, у англичан и датчан, а третья новая подгруппа, R1a1a3, – на юге Европы, у болгар и итальянцев. Возраст всей R1a1a составляет 8.5–12.4 тыс. лет, а ее подгрупп, R1a1a1, R1a1a2 и R1a1a3, – 4.2–5.5, 2.3–3.5 и 1.3–2.3 тыс. лет соответственно. Такая хронология указывает на возможное участие этих гаплогрупп мтДНК в миграционных событиях эпохи бронзового века. Филогеография R1a1a-гаплотипов свидетельствует о северо-западном направлении миграций, так как подгруппа R1a1a1 наблюдается среди новгородских русских и мордвы, а подгруппа R1a1a2 – среди северных германцев. Вместе с тем присутствие подгруппы R1a1a3 у болгар и итальянцев указывает также и на южноевропейское направление миграций.

Анализ палеогеномных данных демонстрирует, что наиболее древние R1a-гаплотипы обнаруживаются в энеолите Северного Кавказа (~4.5 тыс. лет до н.э.) у представителей майкопской (4–3.5 тыс. лет до н.э.) и катакомбной (~3 тыс. лет до н.э.) культур (Wang et al., 2019). Примерно к этому же времени относятся R1a1-гаплотипы представителей кура-араксской культуры (Армения) (Margaryan et al., 2017; Wang et al., 2019). Самые ранние случаи обнаружения подгруппы R1a1a относятся ко времени формирования северокавказской (кубано-терской) культуры (2.9–2.6 тыс. лет до н.э.) (Wang et al., 2019). Немного позже R1a-гаплотипы были выявлены у пред-

ставителей культуры колоколовидных кубков (2.5–2 тыс. лет до н.э.) Чехии и Англии (Olalde et al., 2018). В наше время наиболее высокие частоты гаплогруппы R1a наблюдаются на Северном Кавказе – вплоть до 5 % у адыгейцев, абхазов и карачаевцев (Джаубермезов и др., 2019). Таким образом, данные о распространенности древних и современных R1a-гаплотипов показывают, что эта гаплогруппа маркирует миграции населения с юга на север Европы в эпоху бронзового века и тем самым подтверждает участие кавказского генетического компонента в формировании генофонда европейцев.

В проведенных ранее исследованиях отмечалось, что маркеры мтДНК вполне информативны для обнаружения следов миграций эпохи бронзового века в евразийском масштабе. Так, филогенетическое родство между митохондриальными линиями гаплогруппы U2e2a1d (с возрастом ~5–7 тыс. лет) у населения Восточной Европы (русские, белорусы, сербы) и Индии находит свое объяснение с помощью представлений об индоевропейских миграциях населения Понто-Каспийских степей в различных направлениях, в том числе на Индийский субконтинент (см. Приложение 2) (Davidovic et al., 2017). Аналогичное объяснение (Silva et al., 2017) предложено и для гаплогруппы H2b, возраст которой составляет ~6–6.5 тыс. лет и включает гаплотипы как европейцев, так и населения Индии, Пакистана, Таджикистана, а также Южной Сибири (алтайцы, алтайские казахи и буряты) (см. Приложение 2). В этом случае основным аргументом в пользу восточноевропейского происхождения этой гаплогруппы являются находки H2b-гаплотипов у представителей ямной и срубной археологических культур (Lazaridis et al., 2014; Naak et al., 2015; Mathieson et al., 2015). Интересна также гаплогруппа U4d2, распространенная в популяциях Восточной Европы, а также среди коренного населения Сибири (Davidovic et al., 2017). Возраст ее составляет 6–7 тыс. лет, и поэтому вполне возможно, что носители этой гаплогруппы были участниками трансевразийских миграций эпохи бронзового века (см. Приложение 2). Гаплогруппа N1a1a1a1 (с возрастом ~7 тыс. лет) распространена у населения Восточной Европы (у русских, татар, эстонцев, финнов), Южной Сибири, Казахстана и даже Индии (Palanichamy et al., 2010). Такой ареал гаплогруппы предполагает, что распространение N1a1a1a1-гаплотипов могло быть связано с миграциями населения ямной культуры. Этим же можно объяснить распространенность гаплогрупп H8b1 (возраст ~5–8 тыс. лет) в популяциях Восточной Европы, Центральной Азии и Сибири (см. Приложение 2) (Derenko et al., 2014). Приведенные примеры свидетельствуют, что диверсификация гаплогрупп мтДНК произошла в начале бронзового века, что может быть связано с миграционными процессами в степном поясе Северной Евразии.

Заключение

Результаты исследований полиморфизма мтДНК у современного русского населения и древнего населения Европы и Кавказа отмечают, что истоки митохондриального генофонда русских нужно искать в эпоху бронзового века, когда имели место широкомасштабные перемещения европейских популяций. Полученные результаты свиде-

тельствуют о том, что ~4–2 тыс. лет до н. э. разнообразие митохондриального генофонда восточных европейцев увеличилось за счет появления митохондриальных линий, характерных для населения Центральной и Западной Европы. Следует подчеркнуть, что ранее предложенный сценарий распространения гаплогрупп мтДНК в неолите и бронзовом веке предполагал миграции носителей некоторых гаплогрупп (U2, U4, U5, I, T1, R) с востока на запад Европы в эпоху ранней бронзы, однако обратного потока гаплогрупп этот сценарий не предусматривал (Brandt et al., 2013), поэтому расширение спектра гаплогрупп мтДНК у восточных европейцев эпохи бронзового века объяснений не находило.

Список литературы / References

- Джаубермезов М.А., Екомасова Н.В., Рейдла М., Литвинов С.С., Габидулина Л.Р., Виллемс Р., Хуснутдинова Э.К. Генетическая характеристика балкарцев и карачаевцев по данным об изменчивости митохондриальной ДНК. *Генетика*. 2019;55(1):110-120. DOI 10.1134/S0016675819010053.
- [Dzaubermезov M.A., Ekomasova N.V., Reidla M., Litvinov S.S., Gabidullina L.R., Vilems R., Khusnutdinova E.K. Genetic characterization of Balkars and Karachays using mtDNA data. *Rus. J. Genet.* 2019;55(1):114-123. DOI 10.1134/S1022795419010058.]
- Малыарчук Б.А., Литвинов А.Н., Деренко М.В. Структура и формирование митохондриального генофонда русского населения Восточной Европы. *Генетика*. 2019;55(5):574-582. DOI 10.1134/S0016675819050102.
- [Malyarchuk B.A., Litvinov A.N., Derenko M.V. Structure and forming of mitochondrial gene pool of Russian population of Eastern Europe. *Rus. J. Genet.* 2019;55(5):622-629. DOI 10.1134/S1022795419050107.]
- Allentoft M.E., Sikora M., Sjögren K.G., Rasmussen S., Rasmussen M., Stenderup J., Damgaard P.B., Schroeder H., Ahlström T., Vinner L., Malaspinas A.S., Margaryan A., Higham T., Chivall D., Lynnerup N., Harvati L., Baron J., Della Casa P., Dąbrowski P., Duffy P.R., Ebel A.V., Epimakhov A., Frei K., Furmanek M., Gralak T., Gromov A., Gronkiewicz S., Grupe G., Hajdu T., Jarysz R., Khartanovich V., Khokhlov A., Kiss V., Kolář J., Kriiska A., Lasak I., Longhi C., McGlynn G., Merkevcicius A., Merkyte I., Metspalu M., Mkrtychyan R., Moiseyev V., Paja L., Pálfi G., Pokutka D., Pospieszny L., Price T.D., Saag L., Sablin M., Shishlina N., Smrčka V., Soenov V.I., Szeverényi V., Tóth G., Trifanova S.V., Varul L., Vicze M., Yepiskoposyan L., Zhitenev V., Orlando L., Sicheritz-Pontén T., Brunak S., Nielsen R., Kristiansen K., Willerslev E. Population genomics of Bronze Age Eurasia. *Nature*. 2015; 522(7555):167-172. DOI 10.1038/nature14507.
- Balanovsky O., Chukhryaeva M., Zaporozhchenko V., Urasin V., Zhabagin M., Hovhannisyan A., Agdzhoyan A., Dibirova K., Kuznetsova M., Koshel S., Pocheshkhova E., Alborova I., Skhalyakho R., Utevska O.; Genographic Consortium, Mustafin K., Yepiskoposyan L., Tyler-Smith C., Balanovska E. Genetic differentiation between upland and lowland populations shapes the Y-chromosomal landscape of West Asia. *Hum. Genet.* 2017;136(4):437-450. DOI 10.1007/s00439-017-1770-2.
- Batini C., Hallast P., Vågane Å.J., Zadik D., Eriksen H.A., Pamjav H., Sajantila A., Wetton J.H., Jobling M.A. Population resequencing of European mitochondrial genomes highlights sex-bias in Bronze Age demographic expansions. *Sci. Rep.* 2017;7(1):12086. DOI 10.1038/s41598-017-11307-9.
- Batini C., Hallast P., Zadik D., Delser P.M., Benazzo A., Ghirotto S., Arroyo-Pardo E., Cavalleri G.L., de Knijff P., Dupuy B.M., Eriksson H.A., King T.E., López de Munain A., López-Parra A.M., Loutradis A., Milasin J., Novelletto A., Pamjav H., Sajantila A., Tolun A., Winney B., Jobling M.A. Large-scale recent expansion of European patrilineages shown by population resequencing. *Nat. Commun.* 2015;6:7152. DOI 10.1038/ncomms8152.
- Bramanti B., Thomas M.G., Haak W., Unterlaender M., Jores P., Tambets K., Antanaitis-Jacobs I., Haidle M.N., Jankauskas R., Kind C.J., Lueth F., Terberger T., Hiller J., Matsumura S., Forster P., Burger J. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*. 2009;326(5949):137-140. DOI 10.1126/science.1176869.
- Brandt G., Haak W., Adler C.J., Roth C., Szécsényi-Nagy A., Karimnia S., Möller-Rieker S., Meller H., Ganslmeier R., Friederich S., Dresely V., Nicklisch N., Pickrell J.K., Sirocko F., Reich D., Cooper A., Alt K.W.; Genographic Consortium. Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. *Science*. 2013;342(6155):257-261. DOI 10.1126/science.1241844.
- Brandt G., Szécsényi-Nagy A., Roth C., Alt K.W., Haak W. Human paleogenetics of Europe – the known knowns and the known unknowns. *J. Hum. Evol.* 2015;79:73-92. DOI 10.1016/j.jhevol.2014.06.017.
- Davidovic S., Malyarchuk B., Aleksic J., Derenko M., Topalovic V., Litvinov A., Skonieczna K., Rogalla U., Grzybowski T., Stevanovic M., Kovacevic-Grujicic N. Mitochondrial super-haplogroup U diversity in Serbians. *Ann. Hum. Biol.* 2017;44(5):408-418. DOI 10.1080/03014460.2017.1287954.
- Derenko M., Malyarchuk B., Denisova G., Perkova M., Litvinov A., Grzybowski T., Dambueva I., Skonieczna K., Rogalla U., Tsybovsky I., Zakharov I. Western Eurasian ancestry in modern Siberians based on mitogenomic data. *BMC Evol. Biol.* 2014;14:217. DOI 10.1186/s12862-014-0217-9.
- Ehler E., Novotný J., Juras A., Chyleński M., Moravčík O., Pačes J. AmtDB: a database of ancient human mitochondrial genomes. *Nucl. Acids Res.* 2019;47:D29-D32. DOI 10.1093/nar/gky843.
- Haak W., Balanovsky O., Sanchez J.J., Koshel S., Zaporozhchenko V., Adler C.J., Der Sarkissian C.S., Brandt G., Schwarz C., Nicklisch N., Dresely V., Fritsch B., Balanovska E., Vilems R., Meller H., Alt K.W., Cooper A.; Members of the Genographic Consortium. Ancient DNA from European early Neolithic farmers reveals their near eastern affinities. *PLoS Biol.* 2010;8(11):e1000536. DOI 10.1371/journal.pbio.1000536.
- Haak W., Lazaridis I., Patterson N., Rohland N., Mallick S., Llamas B., Brandt G., Nordenfelt S., Harney E., Stewardson K., Fu Q., Mittnik A., Bánffy E., Economou C., Francken M., Friederich S., Pena R.G., Hallgren F., Khartanovich V., Khokhlov A., Kunst M., Kuznetsov P., Meller H., Mochalov O., Moiseyev V., Nicklisch N., Pichler S.L., Risch R., Rojo Guerra M.A., Roth C., Szécsényi-Nagy A., Wahl J., Meyer M., Krause J., Brown D., Anthony D., Cooper A., Alt K.W., Reich D. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015; 522(7555):207-211. DOI 10.1038/nature14317.
- Juras A., Chyleński M., Ehler E., Malmström H., Żurkiewicz D., Włodarczyk P., Wilk S., Peška J., Fojtík P., Králík M., Libera J., Bagińska J., Tunia K., Klocho V.I., Dabert M., Jakobsson M., Koško A. Mitochondrial genomes reveal an east to west cline of steppe ancestry in Corded Ware populations. *Sci. Rep.* 2018;8(1): 11603. DOI 10.1038/s41598-018-29914-5.
- Lazaridis I., Patterson N., Mittnik A., ..., Kelso J., Reich D., Krause J. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. 2014;513(7518):409-413. DOI 10.1038/nature13673.
- Malmström H., Gilbert M.T., Thomas M.G., Brandström M., Storå J., Molnar P., Andersen P.K., Bendixen C., Holmlund G., Götherström A., Willerslev E. Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Curr. Biol.* 2009;19(20):1758-1762. DOI 10.1016/j.cub.2009.09.017.
- Malyarchuk B., Litvinov A., Derenko M., Skonieczna K., Grzybowski T., Grosheva A., Shneider Y., Rychkov S., Zhukova O. Mitogenomic diversity in Russians and Poles. *Forensic Sci. Int. Genet.* 2017;30:51-56. DOI 10.1016/j.fsigen.2017.06.003.

- Margaryan A., Derenko M., Hovhannisyan H., Malyarchuk B., Heller R., Khachatryan Z., Avetisyan P., Badalyan R., Bobokhyan A., Melikyan V., Sargsyan G., Piliposyan A., Simonyan H., Mkrtchyan R., Denisova G., Yepiskoposyan L., Willerslev E., Allentoft M.E. Eight millennia of matrilineal genetic continuity in the South Caucasus. *Curr. Biol.* 2017;27(13):2023-2028.e7. DOI 10.1016/j.cub.2017.05.087.
- Mathieson I., Alpaslan-Roodenberg S., Posth C., ..., Krause J., Pinhasi R., Reich D. The genomic history of southeastern Europe. *Nature*. 2018;555(7695):197-203. DOI 10.1038/nature25778.
- Mathieson I., Lazaridis I., Rohland N., Mallick S., Patterson N., Roodenberg S.A., Harney E., Stewardson K., Fernandes D., Novak M., Sirak K., Gamba C., Jones E.R., Llamas B., Dryomov S., Pickrell J., Arsuaga J.L., de Castro J.M., Carbonell E., Gerritsen F., Khokhlov A., Kuznetsov P., Lozano M., Meller H., Mochalov O., Moiseyev V., Guerra M.A., Roodenberg J., Vergès J.M., Krause J., Cooper A., Alt K.W., Brown D., Anthony D., Lalueza-Fox C., Haak W., Pinhasi R., Reich D. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*. 2015;528(7583):499-503. DOI 10.1038/nature16152.
- Olalde I., Brace S., Allentoft M.E., ..., Barnes I., Lalueza-Fox C., Reich D. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature*. 2018;555(7695):190-196. DOI 10.1038/nature25738.
- Översti S., Onkamo P., Stoljarova M., Budowle B., Sajantila A., Palo J.U. Identification and analysis of mtDNA genomes attributed to Finns reveal long-stagnant demographic trends obscured in the total diversity. *Sci. Rep.* 2017;7(1):6193. DOI 10.1038/s41598-017-05673-7.
- Palanichamy M.G., Zhang C.L., Mitra B., Malyarchuk B., Derenko M., Chaudhuri T.K., Zhang Y.P. Mitochondrial haplogroup N1a phylogeography, with implication to the origin of European farmers. *BMC Evol. Biol.* 2010;10:304. DOI 10.1186/1471-2148-10-304.
- Perego U.A., Achilli A., Angerhofer N., Accetturo M., Pala M., Olivieri A., Hooshyar Kashani B., Ritchie K.H., Scozzari R., Kong Q.P., Myres N.M., Salas A., Semino O., Bandelt H.J., Woodward S.R., Torroni A. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr. Biol.* 2009;19(1):1-8. DOI 10.1016/j.cub.2008.11.058.
- Pereira J.B., Costa M.D., Vieira D., Pala M., Bamford L., Harich N., Cherni L., Alshamali F., Hatina J., Rychkov S., Stefanescu G., King T., Torroni A., Soares P., Pereira L., Richards M.B. Reconciling evidence from ancient and contemporary genomes: a major source for the European Neolithic within Mediterranean Europe. *Proc. Biol. Sci.* 2017;284(1851):20161976. DOI 10.1098/rspb.2016.1976.
- Poznik G.D., Xue Y., Mendez F.L., Willems T.F., Massaia A., Wilson Sayres M.A., Ayub Q., McCarthy S.A., Narechania A., Kashin S., Chen Y., Banerjee R., Rodriguez-Flores J.L., Cerezo M., Shao H., Gymrek M., Malhotra A., Louzada S., Desalle R., Ritchie G.R., Cerviera E., Fitzgerald T.W., Garrison E., Marcketta A., Mittelman D., Romanovitch M., Zhang C., Zheng-Bradley X., Abecasis G.R., McCarroll S.A., Flicek P., Underhill P.A., Coin L., Zerbino D.R., Yang F., Lee C., Clarke L., Auton A., Erlich Y., Handsaker R.E.; 1000 Genomes Project Consortium, Bustamante C.D., Tyler-Smith C. Punctuated bursts in human male demography inferred from 1244 worldwide Y-chromosome sequences. *Nat. Genet.* 2016;48(6):593-599. DOI 10.1038/ng.3559.
- Richards M.B., Soares P., Torroni A. Palaeogenomics: mitogenomes and migrations in Europe's past. *Curr. Biol.* 2016;26(6):R243-R246. DOI 10.1016/j.cub.2016.01.044.
- Silva M., Oliveira M., Vieira D., Brandão A., Rito T., Pereira J.B., Fraser R.M., Hudson B., Gandini F., Edwards C., Pala M., Koch J., Wilson J.F., Pereira L., Richards M.B., Soares P. A genetic chronology for the Indian Subcontinent points to heavily sex-biased dispersals. *BMC Evol. Biol.* 2017;17(1):88. DOI 10.1186/s12862-017-0936-9.
- Soares P., Ermini L., Thomson N., Mormina M., Rito T., Rohl A., Salas A., Oppenheimer S., Macaulay V., Richards M.B. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* 2009;84(6):740-759. DOI 10.1016/j.ajhg.2009.05.001.
- Wang C.C., Reinhold S., Kalmykov A., Wissgott A., Brandt G., Jeong C., Cheronet O., Ferry M., Harney E., Keating D., Mallick S., Rohland N., Stewardson K., Kantorovich A.R., Maslov V.E., Peterenko V.G., Erlikh V.R., Atabiev B.C., Magomedov R.G., Kohl P.L., Alt K.W., Pichler S.L., Gerling C., Meller H., Vardanyan B., Yeganyan L., Rezepkin A.D., Mariaschk D., Berezina N., Gresky J., Fuchs K., Knipper C., Schiffels S., Balanovska E., Balanovsky O., Mathieson I., Higham T., Berezina Y.B., Buzhilova A., Trifonov V., Pinhasi R., Belinskij A.B., Reich D., Hansen S., Krause J., Haak W. Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nat. Commun.* 2019;10(1):590. DOI 10.1038/s41467-018-08220-8.

ORCID ID

B.A. Malyarchuk orcid.org/0000-0002-0304-0652

Conflict of interest. The author declares no conflict of interest.

Received March 14, 2019. Revised May 14, 2019. Accepted May 20, 2019.

Association of polymorphism *TP53* Arg72Pro with radon-induced lung cancer in the Kazakh population

O. Bulgakova¹✉, A. Kussainova¹, A. Kakabayev², A. Kausbekova¹, R. Bersimbaev¹

¹ L.N. Gumilyov Eurasian National University, Institute of Cell Biology and Biotechnology, Nur-Sultan, Kazakhstan

² Sh. Ualikhanov Kokshetau State University, Kokshetau, Kazakhstan

✉ e-mail: ya.summer13@yandex.kz

Lung cancer is a problem of great concern and one of the commonest cancer diseases worldwide and in the Republic of Kazakhstan in particular. Radon exposure is classified as the second most important cause of lung cancer. According to the experts, the contribution of natural sources to the average annual radiation dose of the Kazakh population currently stands at 80 %, including 50 % from radon. However, the effect of radon on human health in the Republic of Kazakhstan is almost unknown. The tumor suppressor gene *TP53* is a key mediator of the DNA damage response cascade following cell exposure to ionizing radiation. The common polymorphism *TP53* Arg72Pro (rs1042522) is a risk factor for lung cancer in the Asian population, but until now no genetic association studies have been done in the Kazakh population. No information on the synergistic carcinogenic effect of radon exposure and polymorphism *TP53* Arg72Pro (rs1042522) is available either. This paper presents the results of the study of association between alteration in the *TP53* gene and radon-induced lung cancer risk in the Kazakh population. Genetic association was assessed in a case-control study including 44 radon-exposed patients with lung cancer, 41 patients with lung cancer without radon exposure and 42 age/sex-matched healthy controls. We found that polymorphism *TP53* Arg72Pro (rs1042522) was associated with lung cancer risk in the Kazakh population (OR = 6.95, 95 % CI = 2.41–20.05). Individuals with the Arg72Pro genotype also showed a significantly higher risk of radon-induced lung cancer (OR = 8.6, 95 % CI = 2.6–28.59).

Key words: polymorphism *TP53* Arg72Pro (rs1042522); radon; lung cancer; Kazakh population.

For citation: Bulgakova O., Kussainova A., Kakabayev A., Kausbekova A., Bersimbaev R. Association of polymorphism *TP53* Arg72Pro with radon-induced lung cancer in the Kazakh population. Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding. 2019;23(5):594-599. DOI 10.18699/VJ19.530

Ассоциация полиморфизма *TP53* Arg72Pro с риском развития радон-индуцированного рака легкого в казахской популяции

О. Булгакова¹✉, А. Кусаинова¹, А. Какабаев², А. Каусбекова¹, Р. Берсимбаев¹

¹ Евразийский национальный университет им. Л.Н. Гумилева, Институт клеточной биологии и биотехнологии, Нур-Султан, Казахстан

² Кокшетауский государственный университет им. Ш. Уалиханова, Кокшетау, Казахстан

✉ e-mail: ya.summer13@yandex.kz

Самый распространенный тип рака в структуре онкологических заболеваний – рак легкого, который представляет собой серьезную проблему не только в Республике Казахстан, но и по всему миру. Воздействие радона классифицируется как вторая причина развития онкологии легкого. По оценкам экспертов, естественный радиационный фон составляет 80 % от среднегодовой дозы облучения населения Казахстана, из которых 50 % приходится на радон. Однако следует отметить, что влияние радона на здоровье населения Республики Казахстан остается невыясненным. Ген-супрессор опухолей *TP53* является ключевым медиатором в каскаде реакций активируемых при повреждениях ДНК в результате воздействия на клетку ионизирующего излучения. Известно, что полиморфизм *TP53* Arg72Pro (rs1042522) часто встречается в азиатской популяции и считается фактором риска развития рака легкого, но до настоящего времени не была изучена его ассоциация с патогенезом этого заболевания в казахской популяции. Нет информации относительно синергетического канцерогенного эффекта воздействия радона и полиморфизма *TP53* Arg72Pro (rs1042522). В этой статье представлены результаты поиска связи полиморфного варианта гена *TP53* с риском развития рака легкого, вызванного радоном, в казахской популяции. Генетическая ассоциация была оценена по методу «случай-контроль». В исследование были включены 44 пациента (подвергшиеся воздействию радона) с диагнозом рак легкого, 41 пациент с диагнозом рак легкого из регионов с допустимыми показателями содержания радона и 42 здоровых человека соответствующего возраста и пола. Мы обнаружили корреляцию между полиморфизмом гена *TP53* Arg72Pro (rs1042522)

и риском развития рака легкого в казахской популяции ($OR = 6.95$, $95\% CI = 2.41-20.05$). Необходимо отметить, что лица с генотипом Arg72Pro также показали более высокую вероятность развития радон-индуцированного рака легкого ($OR = 8.6$, $95\% CI 2.6-28.59$).

Ключевые слова: полиморфизм *TP53* Arg72Pro (rs1042522), радон; рак легкого; казахская популяция.

Introduction

Radon is considered to be the second most frequent cause of lung cancer only to tobacco smoking (WHO Handbook on Indoor Radon..., 2009). According to the experts, the contribution of natural sources to the mean annual radiation dose for the Kazakh population currently stands at 80 %, including 50 % from radon (Stegnar et al., 2013; Bersimbaev, Bulgakova, 2015). However, the effect of radon on human health in Kazakhstan is not studied sufficiently.

Lung cancer is the leading cause of cancer death in Kazakhstan and the commonest form of cancer (Bersimbaev, Bulgakova, 2017). The incidence rates of lung cancer in various regions of Kazakhstan vary quite widely. The morbidity of lung cancer in regions with high radon concentrations (North Kazakhstan, Akmola region) is higher than that observed in areas with generally low radon levels (Bersimbaev, Bulgakova, 2017).

Radon emits radiation in the form of alpha particles, which interact with DNA either directly or indirectly through the generation of free radicals, producing double-strand breaks, large chromosomal aberrations, and point mutations (Robertson et al., 2013).

The key role in maintaining the genome stability is played by *TP53*, including the response to damage caused by radiation (Yngveson et al., 1999). *TP53* mutations were spotted in the development of tumors of many locations, including lung cancer (Deben et al., 2016). A number of *TP53* mutations are associated with tobacco smoking-induced lung cancers, and similar mutation hotspots have been identified that are not associated with other types of cancer, e.g., codon 157 (Vähäkangas et al., 2001). These mutation spectra are also different between smokers and non-smokers (Hainaut, Pfeifer, 2001).

Specific “hotspot” mutations in cancer-relevant genes have been described in radon-induced lung cancer. Most papers on *TP53* mutations in radon-associated lung cancer are dedicated to occupational studies on uranium miners (Vähäkangas et al., 1992).

Taylor et al. (1994) were the first to detect a *TP53* mutational spectrum different from those seen in lung cancers caused by tobacco smoke, and they reported a radon-related *TP53* hotspot in codon 249, exon 7. Few studies analyzed *TP53* punctual mutations in lung tumors from residential radon-exposed individuals, and their results were not univocally supportive of the mentioned hotspot in codon 249 (Lo et al., 1995; Yngveson et al., 1999). Radon-induced lung tumor mutations in the *TP53* gene were also found in codons 248 and 245 (Holstein et al., 1997). A recent study shows that alterations in several genes, including *TP53*, are implicated in lung cancer resulting from exposure to radon indoors (Choi et al., 2017).

The published data on the association between polymorphism *TP53* Arg72Pro (rs1042522) and lung cancer risk in Asians remain controversial (Wang et al., 2013). Zhao et al. (2018) have shown that the Pro72Pro genotype is associated with a higher risk of cancer due to poorer ability to induce apoptosis.

However, no studies of the sort have been conducted in the Kazakh population. We have not found case-control studies providing evidence for polymorphism *TP53* Arg72Pro (rs1042522) and radon exposure interaction in the risk of lung cancer. So, in the present study we aim to study the association between *TP53* gene allelism and radon-induced lung cancer in the Kazakh population.

Materials and methods

Study design and population. A case-control study was conducted in 2015–2017. Eighty-five lung cancer patients (cases) and forty-two healthy individuals (controls) from the Akmola region of Kazakhstan were recruited for the study. The Akmola region is characterized for having high indoor radon concentrations due to uranium deposits and uranium mining enterprises located in this area (Bersimbaev, Bulgakova, 2015). All participants of the study had lived in this region for at least five years.

The cases had morphopathologically confirmed lung cancer. Subsequent to the results of the radon measurements, lung cancer patients were divided into two groups: 44 radon-induced lung cancer patients (RLC) and 41 lung cancer patients without high level of radon exposure (LC). The criterion for the RLC group participants was indoor radon level exceeding 80 Bq/m³ in the homes. All participants should have lived for at least five years in their current residence. We chose 80 Bq/m³ as a threshold level in our study because WHO Handbook on Indoor Radon indicated that the threshold level of radon having a biological effect was no more than 100 Bq/m³ (WHO Handbook on Indoor Radon..., 2009).

Controls (C) were recruited from individuals attending hospital for reasons other than cancer. They were matched with the cases of lung cancer (LC) with respect to age, gender, and tobacco consumption. The indoor radon level in the residence of participants from control group should be less than 80 Bq/m³. Characteristics of subjects are presented in Table 1.

Exposure assessment. Radon detectors were placed in the participants' homes. The devices were Canary 222 Digital Electronic Radon Gas Monitor (LR-03) radiometers (Corentium AS, Oslo, Norway). A radiometer was installed away from doors, windows, or electrical devices and within 60–180 cm off the floor in rooms previously unventilated for at least 24 h according to the manufacturer's instructions. In each room, radon in the air was measured for seven days and the average value was used for further calculations.

The annual effective dose (H) was calculated according to the formula (Quarto et al., 2015):

$$H(\text{mSv/y}) = C \times F \times O \times T \times D,$$

where C stands for the average radon concentration Bq/m³; F is the equilibrium factor for indoor, set at 0.4; O is the occupancy factor, taken to be 0.8; T is time in hours in a year (8760 h/y); and D is the dose conversion factor; 1.4×10^{-8} Sv per Bq/m³·h.

Table 1. Characteristics of the subject participating in the study

Parameter		Radon lung cancer (RLC) (n = 44)	Lung cancer (LC) (n = 41)	p	C (n = 42)	p
Gender	Male	34 (77 %)	34 (83 %)	> 0.05 ^b	33 (79 %)	> 0.05 ^a
	Female	10 (23 %)	7 (17 %)		9 (21 %)	
Age, years	≤ 60	14 (32 %)	15 (37 %)	> 0.05 ^a	22 (52 %)	> 0.05 ^a
	> 60	25 (57 %)	25 (61 %)		20 (48 %)	
	Unknown	5 (11 %)	1 (2 %)		–	
Tobacco consumption	Non-smokers	22 (50 %)	10 (24 %)	< 0.05 ^c	16 (38 %)	> 0.05 ^a
	Smokers	22 (50 %)	31 (76 %)		26 (62 %)	

Notes: ^a – *t* test for both cases (RLC and LC) and control (C) groups; ^b – *t* test for RLC and LC groups; ^c – χ^2 test for RLC and LC groups. Columns: RLC – lung cancer patients exposed to radon; LC – lung cancer patients without exposure to radon; C – control group. All information regarding participants was rendered anonymous after data and blood sample collection. Informed consent was obtained from each study participant before interview and blood collection. The present study was approved by the Ethical Committee of the Semey State Medical University (Semey, Kazakhstan; approval No. 2).

Table 2. The PCR amplification and restriction protocols

Alteration in the <i>TP53</i> gene	Primers for PCR (5'→3')	PCR conditions	Amplicon length, bp	Restriction enzyme	Restricted product length and corresponding genotype
<i>TP53</i> Arg72Pro (rs 1042522)	FTTCCGTCCTCAAGCAATGGATGA; rTCTGGAAGGGACAGAAGATGAC	Predenaturation: 94 °C, 5 min; 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s; postextension: 72 °C, 7 min	199	BstUI (cat no. R0518S; NEB, USA)	Arg72Arg – 199 bp; Arg72Pro – 199 bp, 113 bp, 87 bp; Pro72Pro – 113bp, 87 bp

Collection of blood samples, DNA extraction and genotyping. Blood samples were collected from all participants by the vein puncture method, and DNA was extracted from blood by the conventional phenol-chloroform method (Sambrook et al., 1989). The genotyping of polymorphism *TP53* Arg72Pro (rs1042522) was performed by PCR-RFLP as previously described by M.K. Chowdhury et al. (2015). The PCR details and relevant information are provided in Table 2.

Assay of cotinine in plasma. The smoking status of the participants was verified using blood plasma cotinine as a marker. The cotinine level was determined using an ELISA kit (Cotinine ELISA kit; cat no. KA0930; Abnova, Taipei, Taiwan) as described in (Bulgakova et al., 2018).

Statistical analysis. Student’s unpaired *t* test was performed to calculate statistically significant difference in gender, age, and tobacco consumption between the LC and control groups. Student’s *t* test was also used to compare the distribution of variables between RLC and LC cohorts. The Chi-square test was used to compare the distributions of tobacco consumption between the RLC and LC groups. All statistical analyses were performed using GraphPad Prism 6 software (GraphPad Software, Inc., La Jolla, CA, USA). “Case-control Study Estimating Calculator” from Gene Expert Company (State Research Institute of Genetics and Selection of Industrial Microorganisms of the National Research Center “Kurchatov Institute”, Russian Federation, http://gen-exp.ru/calculator_or.php) was used in calculating OR, 95 % CI, and *p* values for association between the case-control status and polymorphism *TP53* Arg72Pro (rs1042522). A *p* value of ≤ 0.05 was considered significant. To assess whether the genetic polymorphism modified the effect of residential radon exposure on lung cancer, logistic regression was performed. We also assessed the possibility of a synergism between radon exposure and tobacco

consumption using the method proposed by D. Hosmer and S. Lemeshow (1992). Analyses were conducted with MedCalc software (Version 18.6, MedCalc Software, Belgium).

Results

Characterization of the sample studied. Initially, the case-control study involved a larger number of participants. The exclusion criteria from our study were occupation of dwelling for less than five years, absence of histological confirmation of lung cancer, and belonging to other ethnic groups (Russians, Ukrainians, Germans, etc.) Participants younger than 30 years were also excluded from the study. Patients diagnosed with lung cancer were included in the study regardless of the histological type of cancer. When selecting the control group of healthy individuals, every attempt was made to match them with the cancer patients according to the basic population characteristics. There were no significant differences in the distribution of age, gender, and smoking status between the cases (LC) and controls (C), but there was a small difference between smokers in the RLC and LC groups (see Table 1).

Effective annual radon exposure dose and decay product inhalation. The average equivalent equilibrium radon volume activity (EEVA) in the RLC group was 307.6 Bq/m³ according to the measurements of radon at the homes of monitored subjects. The range of EEVA was from 105 to 716 Bq/m³.

To evaluate the effect of radon exposure on the lung tissue, we used the effective annual dose, which is the tissue-weighted sum of the equivalent doses. The mean effective annual dose for lung cancer patients living in areas with high radon levels (RLC) was 7.5 mSv/y. The minimum and maximum effective annual doses in the RLC group were 4.12 and 20.76 mSv/y.

The same data in the lung cancer patients living in the area with a low level of radon (LC) were 40.6 Bq/m³ (from 8 to

Table 3. Association between polymorphism *TP53* Arg72Pro (rs1042522) and development of lung cancer

Genotype	LC (n = 41)	C (n = 42)	OR	95 % CI	p
Arg72Arg	11 (27 %)	30 (72 %)	0.15	0.06–0.38	0.0001
Arg72Pro	22 (54 %)	6 (14 %)	6.95	2.41–20.05	
Pro72Pro	8 (19 %)	6 (14 %)	1.45	0.46–4.64	

Table 4. Association between polymorphism *TP53* Arg72Pro (rs1042522) and the risk of radon-induced lung cancer

Genotype	Lung cancer (LC) (n = 41)	Radon lung cancer (RLC) (n = 44)	OR	95 % CI	p
Arg72Arg	11 (27 %)	1 (2 %)	0.06	0.01–0.52	0.0004
Arg72Pro	22 (54 %)	40 (91 %)	8.64	2.61–28.59	
Pro72Pro	8 (19 %)	3 (7 %)	0.30	0.07–1.23	

78 Bq/m³) and 2.0 mSv/y (from 0.31 to 3.06 mSv/y), respectively.

EVA levels in the control group fell within the range from 2 to 80 Bq/m³ with the average value 22.5 Bq/m³. The mean effective annual dose for the control group was 0.88 mSv/y. The minimum and maximum effective annual doses in the control group were 0.08 and 3.14 mSv/y, respectively.

Association of polymorphism *TP53* Arg72Pro (rs1042522) with lung cancer in the Kazakh population. We investigated polymorphism *TP53* Arg72Pro (rs1042522) and lung cancer risk in the Kazakh population regardless of radon exposure. For this purpose, we analyzed the distribution of polymorphism *TP53* Arg72Pro (rs1042522) genotypes Arg72Arg, Arg72Pro and Pro72Pro in LC and control groups. The genotype frequencies in the control and LC groups are shown in Table 3.

Statistical analysis of association between a genetic polymorphism and development of lung cancer was conducted by evaluating the data with regard to the general and additive models. Unfortunately, we could not use the recessive and dominant models in view of the small number of participants. According to the general model (see Table 3), there is a significant risk of lung cancer associated with the Arg72Pro (OR = 6.95, 95 % CI = 2.41–20.05) and Pro72Pro genotypes (OR = 1.45, 95 % CI = 0.46–4.64), for all genotypes $\chi^2 = 18.22$, $p = 0.0001$. The additive model confirmed this finding ($\chi^2 = 9.22$, $p = 0.002$). When we analyze separately non-smokers and smokers, we can observe a higher risk of lung cancer for the Arg72Pro genotype (OR = 4.33 95 % CI = 1.17–15.99) in smokers, but the results are not statistically significant ($p = 0.06$). Cotinine levels < 10 ng/ml were considered the criterion of the absence of current smoking.

Effects of residential radon exposure and polymorphism *TP53* Arg72Pro (rs1042522) on lung cancer risk. To estimate OR for radon-induced lung cancer, we analyzed the distribution of polymorphism *TP53* Arg72Pro (rs1042522) genotypes Arg72Arg, Arg72Pro and Pro72Pro in both lung cancer groups (LC and RLC). The logistic regression employed the effective annual dose, which was modeled with one indicator variable for ≥ 4.12 mSv/y and one indicator variable for missing with < 3.14 mSv/y as the common reference exposure. As shown in Table 4, the risk of radon-induced lung cancer is higher in participants with the Arg72Pro genotype.

The odds ratio is 8.64 (95 % CI = 2.61–28.59; $p = 0.0004$). No additive interaction was observed for Arg72Pro, radon exposure, and tobacco consumption ($\chi^2 = 10.33$, $p = 0.066$). Cotinine levels < 10 ng/ml were considered the criterion for the absence of current smoking.

Discussion

It was shown that polymorphism *TP53* Arg72Pro (rs1042522) plays an important role in the development of different types of cancer (Francisco et al., 2011), including lung cancer (Neumann et al., 2018). Katkoori et al. (2017) demonstrated the high survival of cells that express P72^{wt}. This effect can be explained by the fact that the Pro72Pro variant has a weaker *TP53* transcriptional activity compared to Arg72Arg (Thomas et al., 1999; Zhao et al., 2018). As many of the *TP53* target genes are involved in apoptosis regulation, the Pro72Pro genotype can dysregulate programmed cell death and activate the tumor-promoting phenotype.

On the other hand, the protein product of the Arg72Arg genotype is more effective in inducing apoptosis, although Arg72Arg is associated with faster degradation than Pro72Pro (Storey et al., 1998). Neumann et al. (2018) have shown an adverse prognostic value for the presence of heterozygous genotype Arg72Pro in tumor samples from lung cancer patients. In a meta-analysis of 51 studies, polymorphism *TP53* Arg72Pro (rs1042522) was significantly associated with lung cancer risk in any genetic model (Ye et al., 2014). According to S. Wang et al. (2013), the Arg72Pro genotype positively correlated with lung cancer risk in Asians. Only few studies attempted to determine the impact of polymorphism *TP53* Arg72Pro (rs1042522) on cancer risk in the Kazakh population.

The results of this study indicate a significant risk of lung cancer for Arg72Pro (OR = 6.95, 95 % CI = 2.41–20.05) and Pro72Pro genotypes (OR = 1.45, 95 % CI = 0.46–4.64), for all genotypes $\chi^2 = 18.22$, $p = 0.0001$ in the Kazakh population.

The product of the *TP53* gene is involved in the response to damage caused by radiation (Yngveson et al., 1999).

Radon is a decay product of uranium. It emits α radiation, which is carcinogenic for humans. The effects of α radiation include different types of DNA damage leading to mutations, chromosomal aberrations, and cell transformation. Some studies infer that the radon-induced mutation spectra in lung cancer may differ from the mutations in patients with lung cancer

induced by tobacco smoking (Taylor et al., 1994; Hollstein et al., 1997). Yngveson et al. (1999) have shown that residential exposure to radon seems to contribute to a higher mutation prevalence of the *TP53* gene in lung tumors, especially among nonsmokers.

There are only few studies on radon exposure and gene polymorphisms (Bonner et al., 2006; Ruano-Ravina et al., 2014; Leng et al., 2016). Two of them provided evidence for an interaction between members of the glutathione S-transferase family *GSTT1* (Ruano-Ravina et al., 2014) and *GSTM1* (Bonner et al., 2006; Ruano-Ravina et al., 2014) and radon in lung cancer risk. The association between *IL-6* promoter SNPs and squamous cell carcinoma was studied in uranium miners (Leng et al., 2016).

Choi et al. (2018) identified that *CHD4* rs74790047, *TSC2* rs2121870, and *AR* rs66766408 are found to be common exonic mutations in both lung cancer patients and normal individuals exposed to radon indoors.

Our study shows that the effect of radon on lung cancer risk depends on polymorphism *TP53* Arg72Pro (rs1042522). Thus, radon-inducible lung cancer risk increases in participants with the Arg72Pro genotype as compared to two other genotypes, Arg72Arg and Pro72Pro.

A possible explanation for these findings is that radon causes a genetic damage (Jostes, 1996). Druzhinin et al. (2015) reported that the frequencies of chromosomal aberrations in peripheral blood lymphocytes were significantly elevated in long-term resident children in a boarding school under conditions of high exposure to radon. The polymorphism *TP53* Arg72Pro (rs1042522) can affect the induction of cell apoptosis (Pereira et al., 2011) and the repair of damaged DNA (Hu et al., 2005). It is possible that the Arg72Pro genotype does not promote efficient DNA repair or apoptosis in comparison with Arg72Arg and these events lead to malignant transformation of cells. So, Y. Hu et al. (2005) have shown that the Pro72Pro genotype is associated with higher frequency of *TP53* mutations in non-small-cell lung cancer.

There are some limitations in our study. The number of participants was insufficient for comprehensive study of the effect of the *TP53* gene polymorphism-radon interaction. Therefore, we could not analyze the influence of the Pro72Pro genotype on the risk of radon-induced lung cancer. Further studies are warranted to address the relationships among residential radon exposure, *TP53* gene polymorphisms, and lung cancer risk.

Conclusion

To sum up, the exposure to residential radon interacts with the Arg72Pro genotype to increase the risk of lung cancer in the Kazakh population. Our study supports the hypothesis that polymorphism *TP53* Arg72Pro (rs1042522) can modulate the pathogenic effect of radon in lung tissue.

References

Bersimbaev R., Bulgakova O. The health effects of radon and uranium on the population of Kazakhstan. *Genes Environ.* 2015;37:18. DOI 10.1186/s41021-015-0019-3.
Bersimbaev R., Bulgakova O. Residential radon exposure and lung cancer risk in Kazakhstan. Ed. F. Adrovic. *Radon*. London, 2017; 93-124. DOI 10.5772/intechopen.71135.
Bonner M.R., Bennett W.P., Xiong W., Lan Q., Brownson R.C., Harris C.C., Field R.W., Lubin J.H., Alavanja M.C. Radon, secondhand

smoke, glutathione-S-transferase M1 and lung cancer among women. *Int. J. Cancer.* 2006;119(6):1462-1467.
Bulgakova O., Zhabayeva D., Kussainova A., Pulliero A., Izzotti A., Bersimbaev R. miR-19 in blood plasma reflects lung cancer occurrence but is not specifically associated with radon exposure. *Oncol. Lett.* 2018;15(6):8816-8824.
Choi J.R., Koh S.-B., Kim H.R., Lee H., Kang D.R. Radon exposure-induced genetic variations in lung cancers among never smokers. *J. Korean Med. Sci.* 2018;33(29):e207.
Choi J.R., Koh S.-B., Park S.Y., Kim H.R., Lee H., Kang D.R. Novel genetic associations between lung cancer and indoor radon exposure. *J. Cancer Prev.* 2017;22(4):234-240.
Chowdhury M.K., Moniruzzaman M., Emran A.A., Mostafa M.G., Kuddus R.H., Uddin M.A. *TP53* codon 72 polymorphisms and lung cancer risk in the Bangladeshi population. *Asian. Pac. J. Cancer Prev.* 2015;16(8):3493-3498.
Deben C., Deschoolmeester V., Lardon F., Rolfo C., Pauwels P. *TP53* and *MDM2* genetic alterations in non-small cell lung cancer. Evaluating their prognostic and predictive value. *Crit. Rev. Oncol. Hematol.* 2016;99:63-73.
Druzhinin V.G., Sinitsky M.Y., Larionov A.V., Volobaev V.P., Minina V.I., Golovina T.A. Assessing the level of chromosome aberrations in peripheral blood lymphocytes in long-term resident children under conditions of high exposure to radon and its decay products. *Mutagenesis.* 2015;30(5):677-683.
Francisco G., Menezes P.R., Eluf-Neto J., Chammas R. Arg72Pro *TP53* polymorphism and cancer susceptibility. a comprehensive meta-analysis of 302 case-control studies. *Int. J. Cancer.* 2011;129(4):920-930.
Hainaut P., Pfeifer G. Patterns of p53 G→T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis.* 2001;22:367-374.
Hollstein M., Bartsch H., Wesch H., Kure E.H., Mustonen R., Mühlbauer K.R., Spiethoff A., Wegener K., Wiethege T., Müller K.M. p53 gene mutation analysis in tumors of patients exposed to alpha-particles. *Carcinogenesis.* 1997;18(3):511-516.
Hosmer D.W., Lemeshow S. Confidence interval estimation of interaction. *Epidemiology.* 1992;3:452-456.
Hu Y., McDermott M.P., Ahrendt S.A. The p53 codon 72 proline allele is associated with p53 gene mutations in non-small cell lung cancer. *Clin. Cancer Res.* 2005;11(7):2502-2509.
Jostes R.F. Genetic, cytogenetic, and carcinogenic effects of radon: a review. *Mutat. Res.* 1996;340(2-3):125-139.
Katkoori V.R., Manne U., Chaturvedi L.S., Basson M.D., Haan P., Coffey D., Bumpers H.L. Functional consequence of the p53 codon 72 polymorphism in colorectal cancer. *Oncotarget.* 2017;8(44):76574-76586.
Leng S., Thomas C.L., Snider A.M., Picchi M.A., Chen W., Willis D.G., Carr T., Krzeminski J., Desai D., Shantu A., Lin A., Jacobson M.R., Belinsky S.A. Radon exposure, *IL-6* promoter variants, and lung squamous cell carcinoma in former uranium miners. *Environ. Health Persp.* 2016;124(4):445-451.
Lo Y., Darby S., Noakes L., Whitley E., Silcocks P., Fleming K., Bell J. Screening for codon 249 p53 mutation in lung cancer associated with domestic radon exposure. *Lancet.* 1995;345:360.
Neumann M.P., González M.V., Pitiot A.S., Santamaría Í., Martínez C., Tardón A., Astudillo A., Balbín M. *TP53* p.R72P genotype is a marker of poor prognosis in lung cancer. *Cancer Biomark.* 2018;21(4):747-754.
Pereira L., Carvalho M.R., Fonseca C.G., Lima S.S., Cerqueira E.M., Jorge W., Castro M.C. Influence of Arg72Pro polymorphisms of *TP53* on the response of buccal cells to radiotherapy. *Genet. Mol. Res.* 2011;10(4):3552-3558.
Quarto M., Pugliese M., La Verde G., Loffredo F., Roca V. Radon exposure assessment and relative effective dose estimation to inhabitants of puglia region, South Italy. *Int. J. Environ. Res. Public Health.* 2015;12:14948-14957.

- Robertson A., Allen J., Laney R., Curnow A. The cellular and molecular carcinogenic effects of radon exposure: a review. *Int. J. Mol. Sci.* 2013;14(7):14024-14063.
- Ruano-Ravina A., Pereyra M.F., Castro M.T., Pérez-Ríos M., Abal-Arca J., Barros-Dios J.M. Genetic susceptibility, residential radon, and lung cancer in a radon prone area. *J. Thorac. Oncol.* 2014;9(8):1073-1080.
- Sambrook J., Fritsch E.F., Maniatis T. *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press, 1989.
- Stegnar P., Shishkov I., Burkitbayev M., Tolongutov B., Yunusov M., Radyuk R., Salbu B. Assessment of the radiological impact of gamma and radon dose rates at former U mining sites in Central Asia. *J. Environ. Radioact.* 2013;123:3-13.
- Storey A., Thomas M., Kalita A., Harwood C., Gardiol D., Mantovani F., Breuer J., Leigh I.M., Matlashewski G., Banks L. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature*. 1998;393:229-341.
- Taylor J., Watson M., Devereux T., Michels R., Saccomanno G., Anderson M. p53 mutation hotspot in radon-associated lung cancer. *Lancet*. 1994;343(8889):86-87.
- Thomas M., Kalita A., Labrecque S., Pim D., Banks L., Matlashewski G. Two polymorphic variants of wild-type p53 differ biochemically and biologically. *Mol. Cell Biol.* 1999;19(2):1092-1100.
- Vähäkangas K., Bennett W., Castrén K., Welsh J., Khan M., Blömeke B., Alavanja M., Harris C. p53 and K-ras mutations in lung cancers from former and never-smoking women. *Cancer Res.* 2001;61:4350-4356.
- Vähäkangas K., Samet J., Metcalf R., Welsh J., Bennett W., Lane D., Harris C. Mutations of *p53* and *ras* genes in radon-associated lung cancer from uranium miners. *Lancet*. 1992;339(8793):576-580.
- Wang S., Lan X., Tan S., Wang S., Li Y. P53 codon 72 Arg/Pro polymorphism and lung cancer risk in Asians: an updated meta-analysis. *Tumor Biol.* 2013;34(5):2511-2520.
- WHO Handbook on Indoor Radon: A Public Health Perspective. Eds. W.J. Angell, H. Zeeb, F. Shannon. Geneva: World Health Organization, 2009.
- Ye X.H., Bu Z.B., Feng J., Peng L., Liao X.B., Zhu X.L., Sun X.L., Yu H.G., Yan D.F., Yan S.X. Association between the TP53 polymorphisms and lung cancer risk: a meta-analysis. *Mol. Biol. Rep.* 2014;41(1):373-385.
- Yngveson A., Williams C., Hjerpe A., Lundeberg J., Söderkvist P., Pershagen G. p53 mutations in lung cancer associated with residential radon exposure. *Cancer Epidemiol. Biomarkers Prev.* 1999;8:433-438.
- Zhao Y., Wu L., Yue X., Zhang C., Wang J., Li J., Sun X., Zhu Y., Feng Z., Hu W. A polymorphism in the tumor suppressor p53 affects aging and longevity in mouse models. *eLife*. 2018;7:e34701. DOI 10.7554/eLife.34701.

ORCID ID

O.V. Bulgakova orcid.org/0000-0003-3272-0638
A.A. Kussainova orcid.org/0000-0002-1313-9779
A.A. Kakabayev orcid.org/0000-0002-3825-062
A.Zh. Kausbekova orcid.org/0000-0002-7012-2711
R.I. Bersimbaev orcid.org/0000-0002-9671-1178

Acknowledgements. We thank the Astana Oncology Center (Nur-Sultan city) and Akmola Region Oncology Hospital (Kokshetau city) for their help in collection of samples.

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

Received February 19, 2019. Revised May 15, 2019. Accepted May 15, 2019.

Epigenetics of suicidal behavior

R.N. Mustafin^{1,2}✉, A.V. Kazantseva³, R.F. Enikeeva^{2,3}, Yu.D. Davydova³, S.B. Malykh⁴, V.V. Viktorov¹, E.K. Khusnutdinova^{1,2,3}

¹ Bashkir State Medical University, Ufa, Russia

² Bashkir State University, Ufa, Russia

³ Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre, RAS, Ufa, Russia

⁴ Psychological Institute of the Russian Academy of Education, Moscow, Russia

✉ e-mail: ruji79@mail.ru

Suicide is the second leading cause of death among young people and therefore being a serious global problem worldwide. The study of genetic and epigenetic factors in the development of suicidal behavior plays an important role in the development of advanced methods of diagnosis and treatment of this pathology. The role of hereditary factors in the development of suicidal behavior is estimated at 30–55 %, with a pronounced comorbidity with other psychopathologies. The study of genetic liability to suicidal behavior is based on molecular-genetic methods including association and linkage analyses, chip gene expression arrays, and genome-wide association studies. Published data identified multiple genes including those involved in the functioning of serotonergic (*SLC6A4*, *TPH*, *5-HT1A*), hypothalamic-pituitary-adrenal systems (*FKBP5*) and polyamines (*SAT* and *OATL1*) associated with suicidal behavior. However, the diversity of interacting genetic loci complicates the interpretation of the development of a complex phenotype of pathology and prevents the association from being detected. To solve this problem and interpret the missing relationship between the environment and the genome, promising results were obtained from a study of epigenetic factors, which affected the expression of a number of candidate genes involved in brain functioning in suicidal behavior. The analysis of a brain obtained from suicide victims, representing a unique tool for the analysis of modified genomic processes, revealed a wide range of reprogramming patterns of DNA methylation in promoters of the genes of polyamine (*OAZ1*, *OAZ2*, *AMD1*, *ARG2*, *SKA2*), serotonergic (*SLC6A4*) and GABAergic (*GABRA1*) systems, HPA-axis (*GR*, *NR3C1*), tyrosine kinase (*TrkB*) receptors, brain-derived neurotrophic factor (*BDNF*). The role of histone modifications in distinct genes (*Cx30*, *Cx43*, *TrkB.T1*) and the expression of specific long non-coding RNAs and microRNAs in the development of suicidal behavior, which is promising for the development of diagnostic algorithms and target therapy, is discussed.

Key words: association; brain; methylation; non-coding RNAs; suicide; epigenetics.

For citation: Mustafin R.N., Kazantseva A.V., Enikeeva R.F., Davydova Yu.D., Malykh S.B., Viktorov V.V., Khusnutdinova E.K. Epigenetics of suicidal behavior. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5):600-607. DOI 10.18699/VJ19.531

Эпигенетика суицидального поведения

Р.Н. Мустафин^{1,2}✉, А.В. Казанцева³, Р.Ф. Еникеева^{2,3}, Ю.Д. Давыдова³, С.Б. Малых⁴,
В.В. Викторов¹, Э.К. Хуснутдинова^{1,2,3}

¹ Башкирский государственный медицинский университет, Уфа, Россия

² Башкирский государственный университет, Уфа, Россия

³ Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия

⁴ Психологический институт Российской академии образования, Москва, Россия

✉ e-mail: ruji79@mail.ru

Самоубийства занимают второе место среди причин смерти среди молодежи, в связи с чем являются серьезной глобальной проблемой человечества. Для разработки перспективных методов диагностики и лечения данной патологии важное значение имеет исследование генетических и эпигенетических факторов в развитии суицидального поведения. Роль наследственных факторов в развитии суицидального поведения оценивается в 30–55 %, при этом характерна выраженная коморбидность с другими психическими расстройствами. Для исследования генетической предрасположенности к суициду используются молекулярно-генетические методы, включая контролируемые анализы ассоциаций и сцепления, микроматричные анализы экспрессии генов и полногеномный поиск ассоциаций. В литературе представлены данные об идентификации множества генов, в том числе связанных с изменениями функционирования серотонинергической (гены *SLC6A4*, *TPH*, *5-HT1A*), гипоталамо-гипофизарно-надпочечниковой систем (ген *FKBP5*) и полиаминов (гены *SAT* и *OATL1*), ассоциированных с развитием суицидального поведения. Однако разнообразие взаимодействующих генетических локусов усложняет интерпретацию развития сложного фенотипа патологии и не позволяет определить выраженную ассоциацию. Для разрешения данной проблемы и интерпретации недостающей связи между окружающей средой и геномом были получены многообещающие результаты при изучении эпигенетических факторов, роль которых при суицидальном поведении показана в изменении экспрессии

ряда кандидатных генов, вовлеченных в функционирование головного мозга. Уникальным объектом для прямого исследования изменения геномных процессов является головной мозг умерших от суицида людей, при изучении которого был выявлен широкий спектр репрограммирования паттернов ДНК-метилирования промоторов генов системы полиаминов (*OAZ1*, *OAZ2*, *AMD1*, *ARG2*, *SKA2*), серотонинергической (*SLC6A4*) и ГАМК-ергической (*GABRA1*) систем, глюкокортикоидных (*GR*, *NR3C1*) и тирозинкиназных (*TrkB*) рецепторов, нейротрофического фактора головного мозга (*BDNF*). Показана роль изменений модификации гистонов в области расположения специфических генов (*Cx30*, *Cx43*, *TrkB.T1*) и экспрессии специфических длинных некодирующих РНК и микроРНК в развитии суицидального поведения, что перспективно для разработки программ диагностических алгоритмов и таргетной терапии.

Ключевые слова: ассоциации; головной мозг; метилирование; некодирующие РНК; суицид; эпигенетика.

Introduction

Suicide represents violence against oneself with the serious intention to destroy life. Every 40 seconds in the world 1 individual dies from suicide (Roy, Dwivedi, 2017). According to the WHO, about 1 million individuals commit suicide every year (Cui et al., 2017), and the global prevalence of suicide is 11.4 per 100,000 individuals (Lutz et al., 2017). Suicidal behavior (SB) is a generic term used to denote risk, attempts and committed suicide (Bani-Fatemi et al., 2015). SB is the second cause of death among individuals of young age and takes the 10th place in all age groups worldwide (Roy, Dwivedi, 2017; Fanelli, Serretti, 2018). For example, a longitudinal study of adolescents aged 13–18 revealed that 12.1 % of American adolescents had suicidal thoughts, 4 % were planning suicides, and 4.1 % committed suicide (Nock et al., 2013). These observations together with the absence of appropriate preventive strategies make SB an essential public health problem (Roy, Dwivedi, 2017).

SB is a multifactorial pathology showing high comorbidity with mental illness (MI) and major psychopathologies, including major depressive disorder (MDD), bipolar disorder (Ludwig et al., 2017) and schizophrenia (Bani-Fatemi et al., 2015). For example, about 4 % of MDD patients die as a result of suicide (Serafini et al., 2012), and the youngest individuals with SB are diagnosed with mental illness (Nock et al., 2013). Certain environmental stimuli increase the genetic predisposition to SB in MDD patients (Roy, Dwivedi, 2018). However, it should be taken into account that SB is caused by complex risk factors, which are not universal for each individual (Turecki, 2014).

Over the past decades, a number of theories have been proposed explaining the mechanisms for the development of SB. According to one of the most influential theories proposed by Mann (Mann et al., 1999), individuals with a certain vulnerability (“diathesis”) to SB would develop SB under the influence of psychological crises or mental disorders. According to the interpersonal theory of suicide proposed by Joiner (2005), the main factors of SB include suicidal desire (due to high levels of adherence and impaired affiliation) and the ability to perform suicide (the sum of genetic, epigenetic and environmental factors) (Ludwig et al., 2017). According to twin studies, SB showed heritability ranging from 21 to 50 %, while population studies suggested up to 55 % (Roy, Dwivedi, 2017). Molecular-genetic methods are used to study the genetic susceptibility to SB including controlled analyses of associations and linkages, micromatrix analyses of gene

expression and genome-wide analyses of associations. An important role in SB development belongs to epigenetic factors, since they mediate environmental effects on the degree of phenotypic manifestation of genetic susceptibility to pathology (Tsai et al., 2011). The brain obtained from individuals who died from suicide represents a unique object for direct study of changes in genomic processes in SB (Almeida, Turecki, 2016).

The role of genetic factors in the development of suicidal behavior

The scientific literature presents data on the associations of SB with more than 200 genes including genes associated with changes in the functioning of a number of neurobiological systems. These include genes of the serotonergic, noradrenergic and hypothalamic-pituitary-adrenal (HPA) systems (Lutz et al., 2017). According to functional studies, SB is characterized by decreased levels of serotonin metabolites and changes in the number of receptors and serotonin transporter levels in the brain (Chandley, Ordway, 2012). To date, numerous studies on genetic predisposition to SB have demonstrated association of SB risk with serotonergic system genes such as serotonin transporter (*SLC6A4*), tryptophan hydroxylase (*TPH*), and serotonin receptor 1A (*5-HT1A*) (Bach, Arango, 2012). In 2018, a meta-analysis of 45 studies was conducted, confirming the association of a low-expressing S allele in the serotonin transporter gene (*SLC6A4*) with an increased risk of developing SB (Fanelli, Serretti, 2018).

A number of authors have identified the involvement of the noradrenergic system in SB development. For example, a decreased number of noradrenergic neurons in the locus coeruleus, higher β 2-adrenergic and lower α -adrenergic receptor binding in the cortex, reduced levels of noradrenaline metabolites in the cerebrospinal fluid were detected in SB patients (Mann, Currier, 2010). Therefore, impaired noradrenergic functioning promotes SB development. At the same time, antidepressants acting on the noradrenaline transporter, α 2-adrenoreceptors and other stress-sensitive targets (glutamate transporter and receptors, GABA receptors) can reduce the risk of suicide (Chandley, Ordway, 2012).

A great importance in the phenotypic manifestation of SB belongs to stressful events in childhood, which mediate changes in the functioning of monoaminergic and HPA systems in ontogenesis (Mann, Currier, 2010). The important role of stress in SB development is confirmed by the involvement of HPA-axis genes in the development of this pathology. For

instance, an association of the C-allele *FKBP5* polymorphisms 3800373 with SB was determined. The *FKBP5* gene encodes FK506 binding protein involved in the regulation of HPA-axis activity via its binding to glucocorticoid receptors (Fudalej et al., 2015). According to functional data, a lower level of corticotropin-releasing hormone receptor, one of the most important components of the HPA axis, was detected in the prefrontal cortex of individuals with completed suicide (Mann, Currier, 2010). The SB development is also influenced by changes in the genes of the polyamine system: *SAT* (spermidine/spermine N1-acetyltransferase) and *OATL1* (ornithine aminotransferase like-1) (Fiori et al., 2010). However, the involvement of many genes with a small effect was noted. A variety of interactions between the proteins encoded by these genes is also shown. In addition, heterogeneity in the clinical groups of patients with SB was shown in various studies. These problems indicate insufficient data on SB etiopathogenesis. To solve this problem, a number of recent studies were conducted demonstrating promising results by interpreting the missing link between the environment and the genome via epigenetic factors (Roy, Dwivedi, 2017).

The relationship of epigenetic factors with the expression of genes associated with suicidal behavior

The epigenome mediates the gene $\times$ environment interaction, including the effect of adverse life situations (Schneider et al., 2015). An epigenetic approach examines the effect of environmental stimuli such as stressors, life adversity, and various biological processes on the genome. Epigenetic factors include DNA methylation, histone modifications, RNA interference (RNAi) using non-coding RNA (ncRNA) and changes in the nucleus organization. These epigenetic modifications allow the genome to respond and adapt to internal and external factors through variations in gene expression (Bani-Fatemi et al., 2015).

According to epigenetic studies of SB, the pathogenesis of this disease is based on an impaired plasticity of the neuronal pathways with the brain's inability to provide an appropriate adaptive response to environmental stimuli. For instance, individuals with SB are characterized by changes in the expression of genes crucial for synaptic and structural plasticity (Dwivedi, 2018). A number of researchers have demonstrated the role of changes in the expression of genes of the cytokine system and polyamines in SB pathogenesis. In particular, a significantly higher level of expression of tumor necrosis factor alpha (TNF- α) and miR-19a-3n was observed in the prefrontal cortex of individuals with completed suicide compared to the control. This miRNA has a targeted effect on TNF- α (Wang et al., 2018). The role of epigenetic factors in the regulation of polyamine gene (*SAT1*) (spermidine/spermine N1-acetyltransferase) expression in the prefrontal cortex of individuals with completed suicide has been identified (Fiori, Turecki, 2011).

Early-life adversity (ELA) in childhood as a stress-related component represents an important risk factor for developing SB in adults, even if much time will have passed since exposure to stress. ELA includes child abuse, neglect, loss of parents, low socioeconomic status. Although only negative effects of ELA are often assumed, it is important to note that

stress effects are not deterministic and can cause conflicting effects in adulthood, and in some cases can even result in increased stress sensitivity. However, most studies focus on the negative effects of ELA. As a result of a number of studies, it was concluded that ELA long-term effects might be due to changes in the epigenetic landscape due to impaired DNA methylation, post-translational modification of histones and the expression of non-coding RNA (Burns et al., 2018). The study of DNA methylation demonstrated an increased methylation of 97 % of 1000 differentially methylated regions, including functional categories of genes highly expressed in the brain (*APLP2*, *BDNF*, *HTR1A*, *NUAK1*, *PHACTR3*, *MSMP*, *SLC6A4*, *SYN2*, and *SYNE*), in the cerebral cortex of individuals with completed suicide compared to the control group (Schneider et al., 2015).

Changes in DNA methylation and histone modification

Nowadays the majority of studies on the role of epigenetic factors in SB development are focused on resistant markers, such as DNA methylation and histone modifications. DNA methylation is a dynamic process that occurs throughout ontogenesis, even in postmitotic cells, such as neurons. Traditionally, DNA methylation is defined as the addition of a methyl group to the fifth carbon of the cytosine base (5mC) in CpG islands of gene promoters in mammalian genomes, which is functionally related to gene silencing. Unlike 5mC, hydroxymethylated DNA (5hmC) often found in the genes has a global positive effect on gene expression. TET 2/3 enzymes play a role in the formation of 5hmC (ten-eleven translocation of methylcytosine dioxygenases 2 and 3). This process is described in neural cells. The brain is characterized by a high level of methylation of cytosine outside the CG dinucleotides, the so-called CH-methylation, especially characteristic of the first year of life (Burns et al., 2018).

A number of studies have been carried out on the association of changes in epigenetic factors with SB development in individuals who have experienced severe stress in childhood. In this context the effect of changes in stress response systems, mainly the HPA axis, which is programmed under the influence of environmental factors in early childhood, is actively analyzed. Changes in the HPA axis are associated with an increased risk of suicide (Turecki, 2014). Evidence of a significant role of environmental factors in active demethylation in childhood includes methylation differences in exon 1C of the glucocorticoid receptor (*GR*) gene in the brain of individuals with completed suicide with ELA history (Turecki, 2014). SB patients who experienced early-life stress demonstrated hypermethylation in two CpG islands in the promoter region of the neuron-specific glucocorticoid receptor gene (*NR3C1*), which resulted in suppression of gene expression in the brain of individuals with completed suicide compared to the control (McGowan et al., 2009). Serotonergic system genes, in particular, the serotonin transporter gene (*5-HTT*), are also regulated by individual reaction to stress. Namely, the important role of stress mediating the association between modifications in the promoter region of the *5-HTT* gene and the risk of developing SB was suggested to be caused by changes in *5-HTT* gene methylation (Jimenez-Trevino et al., 2017).

Recently, the expression and methylation of polyamine system genes have been actively studied. Stress-mediated impairments in various components of the polyamine system were found in cortical and subcortical structures in individuals with SB with marked changes in the epigenetic regulation of gene expression (Turecki, 2014). The genome-wide DNA methylation studies demonstrated significant site-specific differences in the methylation patterns of the promoter regions of the polyamine system genes including *OAZ1* (ornithine decarboxylase antienzyme 1), *OAZ2* (ornithine decarboxylase antienzyme 2), *AMD1* (S-adenosylmethionine decarboxylase 2), *ARG2* (arginase 2) (Gross et al., 2013), and *SKA2* (spindle and kinetochore associated complex subunit 2) (Guintivano et al., 2014; Pandey et al., 2016; Olie, Gourtet, 2017) in the brain of individuals with completed suicide compared to the control. The methylation pattern of the *SKA2* gene is specific to SB and can be used as a biomarker for determining suicidal risk (Sadeh et al., 2016).

A large number of studies on the epigenetic regulation of SB were focused on changes in methylation in the gene encoding neurotrophic brain-derived factor (*BDNF*) (Kang et al., 2018). Moreover, SB patients were characterized by a significant increase in DNA methylation in the promoter and non-coding exon 4 of the *BDNF* gene, while a hypermethylation was observed in promoter IV of this gene (Keller et al., 2010, 2011). Noteworthy, the association of the hypermethylated *BDNF* gene with SB risk was observed independently of the presence of potential covariates or a particular genotype (Kim et al., 2014). Moreover, distinct changes in *BDNF* expression are also considered to be a risk factor for developing SB in the elderly (Kim et al., 2014). Considering the data obtained, it is suggested to use the methylation pattern of the *BDNF* gene as a marker of a history of suicidal attempts, as well as to predict the possible inefficacy in SB therapy (Kang et al., 2013). Since the effect of BDNF is present due to its binding to the tropomyosin tyrosine kinase receptor *TrkB-T1*, several authors have reported a decrease in the expression of the *TrkB-T1* gene, which is caused by higher methylation in the promoter and the 3'-untranslated regions (UTR) of this gene in the frontal lobe in patients with completed suicide (Ernst et al., 2009b; Maussion et al., 2014).

Some studies have shown a possible role of modified expression of such genes as *MPP4* (membrane palmitoylated protein 4), nucleoporin (*NUP133*), a member of the TRE2/BuB2/CDC16 family of domains (*TBC1D16*), the alpha1 subunit of the gamma-aminobutyric acid receptor (*GABRA1*) in SB development. In particular, a total decrease in methylation in the 5'-UTR of the *MPP4* gene and in intron 3 of the *TBC1D16* gene together with increased methylation in exon 1 of the *NUP133* gene was observed in patients with bipolar disorder and comorbid SB compared to the control group (Jeremian et al., 2017). In addition, hypermethylation of CpG islands in the promoter of the *GABRA1* gene was associated with changes in mRNA expression of the DNA methyltransferase gene (*DNMT*) in the brain of individuals with completed suicide (Poulter et al., 2008).

Histone modifications also make a significant contribution to the regulation of expression in SB. Data were obtained on the suppression of gene expression of connexins 30 and 43 (*Cx30* and *Cx43*) in the brain of individuals with completed

suicide due to histone methylation (Nagy et al., 2017). Namely, changes in astrocyte interaction occurring due to the channels formed mainly by connexins 30 and 43 are largely regulated by histone modifications in the *Cx30* and *Cx43* genes. The role of histone modifications in the regulation of tropomyosin tyrosine kinase receptor (*TrkB-T1*) (Ernst et al., 2009a) and ornithine decarboxylase antienzyme genes (*OAZ*), which is involved in the synthesis of polyamines, was reported (Fiori et al., 2012). In the former study, a decrease in *TrkB-T1* gene expression was observed due to methylation in the third histone (H3K27) (Ernst et al., 2009a). The latter case, in contrast, reported activated *OAZ* gene expression due to enhanced levels of H3K4me3, a marker of transcriptionally active chromatin (Fiori et al., 2012).

The role of long non-coding RNAs

Long non-coding RNA (lncRNA) represents RNA molecules longer than 200 bp with low protein-coding potential. The lncRNAs are classified mainly on the basis of their interaction with known genes. For example, lncRNAs can be antisense, sense, overlapping, intronic, and intergenic transcripts. lncRNAs are characterized by tissue-specific expression and mediate important biological functions by regulating the functioning of protein-coding genes. It was found that the lncRNA genes formed various isoforms with different functions manifesting *in-cis* and *in-trans* regulatory mechanisms. Therefore, one lncRNA gene can control the functioning of several distal target genes (Zhou et al., 2018). It is important to note that transposable elements (TEs) represent the sources of over 41 % of the functional lncRNA domains (Johnson, Guigo, 2014) and are considered to be stress-sensitive elements (Wheeler, 2013), which in site-specific transposition activate the stress response genes (Feng et al., 2013). Moreover, TEs can serve directly as sources of dsRNA genes, whose transcripts regulate the differentiation of stem cells (Lu et al., 2014). This observation is caused by a significant role of lncRNA in human brain functioning and is determined by lncRNA activity in the hippocampus (simultaneously with TEs expression) during neurogenesis. For example, *lncRNA2393* expression promotes the maturation of neural stem cells in the dentate gyrus (Deng et al., 2017).

More than 14,000 lncRNA genes have been identified in the human genome, at least 67 % of the mature transcripts of which consist of TE sequences, while many of them consist entirely of TE (Kapusta, Feschotte, 2014). TEs represent important sources of epigenetic regulation (Mustafin, Khusnutdinova, 2017), thus providing their study as a promising direction for identifying the mechanisms of SB development. A differential expression of six lncRNAs (*TCONS_00019174*, *ENST00000566208*, *NONHSAG045500*, *ENST00000517573*, *NONHSAT034045*, and *NONHSAT142707*) was observed in peripheral blood leukocytes of patients with SB and MDD (Cui et al., 2017), while 23 different lncRNAs were differentially expressed in the brains of individuals with completed suicide (Zhou et al., 2018) compared to the control group. Protein-encoding genes localized distally from lncRNAs identified in these studies are involved in the organization of the cytoskeleton and plasmatic membrane, cell adhesion, DNA binding and regulation of dendrite development (Zhou

et al., 2018). In another study, the association of *LOC285758* lncRNA expression with SB development was determined. This lncRNA represents an antisense transcript of the region flanking the intragenic CpG island of the *MARCKS* gene (myristoylated alanine-rich C-kinase substrate), of which expression is suppressed during prolonged lithium administration (Punzi et al., 2014).

The role of miRNAs

Recently, a great importance in the study of SB development has been attached to the role of small ncRNAs controlling gene expression. The miRNAs switching highly significant regulators of neuronal plasticity and higher nervous activity are the most studied to date (Dwivedi, 2018). For a number of miRNAs, a relationship with brain functioning was determined. For example, miR-16 affects the expression of *SERT*, miR-18a and miR-124a bind to the 3'-UTR of the *GR* gene, and miR-34a controls the effects of lithium and valproate by interacting with *GRM7*. The miR-96 and miR-510 were demonstrated to inhibit translation of *5-HT1B* and *5-HT3E* receptor subunits, while miR-124-1 is involved in serotonin-induced synaptic transmission via regulating *CREB* (cAMP response element-binding protein). In addition, it was shown that miR-30a-5p and miR-195 target the 3'-UTR of the *BDNF* gene in different brain regions (miR-30a in the third layer of the pyramidal neurons of the prefrontal cortex). The expression of miR-134 and miR-183, which target the splicing factor *SC35*, is enhanced by acute stress, miR-280 and miR-289 regulate the synthesis of synaptic proteins by binding to the *CaMKIIa* sites, miR-134 inhibits the translation of *Limk1* in the dendrites of the hippocampus; miR-137 regulates the proliferation of neuronal stem cells by affecting the transcription factor Sox2 (Serafini et al., 2012).

Studies of changes in miRNA expression in SB allowed us to detect significant variations in the level of their expression in the prefrontal cortex of individuals with completed suicide compared to healthy donors. At the same time, a significant decrease in the expression of 21 different miRNAs involved in the regulation of cell growth and differentiation was shown. The targets for these miRNAs include transcription factors (in particular, E2F1, E2F6, BACH1, SP1, HOXA5, and RUNX1) and other nuclear proteins. At the same time, the *VEGFA* gene associated with developing depression in both human and animals appears to be a target for four different miRNAs (miR-20b, 20a, 34a, 34b*) (Smalheiser et al., 2012). Differential expression of thirteen different miRNAs was detected in the brain of individuals with completed suicide. Among these miRNAs, an enhanced expression was detected in miR-17-5p, miR-20b-5p, miR-106a-5p, miR-330-3p, miR-541-3p, miR-582-5p, miR-890, miR-99b-3p, miR-550-5p, and miR-1179. A reduced expression was determined in miR-409-50, let-7g-3p, and miR-1197. The analysis of the integrated gene regulatory network based on target genes of these miRNAs revealed various associations with mental disorders, including MDD and affective ones. These psychopathologies are assumed to be the most important risk factors for SB. The mapping of cellular pathways mediated by miRNA activity identified a total modification in the cellular signaling causing SB development (Roy et al., 2017).

Because changes in the metabolic pathways of polyamine system enzymes are involved in SB development, miRNAs interacting with the *SAT1* and *SMOX* genes were examined. The relationship between miRNAs and polyamine gene expression in SB was reported and the mechanism of post-transcriptional suppression of *SAT1* and *SMOX* gene activity was demonstrated. Individuals with completed suicide were characterized by a significant increase in the brain levels of miR-34c-5p, miR-139-5p, miR-195, and miR-320c, which target the 3'-UTRs of the *SAT1* and *SMOX* genes (Lopez et al., 2014). An enhanced expression of Hsa-miR-185 and Hsa-miR-491-3p causing suppression of the *TrkB-T1* gene was determined in the prefrontal cerebral cortex of individuals with completed suicide. A target binding site in 3'-UTR of the *TrkB-T1* gene was identified for Hsa-miR-185 (Maussion et al., 2012). Recently, considerable attention was focused on the identification of miRNAs associated with ELA in the context of SB development. ELA affects the activity of various miRNAs genes during brain development. For example, differences in the expression of miR-9, miR-29a, miR-124, and miR-132 were observed in the prefrontal cortex of rats aged 14 days when separated from the mother. The expression of miR-124 and miR-132 was suppressed at the 60th day of postnatal development, which indicates stable changes in miRNAs caused by ELA. At the same time, GR activation inhibits miR-132 expression, which suppresses expression of the *BDNF* gene previously associated with SB (Dwivedi, 2018).

Prospects for epigenetic studies of suicidal behavior

According to the stress diathesis model, suicide is positioned as a result of interactions between environmental stressors and susceptibility to SB, regardless of the present mental disorder. The genetic and epigenetic changes detected in the brain of individuals with completed suicide provide a basis for the possible neurobiological screening of SB patients to prevent suicide (van Heeringen, Mann, 2014). Among epigenetic factors, the study of lncRNAs suggested the mechanisms of action of certain pharmaceuticals used in SB therapy. For example, long-term use of lithium used for SB treatment suppresses the expression of the *MARCKS* gene, the expression of the antisense lncRNA of which is associated with SB (Punzi et al., 2014). These results promote further research of lncRNAs, which makes it possible to develop effective methods for SB prevention and therapy. Analysis of lncRNAs differentially expressed in SB (*TCONS_00019174*, *ENST00000566208*, *NONHSAG045500*, *ENST00000517573*, *NONHSAT034045*, and *NONHSAT142707*) has been proposed as a potential diagnostic and therapeutic SB biomarker for the prevention of suicidal attempts in MDD patients (Cui et al., 2017). The changes in miRNA expression identified in the brain of individuals with completed suicide can represent the basis for both clarifying SB pathogenesis and developing a targeted SB therapy (Maussion et al., 2012; Smalheiser et al., 2012; Roy, Dwivedi, 2017).

The analysis of the methylation pattern of *BDNF* (Kang et al., 2013; Kim et al., 2014) and the *SKA2* genes (Sadeh et al., 2016) made it possible to propose them as epigenetic biomarkers of SB. For SB therapy it was suggested to use the histone

deacetylase inhibitor tetrapeptide FK228, which is able to enhance transcription of the *Rap1* and *ERK1/2* genes known to be reduced in the hippocampus of SB individuals. Rap-1 (Ras-proximate-1) is a short nucleotide triphosphate binding protein expressed in neurons of the cerebral cortex involved in dendrites branching and growth. A significant decrease in Rap-1 mRNA expression was detected in the prefrontal cortex and hippocampus of depressed individuals and those with completed suicide compared to the control group. Reduced brain activity of Rap-1 was characteristic of individuals with completed suicide (Emanuele, 2007).

Nowadays, genome editing technologies allow the functional importance of specific epigenetic modifications and gene regulation to be studied directly and the disturbed epigenetic landscape to be remodeled due to the reversibility of epigenetic modifications. One of the most successful methods for epigenetic editing is CRISPR-Cas9, which allows specific changes to be introduced to DNA methylation (Vojta et al., 2016).

Conclusion

The study of the role of epigenetic factors in the development of suicidal behavior is a modern and promising method for determining reversible changes in the brain of patients. The association of specific expression of genes that play a role in the development of suicidal behavior with specific changes in DNA methylation, modifications of histone and levels of non-coding RNA has been proved. The results obtained indicate the prospects for the development of targeted therapy methods for this serious and socially significant pathology using epigenetic factors. The most successful objects of epigenetic exposure are non-coding RNAs, the use of which has already been started in clinical practice. In addition, it can be assumed that the correction of patients' lifestyles and psychotherapy with the study of the role of changes in the epigenetic regulation of the brain can be an effective treatment for suicidal behavior. This conclusion is due to the fact that epigenetic factors are modulated by environmental, especially stressful, influences.

References

- Almeida D., Turecki G. A slice of the suicidal brain: what have postmortem molecular studies taught us? *Curr. Psychiatry Rep.* 2016;18:98.
- Bach H., Arango V. Ch. 2. Neuroanatomy of Serotonergic Abnormalities in Suicide. In: Dwivedi Y. (Ed.) *The Neurobiological Basis of Suicide*. (Ser. Frontiers in Neuroscience). Boca Raton (FL): CRC Press/Taylor & Francis, 2012.
- Bani-Fatemi A., Howe A.S., De Luca V. Epigenetic studies of suicidal behavior. *Neurocase*. 2015;21:134-143.
- Burns S.B., Szyszkowicz J.K., Luheshi G.N., Lutz P.E., Turecki G. Plasticity of the epigenome during early-life stress. *Semin. Cell. Dev. Biol.* 2018;77:115-132.
- Chandley M.J., Ordway G.A. Ch. 3. Noradrenergic Dysfunction in Depression and Suicide. In: Dwivedi Y. (Ed.) *The Neurobiological Basis of Suicide*. (Ser. Frontiers in Neuroscience). Boca Raton (FL): CRC Press/Taylor & Francis, 2012.
- Cui X., Niu W., Kong L., He M., Jiang K., Chen S., Zhong A., Li W., Lu J., Zhang L. Long noncoding RNA expression in peripheral blood mononuclear cells and suicide risk in Chinese patients with major depressive disorder. *Brain Behav.* 2017;7:e00711.
- Deng B., Cheng X., Li H., Qin J., Tian M., Jin G. Microarray expression profiling in the denervated hippocampus identified long noncoding RNAs functionally involved in neurogenesis. *BMC Mol. Biol.* 2017;18(1):15.
- Dwivedi Y. MicroRNAs in depression and suicide: recent insights and future perspectives. *J. Affect. Disord.* 2018;240:146-154.
- Emanuele E. The histone deacetylase inhibitor FK228 may have therapeutic usefulness to prevent suicidal behavior via upregulation of the guanosine triphosphatase Rap-1. *Med. Hypotheses*. 2007;68:451-452.
- Ernst C., Chen E.S., Turecki G. Histone methylation and decreased expression of TrkB.T1 in orbital frontal cortex of suicide completers. *Mol. Psychiatry*. 2009a;14:830-832.
- Ernst C., Deleva V., Deng X., Sequeira A., Pomarenski A., Klempan T., Ernst N., Quirion R., Gratton A., Szyf M., Turecki G. Alternative splicing, methylation state, and expression profile of tropomyosin-related kinase B in the frontal cortex of suicide completers. *Arch. Gen. Psychiatry*. 2009b;66:22-32.
- Fanelli G., Serretti A. The influence of the serotonin transporter gene 5-HTTLPR polymorphism on suicidal behaviors: a meta-analysis. *Prog. Neuropsychopharmacol. Biol. Psychiatry*. 2018;88:375-387.
- Feng G., Leem Y.E., Levin H.L. Transposon integration enhances expression of stress response genes. *Nucleic Acids Res.* 2013;41(2):775-789.
- Fiori L.M., Gross J.A., Turecki G. Effects of histone modifications on increased expression of polyamine biosynthetic genes in suicide. *Int. J. Neuropsychopharmacol.* 2012;15:1161-1166.
- Fiori L.M., Turecki G. Epigenetic regulation of spermidine/spermine N¹-acetyltransferase (SAT1) in suicide. *J. Psychiatr. Res.* 2011;45:1229-1235.
- Fiori L.M., Wanner B., Jomphe V., Croteau J., Vitaro F., Tremblay R.E., Bureau A., Turecki G. Association of polyaminergic loci with anxiety, mood disorders, and attempted suicide. *PLoS One*. 2010;5:e15146.
- Fudalej S., Kopera M., Wolynczyk-Gmaj D., Fudalej M., Krajewski P., Wasilewska K., Szymanski K., Chojnicka I., Podgorska A., Wojnar M., Ploski R. Association between *FKBP5* functional polymorphisms and completed suicide. *Neuropsychobiology*. 2015;72:126-131.
- Gross J.A., Fiori L.M., Labonte B., Lopez J.P., Turecki G. Effects of promoter methylation on increased expression of polyamine biosynthetic genes in suicide. *J. Psychiatr. Res.* 2013;47:513-519.
- Guintivano J., Brown T., Newcomer A., Jones M., Cox O., Maher B.S., Eaton W.W., Payne J.L., Wilcox H.C., Kaminsky Z.A. Identification and replication of a combined epigenetic and genetic biomarker predicting suicide and suicidal behaviors. *Am. J. Psychiatry*. 2014;171:1287-1296.
- Jeremian R., Chen Y.A., De Luca V., Vincent J.B., Kennedy J.L., Zai C.C., Strauss J. Investigation of correlations between DNA methylation, suicidal behavior and aging. *Bipolar Disord.* 2017;19:32-40.
- Jiménez-Treviño L., Saiz P.A., Garcia-Portilla M.P., Blasco-Fontecilla H., Carli V., Iosue M., Jaussent I., López-Castroman J., Vaguero-Lorenzo C., Sarchiapone M., Baca-Garcia E., Courtet P., Bobes J. 5-HTTLPR-brain-derived neurotrophic factor (BDNF) gene interactions and early adverse life events effect on impulsivity in suicide attempters. *World J. Biol. Psychiatry*. Epub Oct 2017; Publ. 2019;20(2):137-149.

- Johnson R., Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA*. 2014;20(7):959-976.
- Joiner T.E. *Why People Die by Suicide*. Cambridge, MA: Harvard Univ. Press, 2005.
- Kang H.J., Bae K.Y., Kim S.W., Shin I.S., Hong Y.J., Ahn Y., Jeong M.H., Yoon J.S., Kim J.M. BDNF methylation and suicidal ideation in patients with acute coronary syndrome. *Psychiatry Investig.* 2018;15(11):1094-1097. DOI 10.30773/pi.2018.09.20.
- Kang H.J., Kim J.M., Lee J.Y., Kim S.Y., Bae K.Y., Kim S.W., Shin I.S., Kim H.R., Shin M.G., Yoon J.S. BDNF promoter methylation and suicidal behavior in depressive patients. *J. Affect. Disord.* 2013;151:679-685.
- Kapusta A., Feschotte C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. *Trends Genet.* 2014;30(10):439-452.
- Keller S., Sarchiapone M., Zarrilli F., Tomaiuolo R., Carli V., Angrisano T., Videtic A., Amato F., Pero R., di Giannantonio M., Losue M., Lembo F., Castaldo G., Chiarriotti L. TrkB gene expression and DNA methylation state in Wernicke area does not associate with suicidal behavior. *J. Affect. Disord.* 2011;135:400-404.
- Keller S., Sarchiapone M., Zarrilli F., Videtic A., Ferraro A., Carli V., Sacchetti S., Lembo F., Angiolillo A., Jovanovic N., Pisanti F., Tomaiuolo R., Monticelli A., Balazic J., Roy A., Marusic A., Coccozza S., Fusco A., Bruni C.B., Castaldo G., Chiarriotti L. Increased BDNF promoter methylation in the Wernicke area of suicide subjects. *Arch. Gen. Psychiatry.* 2010;67:258-267.
- Kim J.M., Kang H.J., Bae K.Y., Kim S.W., Shin I.S., Kim H.R., Shin M.G., Yoon J.S. Association of BDNF promoter methylation and genotype with suicidal ideation in elderly Koreans. *Am. J. Geriatr. Psychiatry.* 2014;22:989-996.
- Lopez J.P., Fiori L.M., Gross J.A., Labonte B., Yerko V., Mechawar N., Turecki G. Regulatory role of miRNAs in polyamine gene expression in the prefrontal cortex of depressed suicide completers. *Int. J. Neuropsychopharmacol.* 2014;17:23-32.
- Lu X., Sachs F., Ramsay L., Jacques P.E., Goke J., Bourque G., Ng H.H. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.* 2014;21(4):423-425.
- Ludwig B., Roy B., Wang Q., Birur B., Dwivedi Y. The life span model of suicide and its neurobiological foundation. *Front. Neurosci.* 2017;11:74.
- Lutz P.E., Mechawar N., Turecki G. Neuropathology of suicide: recent findings and future directions. *Mol. Psychiatry.* 2017;22:1395-1412.
- Mann J.J., Currier D.M. Stress, genetics and epigenetic effects on the neurobiology of suicidal behavior and depression. *Eur. Psychiatry.* 2010;25:268-271.
- Mann J.J., Waternaux C., Haas G.L., Malone K.M. Toward a clinical model of suicidal behavior in psychiatric patients. *Am. J. Psychiatry.* 1999;156(2):181-189.
- Mausson G., Yang J., Suderman M., Diallo A., Nagy C., Aronovitz M., Mechawar N., Turecki G. Functional DNA methylation in a transcript specific 3'UTR region of TrkB associates with suicide. *Epigenetics.* 2014;9:1061-1070.
- Mausson G., Yang J., Yerko V., Barker P., Mechawar N., Ernst C., Turecki G. Regulation of a truncated form of tropomyosin-related kinase B (TrkB) by Hsa-miR-185* in frontal cortex of suicide completers. *PLoS One.* 2012;7:e39301.
- McGowan P.O., Sasaki A., D'Alessio A.C., Dymov S., Labonte B., Szyf M., Turecki G., Meaney M.J. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat. Neurosci.* 2009;12:342-348.
- Mustafin R.N., Khusnutdinova E.K. Non-coding parts of genomes as the basis of epigenetic heredity. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding.* 2017;21(6):742-749. DOI 10.18699/VJ17.30-o. (in Russian)
- Nagy C., Torres-Platas S.G., Mechawar N., Turecki G. Repression of astrocytic connexins in cortical and subcortical brain regions and prefrontal enrichment of H3K9me3 in depression and suicide. *Int. J. Neuropsychopharmacol.* 2017;20:50-57.
- Nock M.K., Green J.G., Hwang I., McLaughlin K.A., Sampson N.A., Zaslavsky A.M., Kessler R.C. Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: results from the National Comorbidity Survey Replication Adolescent Supplement. *JAMA Psychiatry.* 2013;70:300-310.
- Olié E., Courtet P. Genetics and epigenetics of suicidal behaviors. *Biol. Aujourd'hui.* 2017;211:93-96.
- Pandey G.N., Rizavi H.S., Zhang H., Bhaumik R., Ren X. The expression of the suicide-associated gene SKA2 is decreased in the prefrontal cortex of suicide victims but not of nonsuicidal patients. *Int. J. Neuropsychopharmacol.* 2016;19(8):1-10. DOI 10.1093/ijnp/pyw015.
- Poulter M.O., Du L., Weaver I.C., Palkovits M., Faludi G., Merali Z., Szyf M., Anisman H. GABA_A receptor promoter hypermethylation in suicide brain: implications for the involvement of epigenetic processes. *Biol. Psychiatry.* 2008;64:645-652.
- Punzi G., Ursini G., Shin J.H., Kleinman J.E., Hyde T.M., Weinberger D.R. Increased expression of MARCKS in post-mortem brain of violent suicide completers is related to transcription of a long, noncoding, antisense RNA. *Mol. Psychiatry.* 2014;19:1057-1059.
- Roy B., Dwivedi Y. Understanding epigenetic architecture of suicide neurology: a critical perspective. *Neurosci. Biobehav. Rev.* 2017;72:10-27.
- Roy B., Dwivedi Y. Understanding the neuroepigenetic constituents of suicide brain. *Prog. Mol. Biol. Transl. Sci.* 2018;157:233-262.
- Roy B., Wang Q., Palkovits M., Faludi G., Dwivedi Y. Altered miRNA expression network in locus coeruleus of depressed suicide subjects. *Sci. Rep.* 2017;7:4387.
- Sadeh N., Wolf E.J., Logue M.W., Hayes J.P., Stone A., Griffin L.M., Schichman S.A., Miller M.W. Epigenetic variation at SKA2 predicts suicide phenotypes and internalizing psychopathology. *Depress. Anxiety.* 2016;33:308-315.
- Serafini G., Pompili M., Innamorati M., Giordano G., Monteboni F., Sher L., Dwivedi Y., Girardi P. The role of microRNAs in synaptic plasticity, major affective disorders and suicidal behavior. *Neurosci. Res.* 2012;73:179-190.
- Schneider E., El Hajj N., Muller F., Navarro B., Haaf T. Epigenetic dysregulation in the prefrontal cortex of suicide completers. *Cytogenet. Genome Res.* 2015;146:19-27.
- Smalheiser N.R., Lugli G., Rizavi H.S., Torvik V.I., Turecki G., Dwivedi Y. MicroRNA expression is down-regulated and reorganized in prefrontal cortex of depressed suicide subjects. *PLoS One.* 2012;7:e33201.
- Tsai S.J., Hong C.J., Liou Y.J. Recent molecular genetic studies and methodological issues in suicide research. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* 2011;35:809-917.

- Turecki G. Epigenetics and suicidal behavior research pathways. *Am. J. Prev. Med.* 2014;47:144-151.
- van Heeringen K., Mann J.J. The neurobiology of suicid. *Lancet Psychiatry.* 2014;1:63-72.
- Vojta A., Dobrinic P., Tadic V., Bockor L., Korac P., Julg B., Klasic M., Zoldos V. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* 2016;44:5615-5628.
- Wang Q., Roy B., Turecki G., Shilton R.C., Dwivedi Y. Role of complex epigenetic switching in tumor necrosis factor- α up-regulation in the prefrontal cortex of suicide subjects. *Am. J. Psychiatry.* 2018;175:262-274.
- Wheeler B.S. Small RNAs, big impact: small RNA pathways in transposon control and their effect on the host stress response. *Chromosome Res.* 2013;21:587-600.
- Zhou Y., Lutz P.E., Wang Y.C., Ragoussis J., Turecki G. Global long non-coding RNA expression in the rostral anterior cingulate cortex of depressed suicides. *Transl. Psychiatry.* 2018;8:224.

ORCID ID

R.N. Mustafin orcid.org/0000-0002-4091-382X
A.V. Kazantseva orcid.org/0000-0002-3744-8058
R.F. Enikeeva orcid.org/0000-0002-4301-5283
Yu.D. Davydova orcid.org/0000-0003-3508-4710
S.B. Malykh orcid.org/0000-0002-3786-7447
E.K. Khusnutdinova orcid.org/0000-0003-2987-3334

Acknowledgements. This review was supported by the Russian Science Foundation, project 19-78-30021 "Risks of suicidal behavior development: genetic and epigenetic mechanisms", and the Russian Foundation for Basic Research, project 17-29-02195 "Genomics of aggressive and depressive human behavior".

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

Received November 26, 2018. Revised April 10, 2019. Accepted April 10, 2019.

Genetic variants on chromosome 19 (rs439401 and rs4420638) are associated with obesity and high blood pressure in the Algerian population

H. Boulouar^{1, 2, 3}✉, S. Mediene Benchechor^{2, 4}, H. Ouhaibi Djellouli², S. Larjam Hetraf², L. Houti⁵, I. Hammani-Medjaoui⁶

¹ Département de Médecine, Faculté de Médecine "Dr Benzerdjeb Benaouda", Université Aboubekr Belkaid, Tlemcen, Algérie

² Laboratoire de Génétique Moléculaire et Cellulaire, Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, Oran, Algérie

³ Laboratoire de Recherche CancerLab No. 30, Faculté de Médecine "Dr Benzerdjeb Benaouda", Université Aboubekr Belkaid, Tlemcen, Algérie

⁴ Département de Biotechnologie, Faculté des Sciences de la Nature et de la Vie, Université Oran 1 Ahmed Ben Bella, Oran, Algérie

⁵ Faculté de Médecine, Université d'Oran 1; Laboratoire des Systèmes d'Information en Santé, Université d'Oran 1, Oran, Algérie

⁶ Caisse Nationale des Assurances Sociales des travailleurs salariés, Clinique Spécialisée en Orthopédie et Rééducation des Victimes des Accidents de Travail, Oran, Algérie

✉ e-mail: boulouar.houssam@gmail.com; houssam.boulouar@mail.univ-tlemcen.dz

Metabolic syndrome (MetS) represents a combination of at least three primary metabolic abnormalities from among obesity, hyperglycemia, dyslipidemia, and high blood pressure, once combined, they increase significantly the cardiovascular risk. The *APOE* gene is considered as a genetic risk factor for cardiovascular diseases, it has been linked to MetS or its traits in several populations. Our study aimed to analyze the association of three *APOE* gene polymorphisms with MetS risk and its components in a general population sample, and to highlight the potential influence of these polymorphisms on individual susceptibility to MetS. We performed this work using a population-based, cross-sectional study of a representative sample of 787 individuals (378 men and 409 women, aged between 30 and 64 years) recruited in the city of Oran, Algeria (the ISOR Study); the subjects were genotyped for four polymorphisms, rs7412, rs429358, rs4420638 and rs439401, located in the *APOE* gene, using the KASPar technology. rs439401 showed a significant association with hypertension (HBP). The T allele confers a high risk of hypertension with an odds ratio (OR) of 1.46 (95 % CI [1.12–1.9], $p = 0.006$). rs4420638 was significantly associated with obesity in the general population. The G allele provides protection against obesity, the resulting OR is 0.48 (95 % CI [0.29–0.81], $p = 0.004$). Although *APOE* variants were not associated with the risk of MetS, the *APOE* polymorphism alleles were associated with some of the metabolic parameters in Algerian subjects. The relation of *APOE* rs439401 alleles with a HBP is likely to be indicative of a state of stress of the population. Key words: genetics; high blood pressure; diabetes; metabolic syndrome; obesity; Algerian population.

For citation: Boulouar H., Mediene Benchechor S., Ouhaibi Djellouli H., Larjam Hetraf S., Houti L., Hammani-Medjaoui I. Genetic variants on chromosome 19 (rs439401 and rs4420638) are associated with obesity and high blood pressure in the Algerian population. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2019;23(5): 608-614. DOI 10.18699/VJ19.532

Связь генетических вариантов rs439401 и rs4420638 в хромосоме 19 с ожирением и артериальной гипертензией у жителей Алжира

Х. Бульнуар^{1, 2, 3}✉, С. Медин Бенчекор^{2, 4}, Х. Оухаби Джеллоули², С. Ларжам Хетраф², Л. Хоути⁵, И. Хаммани-Меджауи⁶

¹ Медицинское отделение медицинского факультета им. д-ра Беноуды Бензерджеба Университета им. Абубакра Белькаид, Тлемсен, Алжир

² Лаборатория молекулярной и клеточной генетики Научно-технического университета им. Мухаммеда Будиафа, Оран, Алжир

³ Лаборатория по исследованию рака № 30 медицинского факультета им. д-ра Беноуды Бензерджеба Университета им. Абубакра Белькаид, Тлемсен, Алжир

⁴ Отделение биотехнологий факультета естествознания и природы 1-го Оранского университета им. Ахмеда бен Беллы, Оран, Алжир

⁵ Медицинский факультет 1-го Оранского университета и лаборатория информационных систем здравоохранения 1-го Оранского университета, Оран, Алжир

⁶ Специализированная клиника ортопедии и реабилитации жертв несчастных случаев на производстве – структура Национального фонда социального страхования наемных работников, Оран, Алжир

✉ e-mail: boulouar.houssam@gmail.com; houssam.boulouar@mail.univ-tlemcen.dz

Метаболический синдром (MetS) представляет собой комбинацию по меньшей мере трех основных показателей, ассоциированных с нарушением обмена веществ: ожирения, гипергликемии, дислипидемии и высокого артериального давления. Любое их сочетание значительно увеличивает риск развития сердечно-сосудистых заболеваний. Ген *APOE* считается генетическим фактором риска развития сердечно-сосудистых заболеваний, в некоторых популяциях он связан с метаболическим синдромом или его признаками. Настоящее исследование направлено на анализ ассоциации трех полиморфизмов гена *APOE* с риском развития метаболического синдрома и его компонентов в общей популяции, а также на выявление возможного влияния этих полиморфизмов на индивидуальную

восприимчивость к метаболическому синдрому. В рамках проекта ISOR было проведено популяционное перекрестное исследование репрезентативной выборки из 787 добровольцев (378 мужчин и 409 женщин в возрасте от 30 до 64 лет), проживающих в городе Оране (Алжир). Обследуемые были генотипированы по четырем полиморфизмам (rs7412, rs429358, rs4420638 и rs439401), расположенным в гене *APOE*, с использованием технологии KASPar. Полиморфизм rs439401 показал достоверную связь с повышенным артериальным давлением. Аллель Т обуславливает высокий риск развития гипертензии с отношением шансов (OR) равным 1.46 (95 % CI [1.12–1.9], $p = 0.006$). Полиморфизм rs4420638 достоверно коррелировал с ожирением в общей популяции. Наличие аллеля G препятствует развитию ожирения (OR = 0.48; 95 % CI [0.29–0.81], $p = 0.004$). Хотя связи между вариантами гена *APOE* и риском развития метаболического синдрома не выявлено, однако найдена корреляция между полиморфными аллелями этого гена и рядом показателей обмена веществ у обследованных алжирцев. Связь полиморфизма rs439401 в аллелях гена *APOE* с гипертензией может указывать на состояние стресса у населения. Ключевые слова: генетика; высокое артериальное давление; диабет; метаболический синдром; ожирение; алжирская популяция.

Introduction

The concept of the Metabolic Syndrome (MetS) emerged following the increase of the risk factors associated to cardiovascular diseases and diabetes (Reaven, 1988; Grundy et al., 2004). MetS represents a combination of at least three primary metabolic abnormalities among obesity, hyperglycemia, dyslipidemia, and high blood pressure, once combined, they increases significantly the cardiovascular risk (Kahn et al., 2005; Wilson et al., 2005; Hillier et al., 2006; Gami et al., 2007; Meigs et al., 2007).

In Algeria, the health network improvement led to a progressive aging of the population which allows for the emergence of abnormalities associated with aging and MetS. The TAHINA study (Epidemiological Transition And Health Impact in North Africa) conducted in 2005 showed a high prevalence of hypertension (24.9 %) and diabetes (12.2 %) in the Algerian population. Overweight has become a real public health problem, especially among women, 66.5 % are overweight and 30.1 % are obese. Cardiovascular disease and diabetes accounted for 26.1 and 4.4 % of deaths, respectively, in 2002 (Ministère de la Santé, 2007).

There have been at least six different published definitions for MetS, the most used is that of the National Cholesterol Education Program Adult Treatment Panel III (NCEPATP III) (Balkau et al., 2007). The prevalence of MetS differs according to several parameters: definition, country, sex, age, and even according to the region in the same country. In Algeria, a recent study shows that the prevalence of metabolic syndrome was 20 % in the Oran population, it was higher in women than men (25.9 vs 13.7 %) (Houti et al., 2016).

Metabolic syndrome is a multifactorial disease that implicates both environmental and genetic factors (Chuang, 2008). Given the importance of *APOE* in the metabolism of lipoproteins; indeed, *APOE* gene was identified as genetic determinants of plasma lipid and lipoprotein concentrations in Caucasian and North African populations (Wilson et al., 1994; Boulououar et al., 2013). We aimed to analyze the association of the *APOE* gene polymorphisms with MetS risk and its components in a general population sample from the city of Oran in Algeria, and to highlight the potential influence of these polymorphisms on individual susceptibility to MetS.

Material and methods

Ethical considerations. This study was granted ethics approval by the Algerian National Agency for the Development of Health Research (ATRSS exANDRS). All participants provided informed consent prior to enrolment.

Abbreviations

ANDRS – Agence Nationale De Recherche en Santé
APOE – Apolipoprotein E
ATRSS – Agence Thématique de Recherche en Science de la Santé
BMI – Body mass index
DBP – Diastolic blood pressure
DNA – Deoxyribonucleic acid
d.f. – Degree of freedom
HDL – High-density lipoprotein
ISOR – InSulino-résistance à ORan
LDL – Low-density lipoprotein
MetS – Metabolic Syndrome
SBP – Systolic blood pressure
SNP – Single nucleotide polymorphism
T2D – Type 2 Diabetes

Study population. Participants were recruited during the ISOR (InSulino-résistance à ORan) study, a population-based, cross-sectional study of a representative sample of 787 individuals (378 men and 409 women, aged between 30 and 64 years) recruited in the city of Oran, Algeria, from 2007 to 2009 (Boulououar et al., 2013).

Data collection. Data were collected using a preconceived questionnaire on socioeconomic information, physical activity (the level of physical activity was defined in quartiles as “none”, “low”, “medium” and “high” after summing exercise scores for sporting activities, walking, housework and physical activity at work), tobacco use and alcohol intake, past medical history and family history, current medications, as well as anthropomorphic characteristics including height, weight, waist circumference, hip circumference, and blood pressure. Height and weight were measured while the subject was barefoot and lightly dressed. The BMI was calculated according to the Quetelet equation. Systolic and diastolic blood pressure values (SBP and DBP, respectively) were measured on the right arm with the subject in the sitting position, using a standard mercury sphygmomanometer. Measurements were made before and after completion of the questionnaire, with an interval of at least 10 minutes. The mean value of the blood pressure readings was considered for analysis. Regarding tobacco use, participants were categorized as either smokers (i.e. individuals reporting at least one cigarette per day) or non-smoker. After a 12h overnight fast, blood was collected

aseptically via venipuncture in an EDTA tube for DNA extraction and subsequent molecular analysis, and in a heparin tube for biochemistry tests.

Metabolic syndrome diagnosis criteria. In this study, we have adopted the definition of metabolic syndrome according to the criteria of the NCEP ATP III (National Cholesterol Education Program Adult Treatment Panel III) (Alberti et al., 2005), the metabolic syndrome is diagnosed when a subject has three or more of the following risk factors:

- Abdominal obesity: waist circumference > 102/88 cm (Men/Women);
- Triglyceride level ≥ 1.50 g/l (1.69 mmol/l);
- HDL-cholesterol < 0.40/0.50 g/l (1.04/1.29 mmol/l) (Men/Women);
- Blood pressure $\geq 130/85$ mmHg;
- Fasting glucose ≥ 1.10 g/l (6.1 mmol/l).

Type 2 diabetes diagnosis criteria. The definition adopted in this study is that of the American Diabetes Association (ADA) (Gavin, 1998):

1. Fasting plasma glucose ≥ 1.26 g/l (7 mmol/l) twice after 8 hours of fasting.
2. Occasional blood glucose ≥ 2 g/l (11.1 mmol/l) in the presence of symptoms of hyperglycemia (polyuria, polydipsia, unexplained weight loss).
3. Diabetics declared under treatment including oral antidiabetic and/or insulin.

High blood pressure diagnosis criteria. Hypertension (HBP) has been defined according to WHO criteria (Chalmers et al., 1999): mean systolic blood pressure greater than 140 mmHg and/or mean diastolic blood pressure greater than 90 mmHg and/or self-reported current treatment for hypertension with antihypertensive drugs.

Obesity diagnosis criteria. The body mass index (BMI) is calculated according to the Quetelet equation. A subject is considered obese if he has a BMI greater than or equal to 30 kg/m².

Biochemistry and molecular testing. A multichannel analyzer and dedicated kits (Humastar®, HUMAN Diagnostics, Wiesbaden, Germany) were used for the colorimetric, enzymatic measurement of cholesterol (kit: monotest cholesterol with cholesterol esterase, cholesterol oxidase and peroxidase), triglycerides (kit: peridochrom triglyceride with glycerol phosphate oxidase and peroxidase) and glucose (kit: glucose, glucose oxidase and peroxidase). Plasma LDL-cholesterol levels were calculated according to the Friedewald equation. High-density lipoprotein cholesterol levels were measured after sodium phosphotungstate/magnesium chloride precipitation of chylomicrons and VLDL and LDL-cholesterol and then centrifugation. Plasma insulin levels were measured using a microparticle enzyme immune assay running on an AxSYM analyzer (Abbott Laboratories, Abbott Park, Illinois, USA).

Genomic DNA was extracted from white blood cells by using the Stratagene® kit (Agilent Technologies, Les Ulis, France), according to the manufacturer's protocol. Three genetic polymorphisms were characterized in this study, the Epsilon polymorphism defined by the rs7412 and rs429358 single nucleotide polymorphisms (SNPs), the rs439401 and the rs4420638, genotyping was performed by using KASPar technology (KBioscience, Hoddesdon, UK). The genotyping success rates ranged from 93 to 96 %.

Statistical analyses were performed with SAS 9.1 software (SAS Institute Inc., Cary, NC, USA). The Hardy–Weinberg equilibrium was tested using a χ^2 test with one degree of freedom (d.f.). Some of the biochemical traits (Fasting Glucose levels, Triglycerides and Insulin levels) were not normally distributed, we therefore log-transformed these parameters to obtain normal data distributions. Intergroup comparisons of means were performed with a general linear model, multivariate logistic regression analyses were used to calculate the odds ratios for MetS, Type 2 Diabetes (T2D), High Blood pressure (HBP) and obesity (Obes). The confounding variables were age, gender, smoking status and physical activity. After Bonferroni correction, only associations with an uncorrected *p* value below 0.017 were considered to be statistically significant.

Results

Characteristics of study subjects. The main anthropometric, biochemical and clinical characteristics have been measured, the baseline characteristics of the ISOR population study were described elsewhere (Houti et al., 2016).

Genotype and allele distributions. The allele and genotype distributions of the APOE polymorphisms were described in Table 1. There was no evidence of significant deviation from Hardy–Weinberg equilibrium in any distributions.

Prevalence of the metabolic syndrome and the main cardiometabolic risk factors. These data concerning the Oran population are presented in Table 2.

Diabetes mellitus (T2D) was diagnosed in 80 participants (10.6 %). The distribution of prevalence by sex shows no significant difference (*p* = 0.39), it was 11.6 % for men and 9.7 % for women.

The prevalence of obesity in the general population is 21.2 %. It affects more women (32.5 %) than men (9 %), with a significant difference in the prevalence distribution between men and women (*p* < 0.0001).

The prevalence of the metabolic syndrome in the Oran population is 20 %, the distribution of this pathology is also significantly different between the two sexes (*p* < 0.0001). Indeed, it affects more women (25.9 %) than men (13.7 %).

Hypertension affects 20.3 % of the study population. HBP is present in 21.2 % of men and 19.6 % of women, the prevalence distribution in men and women shows no significant difference (*p* = 0.58).

APOE epsilon polymorphism and cardiometabolic risk. No significant association was reported between genotypes of APOE epsilon polymorphism and the studied cardiovascular risk factors (T2D, Obesity, HBP and MetS status), the *p* values ranged from 0.04 to 0.92 (Table 3).

Polymorphism rs439401 and cardio-metabolic risk. In the ISOR study, rs439401 showed a significant association with hypertension (HBP). The T allele confers a high risk of hypertension with an odds ratio (OR) of 1.46 (95 % CI [1.12–1.9], *p* = 0.006). No associations with T2D, obesity and MetS were detected in the ISOR study (see Table 3).

Polymorphism rs4420638 and cardio-metabolic risk. Logistic regression analysis showed that the rs4420638 polymorphism was significantly associated with obesity in the general population. The G allele provides protection against obesity, the resulting OR is 0.48 (95 % CI [0.29–0.79], *p* = 0.004).

Table 1. Genotype and allele frequencies of epsilon polymorphism, rs439401 and rs4420638 in case and control groups

Parameter	MetS	Non MetS
Epsilon		
Genotype frequency, N(%)		
ε2/ε2	1(0.7)	5(0.9)
ε2/ε3	14(9.7)	43(7.4)
ε2/ε4	0	3(0.5)
ε3/ε3	109(75.7)	457(78.8)
ε3/ε4	16(11.1)	66(11.4)
ε4/ε4	4(2.8)	6(1)
TOTAL	144(100)	580(100)
p _{E-H-W}	0.77	0.68
Allele frequency, %		
ε2	5.5	4.8
ε3	86.1	88.2
ε4	8.4	7
rs439401		
Genotype frequency, N(%)		
CC	56(38.6)	228(39.1)
CT	65(44.8)	265(45.2)
TT	24(16.6)	90(15.4)
TOTAL	145(100)	583(100)
p _{E-H-W}	0.48	0.79
Allele frequency, %		
C	61	62
T	39	38
rs4420638		
Genotype frequency, N(%)		
AA	117(79.6)	473(79.1)
AG	27(18.4)	118(19.7)
GG	3(2)	7(1.2)
TOTAL	147(100)	598(100)
p _{E-H-W}	0.90	0.14
Allele frequency, %		
A	88.8	89
G	11.2	11

Note: H-W, Hardy-Weinberg equilibrium.

(see Table 3). No effects of rs4420638 polymorphism on T2D, MetS, and HBP were detected in the ISOR study.

The associations described for rs439401 and rs4420638 remained significant even after adjusting for the APOE epsilon polymorphism.

Discussion

To our knowledge, this is the first study that evaluates the association of APOE gene polymorphisms (epsilon, rs439401 and rs4420638), with the risk of MetS and the main cardio-metabolic risk factors, within the Algerian population.

We found no association between the three polymorphisms of the APOE gene and the metabolic syndrome in the Algerian population. However, some components of the metabolic syndrome considered as a cardiometabolic risk factors were significantly associated with APOE gene polymorphisms.

The logistic regression results showed that the ε2 allele increases the risk of obesity by 88 % in the ISOR study. Similar results were observed in a study among the population of Croatia's Roma minority (Zeljko et al., 2011).

It is possible that gene-nutrition interactions are responsible for the observed association between the ε2 allele and obesity. Indeed, changes in eating habits during the last decade would be responsible for increasing the prevalence of obesity, interacting with the ε2 allele (Boer et al., 1997; Talmud, 2007).

The polymorphisms rs439401 and rs4420638 have been associated in some of GWAS-type studies with changes in plasma lipid concentrations (Kathiresan et al., 2008; Aulchenko et al., 2009; Teslovich et al., 2010), but few studies have investigated the impact of these polymorphisms on metabolic and cardiovascular traits.

Our results on the Oran population, report for the first time, that the T allele of the rs439401 polymorphism increases the risk of arterial hypertension (OR 1.46, 95 % CI [1.12–1.90], $p = 0.006$). No similar results were reported. In the literature, the T allele of rs439401 is significantly associated with changes in BMI, insulin concentration, waist circumference, and triglyceride concentration. The TT genotype is positively associated with an increase in the values of these parameters only in psychologically stressed individuals (Kring et al., 2010). Our results are perhaps indicative of a state of stress of the population, resulting from the changes made in the Algerian population during the last two decades, particularly with

Table 2. Prevalence of the metabolic syndrome and its components in the ISOR population

Parameter	All (n = 787)		Men (n = 378)		Women (n = 409)		p
	n	%	n	%	n	%	
Abdominal adiposity	233	30.1	45	12.1	188	46.8	<.0001
High triglycerides	103	13.3	62	16.7	41	10.2	0.63
Low HDL-Cholesterol	342	44.2	99	26.6	243	60.4	<.0001
High fasting glucose	161	20.8	88	23.7	73	18.2	0.47
T2D	80	10.6	42	11.6	38	9.7	0.39
Obesity	167	21.2	34	9.0	133	32.5	<.0001
MetS	155	20.0	51	13.7	104	25.9	<.0001
HBP	160	20.3	80	21.2	80	19.6	0.58

Note: T2D, Type 2 Diabetes; MetS, Metabolic Syndrome; HBP, High Blood Pressure.

612 Вавиловский журнал генетики и селекции / Vavilov Journal of Genetics and Breeding • 2019 • 23 • 5

Note: T2D, Type 2 Diabetes; MetS, Metabolic Syndrome; HBP, High Blood Pressure; OR adjusted on age, sexe, physical activity, smoking status; OR* obtained for additive modele.

the security crisis in the country. These hypotheses require investigations on a larger sample and in which, the stress level must be measured accurately.

The G allele of rs4420638 seems to confer a protective effect against obesity (OR 0.48, 95 % CI [0.29–0.79], $p = 0.004$), no study was interested in measuring association between rs4420638 polymorphism and obesity previously. No association was reported for the rs4420638 with MetS, T2D and HBP similar results were observed in Tunisian population (Elouej et al., 2016).

The fact that rs4420638 has low linkage disequilibrium with the epsilon polymorphism in our population gives it an advantage over European populations, where these two polymorphisms are in strong linkage disequilibrium (Boulennouar et al., 2013). Thus, the study of the impact of rs4420638 would be independent of the effect of epsilon polymorphism, which makes our population very interesting from a genetic point of view for association analyzes involving rs4420638 polymorphism.

Conclusion

Although APOE variants were not associated with the risk of MetS, the APOE polymorphism alleles were associated with some of the metabolic parameters in Algerian subjects. The relation of APOE rs439401 alleles with a HBP seems been perhaps indicative of a state of stress of the population. These hypotheses require in the future investigations on a larger sample and in which, the stress level must be measured accurately.

The interaction gene-nutrition must be investigated, in the future; indeed, Algerian population known many changes in eating habits during the last decade, which would be responsible for increasing prevalence of obesity in our population and which can influence the effect of APOE polymorphism on the studied parameters.

References

Alberti K.G., Zimmet P., Shaw J. The metabolic syndrome-a new worldwide definition. [Research Support, Non-U.S. Gov't]. *Lancet*. 2005;366(9491):1059-1062.

Aulchenko Y.S., Ripatti S., Lindqvist I., Boomsma D., Heid I.M., Pramstaller P.P., Penninx B.W.J.H., Janssens A.C.J.W., Wilson J.F., Spector T., Martin N.G., Pedersen N.L., Kyvik K.O., Kaprio J., Hofman A., Freimer N.B., Jarvelin M.-R., Gyllenstein U., Campbell H., Rudan I., Johansson Å., Marroni F., Hayward C., Vitart V., Jonasson I., Pattaro C., Wright A., Hastie N., Pichler I., Hicks A.A., Falchi M., Willemsen G., Hottenga J.-J., de Geus E.J.C., Montgomery G.W., Whitfield J., Magnusson P., Saharinen J., Perola M., Silander K., Isaacs A., Sijbrands E.J.G., Uitterlinden A.G., Witteman J.C.M., Oostra B.A., Elliott P., Ruokonen A., Sabatti C., Gieger C., Meitinger T., Kronenberg F., Döring A., Wichmann H.-E., Smit J.H., McCarthy M.I., van Duijn C.M., Peltonen L., for the ENGAGE Consortium. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat. Genet.* 2009;41:47-55.

Balkau B., Valensi P., Eschwege E., Slama G. A review of the metabolic syndrome. *Diabetes Metab.* 2007;33(6):405-413.

Boer J.M., Ehnholm C., Menzel H.J., Havekes L.M., Rosseneu M., O'Reilly D.St.J., Tirt L., and for the EARS Group. Interactions between lifestyle-related factors and the ApoE polymorphism on plasma lipids and apolipoproteins. *Arterioscler. Thromb. Vasc. Biol.* 1997;17(9):1675-1681.

Boulennouar H., Benckekor S.M., Meroufel D.N., Hetraf S.A.L., Djelloul H.O., Hermant X., Grenier-Boley B., Medjaoui I.H., Mehtar N.S., Amouyel P., Houti L., Meirhaeghe A., Goumidi L. Impact of APOE gene polymorphisms on the lipid profile in an Algerian population. *Lipids Health Dis.* 2013;12:155.

Chalmers J., MacMahon S., Mancia G., Whitworth J., Beilin L., Hansson L., Neal B., Rodgers A., Ni Mhurchu C., Clark T. 1999 World Health Organization-International Society of Hypertension Guidelines for the management of hypertension. Guidelines sub-committee of the World Health Organization. *Clin. Exp. Hypertens.* 1999; 21(5-6):1009-1060.

Chuang L.M. Human genetics of the metabolic syndrome. *Asia Pac. J. Clin. Nutr.* 2008;17(Suppl.1):43-46.

Elouej S., Rejeb I., Attaoua R., Nagara M., Sallem O.K., Kamoun I., Chargui M., Jamoussi H., Turki Z., Abid A., Ben Slama C., Bahri S., Ben Romdhane H., Abdelhak S., Kefi R., Grigorescu F. Gender-specific associations of genetic variants with metabolic syndrome components in the Tunisian population. *Endocrin. Res.* 2016;41(4): 300-309.

Gami A.S., Witt B.J., Howard D.E., Erwin P.J., Gami L.A., Somers V.K., Montori V.M. Metabolic syndrome and risk of incident cardiovascular events and death: a systematic review and meta-analysis of longitudinal studies. *J. Am. Coll. Cardiol.* 2007;49(4):403-414.

Gavin J.R., 3rd. New classification and diagnostic criteria for diabetes mellitus. *Clin. Cornerstone.* 1998;1(3):1-12.

Grundy S.M., Brewer H.B., Jr., Cleeman J.I., Smith S.C., Jr., Lenfant C. Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. *Circulation.* 2004; 109(3):433-438.

Hillier T.A., Rousseau A., Lange C., Lépinay P., Cailleau M., Novak M., Calliez E., Ducimetière P., Balkau B. Practical way to assess metabolic syndrome using a continuous score obtained from principal components analysis. *Diabetologia.* 2006;49(7):1528-1535.

Houti L., Hamani-Medjaoui I., Lardjam-Hetraf S.A., Ouhaibi-Djelloul H., Chougrani S., Goumidi L., Mediène-Benckekor S. Prevalence of metabolic syndrome and its related risk factors in the city of Oran, Algeria: the ISOR Study. *Ethn. Dis.* 2016;26(1):99-106.

Kahn R., Buse J., Ferrannini E., Stern M. The metabolic syndrome: time for a critical appraisal: joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care.* 2005;28(9):2289-2304.

Kathiresan S., Melander O., Guiducci C., Surti A., Burt N.P., Rieder M.J., Cooper G.M., Roos C., Voight B.F., Havulinna A.S., Wahlstrand B., Hedner T., Corella D., Tai E.S., Ordovas J.M., Berglund G., Vartiainen E., Jousilahti P., Hedblad B., Taskinen M.-R., Newton-Cheh C., Salomaa V., Peltonen L., Groop L., Altshuler D.M., Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat. Genet.* 2008;40:189-197.

Kring S.I., Brummett B.H., Barefoot J., Garrett M.E., Ashley-Koch A.E., Boyle S.H., Siegler I.C., Sørensen T.I.A., Williams R.B. Impact of psychological stress on the associations between apolipoprotein E variants and metabolic traits: findings in an American sample of caregivers and controls. *Psychosom. Med.* 2010;72(5):427-433.

Meigs J.B., Rutter M.K., Sullivan L.M., Fox C.S., D'Agostino R.B., Sr., Wilson P.W. Impact of insulin resistance on risk of type 2 diabetes and cardiovascular disease in people with metabolic syndrome. *Diabetes Care.* 2007;30(5):1219-1225.

Ministère de la Santé, de la population et de la réforme hospitalière. In: Transition épidémiologique et système de santé. Projet TAHINA Résultats Enquête Nationale Santé. 2007.

Reaven G.M. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes.* 1988;37(12):1595-1607.

Talmud P.J. Gene-environment interaction and its impact on coronary heart disease risk. *Nutr. Metab. Cardiovasc. Dis.* 2007;17(2):148-152.

- Teslovich T.M., Musunuru K., Smith A.V. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707-713.
- Wilson P.W., D'Agostino R.B., Parise H., Sullivan L., Meigs J.B. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation*. 2005;112(20):3066-3072.
- Wilson P.W., Myers R.H., Larson M.G., Ordovas J.M., Wolf P.A., Schaefer E.J. Apolipoprotein E alleles, dyslipidemia, and coronary heart disease. The Framingham Offspring Study. *JAMA*. 1994;272(21):1666-1671.
- Zeljko H.M., Skaric-Juric T., Narancic N.S., Tomas Ž., Barešić A., Salihović M.P., Starčević B., Jančićjević B. E2 allele of the apolipoprotein E gene polymorphism is predictive for obesity status in Roma minority population of Croatia. *Lipids Health Dis*. 2011;10:9. DOI 10.1186/1476-511X-10-9.

Acknowledgements. The ISOR project was funded through a collaboration agreement between (DPGRF) the Direction de la Post-Graduation et de la Recherche-Formation (Algeria) and (INSERM) the Institut National de la Santé et de la Recherche Médicale (France). The work in France was also part-funded by INSERM. The work in Algeria was also part-funded by the Algerian National Agency for the Development of Health Research (ANDRS) and a grant from the Projets Nationaux de Recherche (PNR) program run by the Algerian Direction Générale de la Recherche Scientifique et du Développement Technologique/ Ministère de l'Enseignement Supérieur et de la Recherche Scientifique (DGRSDT/MESRS).

We are grateful to Dr. Louisa Goumidi, Dr. Aline Meirhaeghe and Pr. Philippe Amouyel for technical support in DNA extraction, genotyping and statistical analysis, which were performed in their lab (INSERM, U744; Institut Pasteur de Lille, Université Lille Nord de France, Lille, France).

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

Authors' contributions. SMB and LH, designed research; SMB, LH and IHM conducted research; HOD, SLH, IHM, SMB and LH participated in the recruitment of subjects; HB built the database; HB performed the DNA extraction and genotyping; HB and SMB performed the statistical analyses; HB and SMB interpreted the results. IHM assayed biochemical parameters; HB wrote the paper under the supervision of SMB; HB and SMB had primary responsibility for final content. All authors read and approved the final manuscript.

Received March 17, 2019. Revised July 8, 2019. Accepted July 9, 2019.

Prediction of some peroxidase functions in *Arabidopsis thaliana* L. by bioinformatic search

A.S. Tugbaeva , A.A. Ermoshin, I.S. Kiseleva

Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia

 e-mail: anastasia.tugbaeva@gmail.com

Peroxidases of class III are common in various organisms. They are involved in lignin biosynthesis and plant protection against stressors. Peroxidases are presented in many isoforms, whose role is not always clear. The aim of this study is to analyze the amino acid sequences of reference peroxidases with known functions and peroxidases from *Arabidopsis thaliana* L. whose functions are unknown and to consider their putative roles in lignin biosynthesis. The structural and functional organization of peroxidases was analyzed by bioinformatical methods applied to open Internet sources. Seven reference peroxidases were chosen from four plant species: *Zinnia* sp., *Armoracia rusticana* P.G. Gaertn., *Lycopersicon esculentum* L. и *Populus alba* L. Twenty-four amino acid sequences of homologous peroxidases from *A. thaliana* were selected for the analyses with the BLAST service. Their molecular weights and isoelectric points were calculated. Multiple alignments of amino acid sequences and phylogenetic analysis were done. Sites of binding to monolignol substrates were identified in seven peroxidases from *A. thaliana*, and the enzymes were assigned to the groups of S- or G-peroxidases. Amino acid replacements in the primary structures of peroxidases were analyzed. Peroxidases from *A. thaliana* were clustered with reference peroxidases. They formed six clusters on the phylogenetic tree, three of which contained only *A. thaliana* peroxidases. Peroxidases within each cluster had similar molecular weights and isoelectric points, common localization of expression, and similar functions. Thus, the use of bioinformatics, databases, and published data bring us to assumptions as to the functions of several *A. thaliana* class III peroxidases. AtPrx39 peroxidase was shown to be affine to sinapyl alcohol; AtPrx54, to *p*-coumaryl and coniferyl alcohols. They are likely to participate in lignin biosynthesis.

Key words: peroxidase; lignification; *Arabidopsis thaliana* L.; bioinformatics; multiple alignments.

For citation: Tugbaeva A.S., Ermoshin A.A., Kiseleva I.S. Prediction of some peroxidase functions in *Arabidopsis thaliana* L. by bioinformatic search. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2019;23(5):615-623. DOI 10.18699/VJ19.533

Предсказание функций некоторых пероксидаз *Arabidopsis thaliana* L. на основе биоинформатического поиска

А.С. Тугбаева , А.А. Ермошин, И.С. Киселева

Уральский федеральный университет им. первого Президента России Б.Н. Ельцина, Екатеринбург, Россия

 e-mail: anastasia.tugbaeva@gmail.com

Пероксидазы III класса являются распространенными в разных группах организмов ферментами, участвуют в биосинтезе лигнина, защите растений от патогенов и абиотических стрессоров. Peroxidases имеют множество изоформ, роль которых в клеточных процессах не всегда ясна. В работе проанализированы аминокислотные последовательности референсных пероксидаз с известными функциями и пероксидаз *Arabidopsis thaliana* L., функции которых неясны, выдвинуто предположение о роли последних в биосинтезе лигнина. Проведен биоинформатический анализ информации о структурно-функциональной организации пероксидаз из открытых интернет-источников. Было выбрано семь референсных пероксидаз из травянистых и древесных растений четырех видов (*Zinnia* sp., *Armoracia rusticana* P.G. Gaertn., *Lycopersicon esculentum* L. и *Populus alba* L.), для которых показано их участие в лигнификации клеточной стенки. С помощью сервиса BLAST были отобраны 24 аминокислотные последовательности гомологичных пероксидаз из *A. thaliana*. Для каждого фермента рассчитаны молекулярная масса и изоэлектрическая точка. Построены множественные выравнивания аминокислотных последовательностей и проведен филогенетический анализ. Для семи пероксидаз *A. thaliana* выявлены сайты связывания с субстратом (синаповый, *para*-кумаровый и кониферилловый спирты), их принадлежность группам S- или G-пероксидаз. Проанализированы аминокислотные замены в первичной структуре белка. Peroxidases *A. thaliana* кластеризуются с референсными пероксидазами и образуют шесть групп на филогенетическом дереве, три из которых образованы исключительно пероксидазами *A. thaliana*. Peroxidases, объединенные в общий кластер, характеризуются близким значением молекулярной массы и величиной

изоэлектрической точки, имеют общую локализацию экспрессии и выполняют схожие функции. Таким образом, использование биоинформатических методов, анализ литературы и материалов в базах данных позволили предположить не известные ранее функции для некоторых пероксидаз *A. thaliana*, относящихся к III классу. Показано, что пероксидаза AtPrx39 обладает сродством к окислению синапового спирта; AtPrx54 – к окислению *пара*-кумарового и кониферилового спиртов, и, предположительно, они могут участвовать в биосинтезе лигнина.

Ключевые слова: пероксидаза; лигнификация; *Arabidopsis thaliana* L.; биоинформатика; множественные выравнивания.

Introduction

Peroxidases are the group of enzymes that catalyze the oxidation of a substrate with the presence of hydrogen peroxide. The superfamily of “plant” peroxidases (those of plants, fungi, and bacteria) is divided into three classes based on their structural and catalytic properties. All peroxidases contain 10 homologous α -helices. Class I and class II have one specific α -helix, and class III peroxidases have three specific α -helices (Hiraga et al., 2001).

Living organisms contain many peroxidase isoforms, and their amino acid sequences are similar by less than 20 %. A high level of conservation characterizes five amino acid positions essential for the folding of α -helices, assembly of subunits, and catalytic properties of the enzymes (Hiraga et al., 2001).

Higher plants contain class I and III peroxidases, which differ in structure, function, and location in the plant cell. Ascorbate peroxidase (EC 1.11.1.11) and glutathione peroxidase (EC 1.11.1.9) belong to class I. They are located in chloroplasts, peroxisomes, and cytoplasm. Class I peroxidases are distinguished by high specificity to an oxidizable substrate. Class III peroxidases (EC 1.11.1.7) include enzymes that are located in vacuoles and secreted into the apoplast. They oxidize various substrates. Class III plant peroxidases are encoded by a large family of genes: 73 have been identified in *Arabidopsis thaliana* L. and 138 in *Oryza sativa* L. (Welinder et al., 2002; Passardi et al., 2004a). Class III peroxidases act as components of the antioxidant system of plants and, at the same time, can form reactive oxygen species (Passardi et al., 2004a). The dual functions of peroxidases allow them to take part in many physiological processes: protection against pathogens (Passardi et al., 2004b), wound healing, auxin and anthocyanin catabolism, and porphyrin metabolism (Cosio, Dunand, 2009; Jovanovic et al., 2018).

Apoplastic peroxidases are involved in the biosynthesis of cell wall components, such as lignin and suberin. Lignin is an aromatic phenolic heteropolymer with a disordered structure, covalently associated with polysaccharides of the secondary cell wall and responsible for its strength and hydrophobicity. The composition and amount of lignin in the cell wall change in the course of plant ontogenesis (Boerjan et al., 2003) and in response to different stress factors (Liu et al., 2018).

Peroxidase and laccase are involved in the formation of lignin precursors: *p*-coumaryl, coniferyl, and sinapyl radicals. Laccases (EC 1.10.3.2) catalyze the formation of guaiacyl (G) units, whereas peroxidases are involved in the generation of syringyl (S), *p*-hydroxyphenyl (H), and guaiacyl (G) units of lignin (Berthet et al., 2012).

Despite the large amount of research focusing on class III peroxidases, only few isoforms have been shown to participate

in lignin biosynthesis in herbaceous (*Zinnia* sp., *Armoracia rusticana* P.G. Gaertn., *Lycopersicon esculentum* L.) and woody (*Populus alba* L.) plants (Quiroga et al., 2000; Aoyama et al., 2002; Sasaki et al., 2004; Sato et al., 2006; Marjamaa et al., 2009). Class III peroxidases can oxidize three monolignols; however, most isoforms oxidize coniferyl and *p*-coumaryl alcohols and only few of them use sinapyl alcohol as a substrate *in vitro* (Barcelo et al., 2007).

The structures of peroxidases ZePrx34, ZPO-C, CWPO-C, HRP, HRP-A2A, HRP-C1C, and TPX1, which can be considered reference ones, have been studied in detail (Quiroga et al., 2000; Aoyama et al., 2002; Sasaki et al., 2004; Gabaldon et al., 2005; Sato et al., 2006). The attention to peroxidases is due to their function in the formation of plant resistance to oxidative stress caused by both abiotic and biotic factors, as well as to their participation in lignin biosynthesis and plant growth. Bioinformatic analysis of peroxidases with unknown functions is of fundamental (determination of the enzyme functions) and practical (design of genetic constructs to create resistant plants or plants with a modified cell wall) significance. The purpose of this work is to analyze the functions of *A. thaliana* peroxidases based on the similarity to the amino acid sequences of reference plant peroxidases for which the involvement in cell wall lignification is known.

Materials and methods

Amino acids sequences of plant peroxidases from *Zinnia elegans* Jacq. (ZePrx34.70, NCBI identifier – Q4W1I8.1), *Zinnia violacea* Cav. (ZPO-C, BAD93164.1), *P. alba* (CWPO-C, BAE16616.1), *A. rusticana* (HRP, CCJ34837.1; HRP-A2A, CCJ34825.1; HRP-C1C, P15233.1) and *L. esculentum* (TPX1, NP_001289850.1) were sought in the Protein NCBI database (<https://www.ncbi.nlm.nih.gov/protein/>). They were aligned with the amino acid sequences of *A. thaliana* with the Protein BLAST tool (<https://blast.ncbi.nlm.nih.gov>). Search route: database – model organism (landmark), organism – *A. thaliana* (taxid: 3702), algorithm PSI-BLAST (Position-Specific Iterated BLAST). A library (Suppl. 1)¹ was formed from 24 amino acid sequences of *A. thaliana* peroxidases with high levels of similarity to reference peroxidases (E-value less than $1e^{-80}$). The online program EMBOSS Pepstats (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/) was applied to calculate the molecular masses and isoelectric points of the proteins.

The phylogenetic tree based on *A. thaliana* peroxidase proteins was built by the Neighbor-Joining method (Sanou, Nei, 1981) in the MEGA 7 program based on sequence alignments of the encoded protein. The evolutionary distances were

¹ Supplementary Materials 1–3 are available in the online version of the paper: <http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx11.pdf>

computed by the *p*-distance method (Nei, Kumar, 2000). The bootstrap test included 1000 replicates, and the results are shown nearby the branches (Kumar et al., 2016).

Information about the expression of *A. thaliana* peroxidase genes at different stages of development was obtained from the bio-array resource for plant functional genomics (<http://bar.utoronto.ca>). Analysis of peroxidase functions was carried out with regard to information stored in the Arabidopsis Information Resource (www.arabidopsis.org) by gene identifiers in TAIR. We considered information from the Annotations, GO Biological Process section on the involvement of peroxidases in stress reactions, growth, and lignification of the cell wall. Amino acid sequence alignments were built with the CLUSTAL algorithm for multiple sequence alignment in MUSCLE 3.8 (<https://www.ebi.ac.uk/Tools/msa/>). Highly conservative and semiconservative domains, structural motifs were identified.

Results

The bioinformatic search with Protein BLAST showed that plant peroxidases from *Z. elegans* (ZePrx34.70, ZPO-C), *P. alba* (CWPO-C), *A. rusticana* (HRP, HRP-A2A, HRP-C1C) and *L. esculentum* L. (TPX1) have high levels of similarity to 24 *A. thaliana* peroxidases. Molecular weights and isoelectric points were calculated for these enzymes (Table 1).

The isoelectric points (pIs) and molecular weights of *A. thaliana* peroxidases differ from those of reference enzymes. In particular, the pI value of peroxidase AtPrx36 is in the more acidic pH range compared to ZePrx34.70, and the protein has a higher molecular weight (38.24 vs. 34.24 kDa, respectively). AtPrx13 peroxidase is characterized by an acidic pI value (4.74), whereas pI for HRP is 8.35. AtPrx32, 37 and 23 peroxidases have pIs within 6.62–7.97, and their molecular weights vary from 38.10 to 38.85 kDa, whereas the pI and molecular weight of HRP_A2A protein are 4.62 and 35.03 kDa, respectively.

It is seen that *A. thaliana* peroxidases differ in pI values and molecular weights from reference enzymes and they are expected to differ in their affinity to the substrate and in functions. It is known that basic peroxidases (isoelectric point > 7.0) can oxidizing *p*-coumaryl, coniferyl and sinapyl alcohols (Kukavica et al., 2012), while acidic peroxidases (isoelectric point < 7.0) are poorly capable of oxidizing sinapyl alcohol (Barcelo et al., 2004). Therefore, the roles of basic and acidic peroxidases in cell wall lignification may be different. Plant peroxidases with high ability to oxidize coniferyl alcohol (CWPO-A, HRP-C1C and AtPrx53) or sinapyl alcohol (CWPO-C from *P. alba*, ZePrx from *Z. elegans*, AtPrx4) are described.

The phylogenetic tree of these peroxidases shown in Figure is constructed by amino acid alignment. The reference peroxidases and *A. thaliana* enzymes form six clusters. The first one includes HRP-C1C, AtPrx33, 34, and 32 peroxidases with a high bootstrap support value of 72–100 %. The second cluster groups peroxidases, homologous to HRP_A2A: AtPrx2 and 54 (bootstrap support 100 %), AtPrx52 and 4 peroxidases, homologous to ZePrx34.70, form the third cluster with bootstrap support of 98–99 %. AtPrx47, 64, and 66 peroxidases group in the fourth cluster together with ZPO-C peroxidase (92–100 % bootstrap support). The fifth cluster on the phylo-

genetic tree combines peroxidases TPX1 and AtPrx3 and 39 with a bootstrap support of 100 %. The sixth cluster consists of HRP, CWPO-C, AtPrx71, 62, and 69 peroxidases with bootstrap support value of 72–100 %.

Separate clusters on the phylogenetic tree include *A. thaliana* peroxidases: AtPrx38 and AtPrx37 (cluster A), AtPrx22 and AtPrx23 (cluster B), AtPrx36 and AtPrx72 (cluster C) with 100 % bootstrap support.

With materials from the BAR and TAIR databases, the peroxidase functions and expression sites were identified for the enzymes of clusters 1–6 (Table 2). The Table 2 does not include clusters A, B, or C, formed by homologous proteins of *A. thaliana*.

Functions of some *A. thaliana* peroxidases have been studied in mutants with knocked-out genes and transgenic plants. According to experimental studies, peroxidase HRP-C1C from *A. rusticana* most effectively oxidizes coniferyl alcohol *in vitro* (Sasaki et al., 2004). The most homologous HRP-C1C peroxidases AtPrx33 and AtPrx34 are involved in root growth and cell elongation (Irshad et al., 2008) and in an oxidative burst, when pathogens penetrate into the cell (Bindschedler et al., 2006). AtPrx32 peroxidase is involved in cell elongation (Irshad et al., 2008). Thus, there is no data on the participation of cluster 1 enzymes in cell wall lignification.

Purified peroxidase HRP_A2A from *A. rusticana* efficiently oxidizes guaiacol *in vitro* (Krainer et al., 2014). According to the BAR database, the AtPRX2 and AtPRX54 genes are expressed in its seedling roots and hypocotyl and in the roots of juvenile plants (see Table 2). Mutants of *A. thaliana* *atprx2* are characterized by a reduced total lignin content, changes in lignin composition, and plant biomass decrease (Shigeto et al., 2013).

The isoform ZePrx34.70 from *Z. elegans*, catalyzing the oxidation of sinapyl alcohol, is expressed in roots and hypocotyl and involved in lignification (Gabaldon et al., 2005). Among the analyzed peroxidases, AtPrx4 and AtPrx52 from *A. thaliana* are homologous to ZePrx34.70, as confirmed in (Herrero et al., 2013a). According to (Fernandez-Perez et al., 2015), the *AtPrx4* gene is expressed in roots, stems, and leaves, and it affects the plant growth on long days. The product of its expression is involved in syringol polymerization.

Purified ZPO-C peroxidase from *Z. violacea* uses both synapyl and coniferyl alcohol as a substrate *in vitro* (Sato et al., 2006). Homologous AtPrx66 takes part in cell wall lignification of forming vessels (Sato et al., 2006). It was shown that the homolog AtPrx64 also plays a role in xylem lignification (Yokoyama, Nishitani, 2006).

The gene for the basic peroxidase TPX1 of *L. esculentum* is specifically expressed in root xylem and involved in lignification and suberization (Quiroga et al., 2000). An increase in lignin content has been shown in transgenic *L. esculentum* plants with overexpression of TPX1 (Mansouri et al., 1999). Homologous peroxidase AtPrx3 is involved in lignification (see Table 2). AtPrx3 cationic peroxidase transcripts were found in the seedlings and roots, and their participation in the response of plants to salt stress and drought was shown (Llorente et al., 2002). The role of AtPrx39 peroxidase in cell wall lignification has not been studied. However, the *AtPRX39* gene is expressed in the root transport zone. It affects the development of the root system (Tsukagoshi et al., 2010).

Table 1. Annotation of class III peroxidases from *A. thaliana* with high levels of similarity (Score and E-value) to reference peroxidases. The TAIR acc. no., NCBI acc. no., molecular weight, and isoelectric point are indicated for each protein

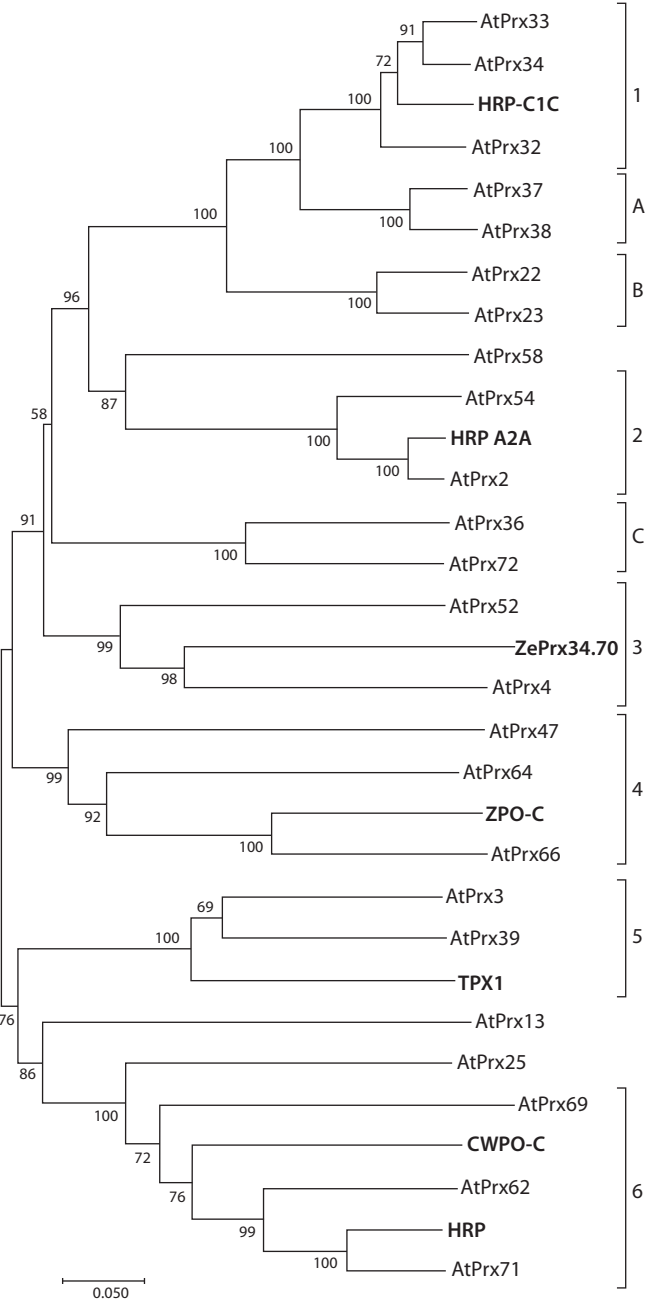
Reference peroxidase, protein ID, pl, molecular weight, kDa	<i>A. thaliana</i> peroxidase, protein ID, pl, molecular weight, kDa	TAIR ID	Score	E-value
ZePrx34.70, <i>Zinnia elegans</i> Q4W1I8.1 8.22 34.24	AtPrx4 NP_172906.1 7.74 34.41	AT1G14540	342	1e ⁻¹¹⁷
	AtPrx52 NP_196153.1 8.30 34.21	AT5G05340	327	3e ⁻¹¹¹
	AtPrx36 NP_190668.2 4.66 38.24	AT3G50990	262	1e ⁻⁸⁵
ZPO-C, <i>Zinnia violacea</i> BAD93164.1 8.64 34.79	AtPrx66 NP_200002.3 9.76 35.56	AT5G51890	452	5e ⁻¹⁶¹
	AtPrx64 NP_199033.1 9.04 34.71	AT5G42180	328	5e ⁻¹¹²
	AtPrx47 NP_001320124.1 8.29 35.97	AT4G33420	283	2e ⁻⁹⁴
CWPO-C, <i>Populus alba</i> BAE16616.1 8.30 34.63	AtPrx71 NP_201217.1 8.18 34.89	AT5G64120	425	8e ⁻¹⁵⁰
	AtPrx25 NP_181679.4 7.53 35.89	AT2G41480	417	6e ⁻¹⁴⁷
TPX1, <i>Lycopersicon esculentum</i> NP_001289850.1 7.51 35.99	AtPrx3 NP_172018.1 8.41 34.91	At1G05260	449	2e ⁻¹⁵⁹
	AtPrx39 NP_192868.1 6.92 35.60	AT4G11290	430	9e ⁻¹⁵²
	AtPrx72 NP_201440.1 8.47 37.43	AT5G66390	261	3e ⁻⁸⁵
HRP, <i>Armoracia rusticana</i> CCJ34837.1 8.35 34.79	AtPrx62 NP_198774.1 8.42 34.13	AT5G39580	459	2e ⁻¹⁶³
	AtPrx69 NP_201215.1 9.71 35.68	AT5G64100	326	8e ⁻¹¹¹
	AtPrx13 NP_177835.3 4.74 34.76	AT1G77100	278	3e ⁻⁹²

End of Table1

Reference peroxidase, protein ID, pl, molecular weight, kDa	<i>A. thaliana</i> peroxidase, protein ID, pl, molecular weight, kDa	TAIR ID	Score	E-value
HRP_A2A, <i>Armoracia rusticana</i> CCJ34825.1 4.62 35.03	AtPrx2 NP_196290.1 4.52 34.99	AT5G06720	635	0.0
	AtPrx54 NP_196291.1 4.27 37.29	AT5G06730	578	0.0
	AtPrx22 NP_181372.1 5.76 38.11	AT2G38380	380	1e ⁻¹³¹
	AtPrx58 NP_197488.1 4.92 35.43	AT5G19880	363	2e ⁻¹²⁵
	AtPrx37 OAP01113.1 7.85 38.20	AT4G09970	357	8e ⁻¹²³
	AtPrx23 NP_181373.1 7.97 38.10	AT2G38390	356	3e ⁻¹²²
	AtPrx32 NP_850652.1 6.62 38.85	AT3G32980	354	2e ⁻¹²¹
	AtPrx33 NP_190480.1 6.80 38.94	AT3G49110	629	0.0
HRP-C1C, <i>Armoracia rusticana</i> P15233.1 6.62 36.54	AtPrx34 NP_190481.1 7.56 38.83	AT3G49120	624	0.0
	AtPrx38 NP_192618.1 7.57 38.09	AT4G08780	526	0.0

Peroxidase CWPO-C from *P. alba* is a cationic isoform of the enzyme efficiently polymerizing sinapyl alcohol *in vitro* (Aoyama et al., 2002). HRP is a cationic isoform with high ability to oxidize coniferyl alcohol. Peroxidase isoenzymes, such as HRP and CWPO-C, have been shown to catalyze single-electron oxidation of sinapyl alcohol using coniferyl alcohol as a radical mediator (Aoyama et al., 2002). AtPrx71 peroxidase is the closest homolog of HRP. It participates in the formation of secondary xylem (Yokoyama, Nishitani, 2006) and in the response to biotic factors (Chassot et al., 2007). AtPrx62 peroxidase expression increases in response to heavy metal ions and plant pathogens (Cosio, Dunand, 2009).

The genes encoding peroxidase AtPrx32 and 37 are expressed in the root and hypocotyl and involved in lignification (see Table 2). Overexpression of *AtPRX37* in transgenic *A. thaliana* causes a decrease in plant growth rate, affects the development of xylem, and ultimately leads to the formation of a dwarf phenotype. Presumably, AtPrx37 peroxidase is involved in the regulation of plant growth through the cell wall lignification process (Pedreira et al., 2011). AtPrx72 and 36 peroxidases differ in localization in plant tissues and their functions. The *AtPRX72* gene is expressed in roots and stems (Valerio et al., 2004). The *AtPRX36* gene is expressed in the hypocotyl, where it takes part in cell elongation (Irshad et al., 2008); in the endosperm; and the seed coat (Kunieda et



The peroxidase protein phylogenetic tree from *A. thaliana* based on sequence alignments of the encoded protein (Sanou, Nei, 1981). The evolutionary history was reconstructed by the Neighbor-Joining method. The optimal tree with the sum of branch length 5.06962574 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown nearby the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to construct the phylogenetic tree. The evolutionary distances were computed by the *p*-distance method (Nei, Kumar, 2000). They are presented as numbers of amino acid differences per site. The analysis involves 31 amino acid sequences. All positions containing gaps and missing data are eliminated. A total of 268 positions are present in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016). For protein database NCBI accession numbers of the sequences used for the building of phylogenetic tree see Table 1, Suppl. 1.

al., 2013). Analysis of *A. thaliana* mutants for the *AtPRX72* gene showed a decrease in lignin content and the number of syringyl units. In addition, the mutants were characterized by slow growth, decrease in stem diameter, and smaller numbers of shoots and leaves (Herrero et al., 2013b). Thus, several peroxidases, clustered in groups 2–6, can oxidize monolignols and participate in the lignification of cell walls.

We anticipate that peroxidases combined into a common cluster on a phylogenetic tree are predominantly expressed in the same plant organs and perform similar functions (see Table 2). Of all peroxidases, the participation of *AtPrx54* and 39 in physiological processes is least understood. We infer from the results of data analysis that *AtPrx54* peroxidase performs functions similar to *AtPrx2*: it is involved in lignification, growth, and response to abiotic stress. *AtPrx39* peroxidase is involved in lignification and responses to biotic and abiotic stress. Assuming the concept of the evolutionary origin of proteins (Gabaldon, Koonin, 2013), we conjecture the participation of *AtPrx54* and 39 peroxidases in cell wall lignification processes, since they are orthologs of *HRPA2A* and *TPX1*, respectively. *AtPrx69* and 62 peroxidases, combined into a common cluster with *AtPrx71* peroxidase, are also highly likely to be involved in cell wall lignification.

The functions of homologous proteins can be determined based on their domain structure and substrate binding sites. Multiple alignment of amino acid sequences was previously performed (data not shown). It included reference and *A. thaliana* peroxidases combined into a common cluster (clusters 1–6 on the phylogenetic tree, see Figure). *AtPrx54* peroxidase was shown to have structures like *AtPrx4* and *HRP_A2A*; and *AtPrx39* peroxidase was highly homologous to *TPX1*, *CWPO-C*, and *HRP* proteins. Highly conservative and semiconservative sections, structural motifs characteristic of the analyzed class III peroxidases were identified (see Suppl. 2, 3).

Thus, peroxidases are distinguished by structural motifs, and, accordingly, affinity to oxidizable substrates. The structural motifs that are necessary and sufficient to polymerize G-monolignols are V78, 95-VSCSD, S98, 105-SEA, F185, and N281. S-Peroxidases have motifs absent from G-peroxidases: I80, 95-VSCAD, A98, 105-ARD, Y178, and K268. The 95-VSCAD motif determines the ability of peroxidases to polymerize both syringaldazine and sinapyl alcohol (Barcelo et al., 2007). In addition, the affinity of peroxidase to the substrate is affected by hydrophobic interactions between the substrate and the enzyme, which involve the amino acids at positions P69, I138, P139, S140, R175, and V178 (Barcelo et al., 2007). Studies of the ATP A2 peroxidase structure from *A. thaliana* have shown that hydrophobic interactions between the sinapyl alcohol and the amino acid residues in position I138 and P139 do not allow the enzyme to use sinapyl alcohol as a substrate (Ostergaard et al., 2000).

The G-peroxidases whose amino acid sequences were analyzed include *AtPrx2* and 54, *HRP_A2A*. They are characterized by structural motifs V78, 95-VSCSD, S98, 105-SEA, F186, and N281. The hydrophobicity of the site of substrate binding is determined by amino acid substitutions increasing hydrophobicity: proline to alanine at position 96, isoleucine to leucine at position 138, and isoleucine to

Table 2. The function of class III peroxidases from *A. thaliana*

Cluster	Peroxidase	Function	Localization of expression
1 Homology HRP-C1C	AtPrx33	Cell elongation, biotic stress	Root, hypocotyl, 2nd internode, leaves, stage 15 flowers, sepals
	AtPrx34	Cell elongation, biotic and abiotic stress	Root, hypocotyl, 2nd internode, leaves, stage 15 flowers, sepals
	AtPrx32	Cell elongation	Root 1-day-old-seedling
2 Homology HRP A2A	AtPrx54	No data	Root, hypocotyl
	AtPrx2	Lignification, cell elongation, abiotic stress	Root, hypocotyl, 1st node, carpels, sepals, stamen
3 Homology ZePrx34.70	AtPrx52	Lignification, abiotic and biotic stress	Flowers, sepals
	AtPrx4	Lignification, cell elongation	Root, cotyledon, hypocotyl, 2nd internode, leaves, stage 10–15 flowers, sepals, stamen
4 Homology ZPO-C	AtPrx47	Lignification	Root, petals, stage 15 flowers
	AtPrx64	Lignification, abiotic stress	Seeds, 1st node, 2nd internode
	AtPrx66	Lignification	Leaves, stage 3–4 seeds, sepals, petals, stamen
5 Homology TPX1	AtPrx3	Lignification, abiotic and biotic stress	Root, hypocotyl, stage 4–7 seeds
	AtPrx39	No data	Root, hypocotyl, carpels
6 Homology CWPO-C и HPR	AtPrx69	Cell elongation, abiotic and biotic stress	Root, hypocotyl
	AtPrx62	Abiotic and biotic stress	Root, hypocotyl, 24 h imbibed seed
	AtPrx71	Lignification, abiotic and biotic stress	Leaves, hypocotyl, sepals, petals, carpels, stage 3 seeds

Note: The data about peroxidases functions were obtained from the Arabidopsis Information Resource (www.arabidopsis.org). Localization of expression on different stages of development was analyzed using the bio-array resource for plant functional genomics (<http://bar.utoronto.ca>).

phenylalanine at position 142. Replacements at sites with hydrophilic amino acids (glycine to proline at position 68, isoleucine to leucine at position 138, arginine to glutamine at position 175, glycine to valine at position 177, or valine to threonine at position 178) do not change the properties of the substrate-binding sites. The 138-IPS hydrophobic motif determines the conformation of the protein and the hydrophobicity of the substrate-binding site. Thus, AtPrx54 peroxidase has sites that enable it to polymerize *p*-coumaryl and coniferyl alcohols.

S-Peroxidases include reference enzymes CWPO-C, TPX1, HPR, and AtPrx3, 39, 62, 69, and 71 from *A. thaliana*. Their catalytic properties are determined by the motifs I78, 92-VSCAD, A96, 103-ARD, Y182, and K282. Amino acid substitutions of leucine for isoleucine at position 135 and tyrosine for phenylalanine at position 231 ensure the

hydrophobicity of the substrate-binding site. Presumably, peroxidases AtPrx39, 69, and 62 are involved in the polymerization of sinapyl alcohol.

Conclusion

Plant peroxidases of class III from different plant families are similar to each other in amino acid sequences, tissue localization, and functions. Structure-functional regions were identified in the peroxidases on the base of amino acid sequence homology. These regions allow inferences as to the substrate specificity of the peroxidases. The results show that AtPrx39 oxidizes sinapyl alcohol and belongs to S-peroxidases; AtPrx54 oxidizes *p*-coumaryl and coniferyl alcohols and belongs to G-peroxidases. Therefore, AtPrx39 and 54 peroxidases can participate in the polymerization of monolignols in lignin biosynthesis. Thus, the use of bioinformatic methods

and the analysis of literature and materials in databases suggest previously unknown functions of *A. thaliana* peroxidases belonging to class III.

References

- Aoyama W., Sasaki S., Matsumura S., Mitsunaga T., Hirai H., Tsutsumi Y., Nishida T. Sinapyl alcohol-specific peroxidase isoenzyme catalyzes the formation of the dehydrogenative polymer from sinapyl alcohol. *J. Wood Sci.* 2002;6(48):497-504. DOI 10.1007/BF00766646/.
- Barcelo A.R., Gomez Ros L.V., Carrasco A.E. Looking for syringyl peroxidases. *Trends Plant Sci.* 2007;12(1):486-491. DOI 10.1016/j.tplants.2007.09.002.
- Barcelo A.R., Gomez Ros L.V., Gabaldon C., Lopez-Serrano M., Pomar F., Carrion J.S., Pedreño M.A. Basic peroxidases: the gateway for lignin evolution? *Phytochem. Rev.* 2004;3:61-78. DOI 10.1023/B:PHYT.0000047803.49815.1a/.
- Berthet S., Thevenin J., Baratin D., Demont-Caulet N., Debeaujon I., Bidzinski P., Leple J.C., Huis R., Hawkins S., Gomez L.D., Lapiere C., Jouanin L. Role of plant laccases in lignin polymerization. *Adv. Botan. Res.* 2012;61:145-172. DOI 10.1016/B978-0-12-416023-1.00005-7.
- Bindschedler L.V., Dewdney J., Blee K.A., Stone J.M., Asai T., Plotnikov J., Denoux C., Hayes T., Gerrish C., Davies D.R., Ausubel F.M., Bolwell G.P. Peroxidase-dependent apoplastic oxidative burst in *Arabidopsis* required for pathogen resistance. *Plant J.* 2006;47:851-863. DOI 10.1111/j.1365-3113X.2006.02837.x.
- Boerjan W., Ralph J., Baucher M. Lignin biosynthesis. *Annu. Rev. Plant Biol.* 2003;54:519-546. DOI 10.1146/annurev.arplant.54.031902.134938.
- Chassot C., Nawrath C., Metraux J.P. Cuticular defects lead to full immunity to a major plant pathogen. *Plant J.* 2007;49:972-980. DOI 10.1111/j.1365-3113X.2006.03017.x.
- Cosio C., Dunand C. Specific functions of individual class III peroxidase genes. *J. Exp. Bot.* 2009;2(60):391-408. DOI 10.1093/jxb/ern318.
- Fernandez-Pereza F., Vivara T., Pomar F., Pedreño M.A., Novo-Uzal E. Peroxidase 4 is involved in syringyl lignin formation in *Arabidopsis thaliana*. *J. Plant Physiol.* 2015;175:86-94. DOI 10.1016/j.jplph.2014.11.006.
- Gabaldon C., Lopez-Serrano M., Pedreño M.A., Barcelo A.R. Cloning and molecular characterization of the basic peroxidase isoenzyme from *Zinnia elegans*, an enzyme involved in lignin biosynthesis. *Plant Physiol.* 2005;3(139):1138-1154. DOI 10.1104/pp.105.069674.
- Gabaldon T., Koonin E.V. Functional and evolutionary implications of gene orthology. *Nat. Rev. Genet.* 2013;14(5):360-366. DOI 10.1038/nrg3456.
- Herrero J., Esteban-Carrasco A., Zapata J.M. Looking for *Arabidopsis thaliana* peroxidases involved in lignin biosynthesis. *Plant Physiol. Biochem.* 2013a;67:77-86. DOI 10.1016/j.plaphy.2013.02.019.
- Herrero J., Fernandez-Perez F., Yebra T., Novo-Uzal E., Pomar F., Pedreño M.A., Cuello J., Guera A., Esteban-Carrasco A., Zapata J.M. Bioinformatic and functional characterization of the basic peroxidase 72 from *Arabidopsis thaliana* involved in lignin biosynthesis. *Planta.* 2013b;6(237):1599-1612. DOI 10.1007/s00425-013-1865-5.
- Hiraga S., Sasaki K., Ito H., Ohashi Y., Matsui H. A large family of class III plant peroxidases. *Plant Cell Physiol.* 2001;5(42):462-468. DOI 10.1093/pcp/pce061.
- Irshad M., Canut H., Borderies G., Pont-Lezica R., Jamet E. A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. *BMC Plant Biol.* 2008;8:94. DOI 10.1186/1471-2229-8-94.
- Jovanovic S.V., Kukavica B., Vidovic M., Morina F., Menckhoff L. Class III peroxidases: functions, localization and redox regulation of isoenzymes. *Antioxidants and Antioxidant Enzymes in Higher Plants.* 2018;269-300. DOI 10.1007/978-3-319-75088-0_13.
- Kraimer F.W., Pletzenauer R., Rossetti L., Herwig C., Glieder A., Spadiut O. Purification and basic biochemical characterization of 19 recombinant plant peroxidase isoenzymes produced in *Pichia pastoris*. *Protein Expr. Purif.* 2014;100(95):104-112. DOI 10.1016/j.pep.2013.12.003.
- Kukavica B., Veljovic-Jovanovic S., Menckhoff L., Luthje S. Cell wall-bound cationic and anionic class III isoperoxidases of pea root: biochemical characterization and function in root growth. *J. Exp. Bot.* 2012;63:4631-4645. DOI 10.1093/jxb/ers139.
- Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 2016;33(7):1870-1874. DOI 10.1093/molbev/msw054.
- Kunieda T., Shimada T., Kondo M., Nishimura M., Nishitani K., Hara-Nishimura I. Spatiotemporal secretion of PEROXIDASE36 is required for seed coat mucilage extrusion in *Arabidopsis*. *Plant Cell.* 2013;4(25):1355-1367. DOI 10.1105/tpc.113.110072.
- Liu Q., Luo L., Zheng L. Lignins: biosynthesis and biological functions in plants. *Int. J. Mol. Sci.* 2018;19(2):335. DOI 10.3390/ijms19020335.
- Llorente F., Lopez-Cobollo R.M., Catala R., Martinez-Zapater J.M., Salinas J. A novel cold-inducible gene from *Arabidopsis*, RC13, encodes a peroxidase that constitutes a component for stress tolerance. *Plant J.* 2002;32:13-24. DOI 10.1046/j.1365-3113X.2002.01398.x.
- Mansouri I.E., Mercado J.A., Santiago-Domenech N., Pliego-Alfaro F., Valpuesta V., Quesada M.A. Biochemical and phenotypical characterization of transgenic tomato plants overexpressing a basic peroxidase. *Physiol. Plant.* 1999;106:355-362. DOI 10.1034/j.1399-3054.1999.106401.x.
- Marjamaa K., Kukkola E.M., Fagerstedt K.V. The role of xylem class III peroxidases in lignification. *J. Exp. Bot.* 2009;60(2):367-376. DOI 10.1093/jxb/ern278.
- Nei M., Kumar S. *Molecular Evolution and Phylogenetics.* New York: Oxford University Press, 2000.
- Ostergaard L., Teilum K., Mirza O., Mattsson O., Petersen M., Welinder K.G., Mundy J., Gajhede M., Henriksen A. *Arabidopsis* ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. *Plant Mol. Biol.* 2000;44:231-243. DOI 10.1023/A:1006442618860.
- Passardi F., Longet D., Penel C., Dunand C. The class III peroxidase multigenic family in rice and its evolution in land plants. *Phytochemistry.* 2004a;65:1879-1893. DOI 10.1016/j.phytochem.2004.06.023.
- Passardi F., Penel C., Dunand C. Performing the paradoxical: how plant peroxidases modify the cell wall. *Trends Plant Sci.* 2004b;9:534-540. DOI 10.1016/j.tplants.2004.09.002.
- Pedreira J., Herrera M.T., Zarra I., Revilla G. The overexpression of *AtPrx37*, an apoplastic peroxidase, reduces growth in *Arabidopsis*. *Physiol. Plant.* 2011;141:177-187. DOI 10.1111/j.1399-3054.2010.01427.x.
- Quiroga M., Guerrero C., Botella M.A., Barcelo A., Amaya I., Medina M.I., Alonso F.J., Milrad de Forchetti S., Tigier H., Valpuesta V. A tomato peroxidase involved in the synthesis of lignin and suberin. *Plant Physiol.* 2000;122:1119-1127. DOI 10.1104/pp.122.4.1119.
- Sanou N., Nei M. The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 1981;4:406-425. DOI 10.1093/oxfordjournals.molbev.a040454.
- Sasaki S., Nishida T., Tsutsumi Y., Kondo R. Lignin dehydrogenative polymerization mechanism: a poplar cell wall peroxidase directly oxidizes polymer lignin and produces *in vitro* dehydrogenative polymer rich in beta-O-4 linkage. *FEBS Lett.* 2004;562:197-201. DOI 10.1016/S0014-5793(04)00224-8.
- Sato Y., Demura T., Yamawaki K., Inoue Y., Sato S., Sugiyama M., Fukuda H. Isolation and characterization of a novel peroxidase gene *ZPO-C* whose expression and function are closely associated with

- lignification during tracheary element differentiation. *Plant Cell Physiol.* 2006;4(47):493-503. DOI 10.1093/pcp/pcj016.
- Shigeto J., Kiyonaga Y., Fujita K., Kondo R., Tsutsumi Y. Putative cationic cell-wall-bound peroxidase homologues in *Arabidopsis*, AtPrx2, AtPrx25, and AtPrx71, are involved in lignification. *J. Agric. Food Chem.* 2013;16(61):3781-3788. DOI 10.1021/jf400426g.
- Tsakagoshi H., Busch W., Benfey P.N. Transcriptional regulation of ROS controls transition from proliferation to differentiation in the root. *Cell.* 2010;4(143):606-616. DOI 10.1016/j.cell.2010.10.020.
- Valerio L., De Meyer M., Penel C., Dunand C. Expression analysis of the *Arabidopsis* peroxidase multigenic family. *Phytochemistry.* 2004;65:1331-1342. DOI 10.1016/j.phytochem.2004.04.017.
- Welinder K.G., Justesen A.F., Kjaersgard I.V.H., Jensen R.B., Rasmussen S.K., Jespersen H.M., Duroux L. Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. *Eur. J. Biochem.* 2002;269:6063-6081. DOI 10.1046/j.1432-1033.2002.03311.x.
- Yokoyama R., Nishitani K. Identification and characterization of *Arabidopsis thaliana* genes involved in xylem secondary cell walls. *J. Plant Res.* 2006;119:189-194. DOI 10.1007/s10265-006-0261-7.

ORCID ID

A.S. Tugbaeva orcid.org/0000-0001-9231-3650
A.A. Ermoshin orcid.org/0000-0003-2784-4898
I.S. Kiseleva orcid.org/0000-0002-8709-5203

Acknowledgements. The work was supported by the Ministry of Science and Higher Education of the Russian Federation (No. 075-02-2019-980).

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

Received March 14, 2019. Revised May 13, 2019. Accepted May 15, 2019.

Тернистый путь макрофаг-активирующего фактора (GcMAF): от открытия к клинической практике

А.А. Останин¹, С.С. Кирикович², Е.В. Долгова², А.С. Проскурина², Е.Р. Черных¹, С.С. Богачев² 

¹ Научно-исследовательский институт фундаментальной и клинической иммунологии, Новосибирск, Россия

² Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

 e-mail: labmolbiol@mail.ru

Витамин D₃-связывающий белок (DBP) представляет собой полифункциональный гликопротеин, основная роль которого заключается в транспорте витамина D₃ и его метаболитов; но он также является предшественником макрофаг-активирующего фактора (GcMAF). DBP конвертируется в GcMAF в результате сайт-специфического селективного дегликозилирования под действием β-галактозидазы и сиалидазы, локализованных на активированных В- и Т-лимфоцитах соответственно. Биологическая активность GcMAF выражается, прежде всего, в его способности активировать макрофаги, усиливая их фагоцитарную функцию и продукцию реактивных форм кислорода. В результате активации на макрофагах повышается экспрессия специфических рецепторов, участвующих в распознавании опухоль-ассоциированных антигенов, а также в реализации прямой противораковой активности через индукцию апоптоза/некроза опухолевых клеток. Повышенный интерес к GcMAF связан с его потенциальной возможностью использования в клинике в качестве нового противоопухолевого препарата. Роль GcMAF проявляется не только при онкологических, но и при целом ряде вирусных и нейродегенеративных заболеваний, при которых в сыворотке больных повышена активность N-ацетилгалактозаминидазы (нагалазы). Нагалаза – это фермент, который полностью, а не селективно дегликозилирует DBP и блокирует, таким образом, образование GcMAF, что приводит к иммунным нарушениям. В обзоре подробно рассмотрены современные данные о структуре и функциях DBP как основного предшественника GcMAF. По своему составу находящийся в циркуляции DBP – это смесь немодифицированных и О-гликозилированных молекул, степень гликозилирования которых определяется генотипом по гену, кодирующему DBP. На роль DBP в устойчивости организма к ряду заболеваний указывает тот факт, что у индивидуумов, гомозиготных по аллелю, кодирующему дефектный DBP, не образуется ни одной молекулы GcMAF, вследствие чего эти индивидуумы имеют высокий риск развития различных тяжелых заболеваний (боковой амиотрофический склероз, колоректальный рак и др.). В обзоре представлены данные об основных механизмах противоопухолевого эффекта GcMAF, опухолевой стратегии нейтрализации активности GcMAF, результаты клинических испытаний GcMAF при различных нозологических формах рака, а также обсуждены имеющиеся противоречия относительно позиционирования GcMAF в качестве эффективного противоопухолевого препарата.

Ключевые слова: витамин D₃-связывающий белок (DBP); макрофаг-активирующий фактор (GcMAF); N-ацетилгалактозамин (GalNAc); α-N-ацетилгалактозаминидаза (нагалаза); противоопухолевая терапия.

Для цитирования: Останин А.А., Кирикович С.С., Долгова Е.В., Проскурина А.С., Черных Е.Р., Богачев С.С. Тернистый путь макрофаг-активирующего фактора (GcMAF): от открытия к клинической практике. Вавиловский журнал генетики и селекции. 2019;23(5):624-631. DOI 10.18699/VJ19.535

A thorny pathway of macrophage activating factor (GcMAF): from bench to bedside

A.A. Ostanin¹, S.S. Kirikovich², E.V. Dolgova², A.S. Proskurina², E.R. Chernykh¹, S.S. Bogachev² 

¹ Institute of Fundamental and Clinical Immunology, Novosibirsk, Russia

² Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

 e-mail: labmolbiol@mail.ru

Vitamin D₃ Binding Protein (DBP) is a multifunctional glycoprotein whose main role is to transport vitamin D₃ and its metabolites, but it also is the precursor of the macrophage activating factor (GcMAF). DBP is converted to GcMAF as a result of site-specific selective deglycosylation under the action of β-galactosidase and sialidase, localized on activated B and T cells, respectively. GcMAF exerts its biological activity primarily as the capability of activating macrophages by enhancing their phagocytic function and producing ROS. Activation results in elevated expression of the specific macrophageal surface receptors involved in the recognition of tumor-associated antigens, as well as in the implementation of direct anticancer activity by inducing the apoptosis or necrosis of tumor cells. Increased interest in GcMAF is associated with its potential to be used in the clinic as a new antitumor drug. Besides its anti-tumor activity, GcMAF exerts a potential against a number of viral and neurodegenerative diseases associated with increased activity of N-acetylgalactosaminidase (nagalase) in the blood serum of patients. Nagalase is an enzyme that completely (rather than selectively)

deglycosylates DBP so it cannot be converted to GcMAF, leading to immunodeficiency. Circulating DBP is composed of unmodified and O-glycosylated molecules with the glycosylation degree being dependent on the allelic variants of the gene encoding DBP. The role of DBP in the resistance of organism against a number of diseases is supported by the increased risk of a variety of severe illnesses (amyotrophic lateral sclerosis, colorectal cancer etc.) in patients deficient for GcMAF due to homozygosity for defective DBP alleles. In this review, we also will examine in detail the current data i) on the structure and functions of DBP, as the main precursor of GcMAF, ii) on the main mechanisms of GcMAF anticancer effect, iii) on the tumor strategy for neutralizing GcMAF activity, iv) on the results of GcMAF clinical trials in various cancers; and will discuss the available controversies regarding the positioning of GcMAF as an effective antitumor drug. Key words: vitamin D₃-binding protein (DBP); Gc protein-derived macrophage activating factor (GcMAF); N-acetylgalactosamine (GalNAc); α-N-acetylgalactosaminidase (nagalase); anticancer therapy.

For citation: Ostanin A.A., Kirikovich S.S., Dolgova E.V., Proskurina A.S., Chernykh E.R., Bogachev S.S. A thorny pathway of macrophage activating factor (GcMAF): from bench to bedside. *Vavilovskii Zhurnal Genetiki i Selekcii* = Vavilov Journal of Genetics and Breeding. 2019;23(5):624–631. DOI 10.18699/VJ19.535 (in Russian)

Введение

Специфический активатор макрофагов, GcMAF, – компонент плазмы крови, который образуется в результате сайт-специфического селективного дегликозилирования витамин D₃-связывающего белка (DBP) под действием ферментов β-галактозидазы и сиалидазы, локализованных на клеточных мембранах активированных В- и Т-лимфоцитов соответственно. Многочисленные экспериментальные данные, а также результаты пилотных клинических исследований свидетельствуют о выраженной противоопухолевой активности GcMAF, что открывает широкие перспективы его потенциального использования в качестве нового лекарственного средства при онкологических заболеваниях.

Цель настоящего обзора – представление данных о структуре и функции DBP как предшественника GcMAF, основных механизмах противоопухолевого эффекта GcMAF, опухолевой стратегии нейтрализации активности GcMAF, результатах клинических испытаний GcMAF при различных нозологических формах рака, а также обсуждение имеющихся критических замечаний в отношении позиционирования GcMAF в качестве эффективного противоопухолевого препарата.

Структура и функции DBP как предшественника GcMAF

Витамин D₃-связывающий белок (DBP) – это полифункциональный гликопротеин, относящийся к семейству белков крови (Group-specific component, Gc-белки размером 51–58 кДа), представленных в том числе альбумином, α-фетопротеином и афамином (α-альбумин, витамин Е-связывающий белок). DBP синтезируется гепатоцитами и попадает в кровоток в форме зрелого мономера, несущего три функциональных домена. Домен, связывающий витамин D₃ и жирные кислоты, расположен между аминокислотными остатками 35–49, актин-связывающий домен находится между аминокислотными остатками 350–403. Два сайта связывания с клеточной мембраной нейтрофилов расположены на N- и C-концах молекулы гликопротеина (Haddad et al., 1992; Otterbein et al., 2002; Verboven et al., 2002; Malik et al., 2013).

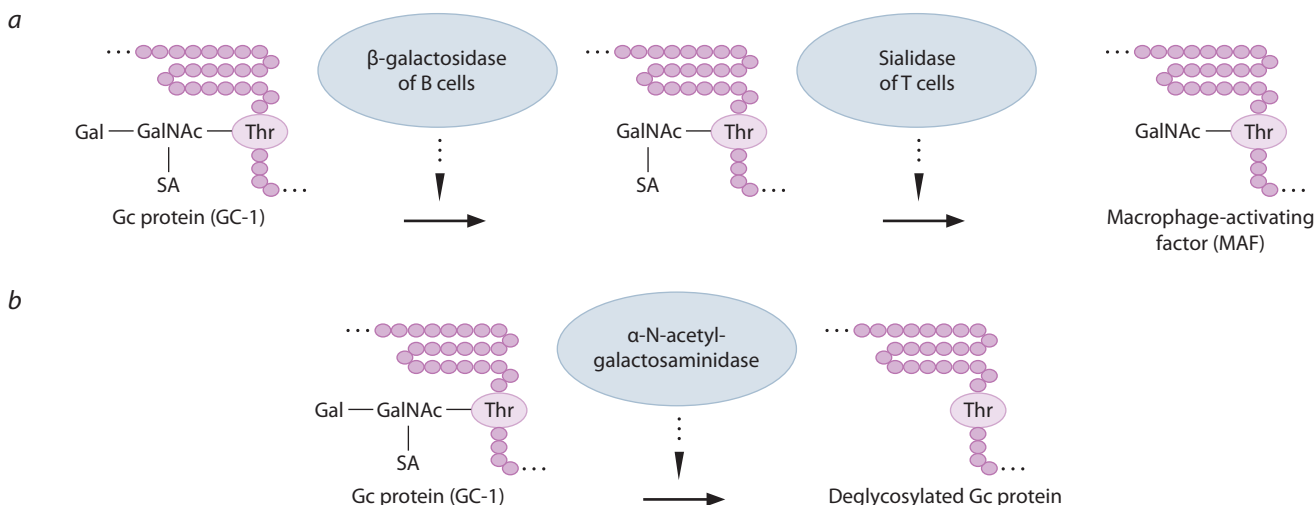
Белок DBP кодируется геном, находящимся на длинном плече 4-й хромосомы (4q11-q13) и представленным единственной копией (NCBI ID: 2638), состоящей из 15 экзонов и 12 интронов (Song et al., 1999). DBP присутствует в плазме крови человека в большом количестве

(300–600 мг/л) и выступает в качестве основного переносчика витамина D₃ и его производных. К другим функциям DBP относятся: а) связывание и выведение актина, высвобождающегося при некрозе клеток, а также выведение жирных кислот и бактериальных эндотоксинов; б) активация хемотаксиса нейтрофилов; в) активация Т-лимфоцитов через механизм макропиноцитоза; г) участие в метаболизме костной ткани. DBP обнаруживается в различных жидкостях – сыворотке крови, моче, грудном молоке, асцитной жидкости, ликворе, слюне, эякуляте, а также органах человека – мозге, сердце, легких, селезенке, почках, плаценте, семенниках, матке (Malik et al., 2013; Delanghe et al., 2015; Morales, 2017).

Еще одна очень важная функция DBP – его способность активировать макрофаги, которую Gc-белок приобретает в результате сайт-специфического селективного дегликозилирования (рисунок, а).

Гликозилированный DBP содержит один трисахарид, ковалентно связанный с треонином в 420-й позиции и состоящий из N-ацетилгалактозамина (GalNAc) с двумя разветвленными остатками сахаров галактозы и сиаловой кислоты. DBP конвертируется в GcMAF под действием ферментов β-галактозидазы и сиалидазы, локализованных на клеточных мембранах активированных В- и Т-лимфоцитов соответственно. В результате образуется активный белок GcMAF, содержащий остаточный сахар N-ацетилгалактозамин с освободившимися от галактозы и сиаловой кислоты сайтами связывания (Yamamoto, Kumashiro, 1993; Yamamoto et al., 2008a–c). Следует отметить, что такое селективное дегликозилирование DBP происходит естественным образом при развитии воспалительного ответа. Считается, что именно GalNAc в составе GcMAF обеспечивает активацию макрофагов. Полное дегликозилирование DBP под действием фермента N-ацетилгалактозаминидазы (нагалазы) приводит к разрыву связи GalNAc с треонином (см. рисунок, б), и таким образом нагалаза блокирует DBP → GcMAF конверсию.

По своему составу находящийся в циркуляции DBP представляет собой смесь немодифицированных и О-гликозилированных молекул, степень гликозилирования которых определяется генотипом DBP (Malik et al., 2013). Для человека описаны три мажорных аллельных варианта белка: DBP1F, DBP1S и DBP2. Варианты DBP1F и DBP1S могут быть конвертированы в GcMAF под действием β-галактозидазы и сиалидазы, тогда как вариант DBP2, в котором треонин в 420-й позиции заменен на лизин, не



Schematic presentation of DBP deglycosylation (Yamamoto et al., 1996): (a) selective, (b) complete.

может связывать GalNAc, поскольку в нем отсутствует основной сайт О-связанного гликозилирования сахаридов. Поскольку у индивидуумов, гомозиготных по аллелю DBP2, в принципе не может образоваться ни одной молекулы GcMAF, не удивительно, что именно вариант DBP2/DBP2 полиморфизма ассоциируется с повышенным риском развития различных тяжелых заболеваний, таких как боковой амиотрофический склероз, колоректальный рак и др. (Morales, 2017).

Механизмы противоопухолевого эффекта GcMAF

Биологическая активность GcMAF проявляется в его способности активировать макрофаги, усиливая их фагоцитарную функцию и продукцию реактивных форм кислорода. Кроме того, GcMAF стимулирует пролиферацию миелоидных клеток-предшественников и индуцирует их дифференцировку в зрелые макрофаги (Yamamoto, Homma, 1991; Homma et al., 1993; Mohamad et al., 2002a, b; Yamamoto et al., 2008a–c; Uto et al., 2012; Thyer et al., 2013a; Ishikawa et al., 2014). В результате активации на макрофагах повышается экспрессия специфических рецепторов, участвующих в распознавании и презентации опухоль-ассоциированных антигенов, а также в реализации прямой противораковой активности через индукцию апоптоза/некроза опухолевых клеток (Yamamoto et al., 2008a–c; Rehder et al., 2009; Сахно и др., 2016). Так, например, показано, что макрофаги человека, активированные GcMAF в дозе 100 пг/мл, индуцируют гибель 51 и 82 % клеток линии LNCaP (рак простаты человека) через 4 и 18 ч сокультивирования соответственно (Yamamoto et al., 2008b). В культуре *in vitro* GcMAF усиливает также дифференцировку «профессиональных» антиген-презентирующих дендритных клеток (ДК), что проявляется значимым увеличением зрелых HLA-DR⁺CD86⁺ДК до уровня > 80 % в общей популяции.

В работе (Gregory et al., 2010) показано, что GcMAF характеризуется прямым ингибиторным действием на пролиферацию и миграцию клеток линий рака простаты (LNCaP и PC3), а также снижает экспрессию на них uPAR

(urokinase plasminogen activator receptor), активация которого коррелирует с метастазированием опухоли. GcMAF в дозе 40 нг/мл ингибирует пролиферацию клеток линии аденокарциномы молочной железы (MCF-7) в среднем на 50 %. При этом регистрируется снижение экспрессии опухолевыми клетками виментина, маркера прогрессии и метастазирования рака молочной железы, вызванного трансформацией эпителиальных клеток и приобретением ими фенотипа мезенхимальных клеток с повышенной устойчивостью к апоптозу, высокой миграционной активностью и инвазивностью (феномен эпителиально-мезенхимального транзита) (Pacini et al., 2012b; Thyer et al., 2013b). В экспериментах с золотистыми сирийскими хомяками в модели канцерогенеза, индуцированного 9,10-диметил-1,2-бензантраценом (DMBA), также был выявлен ингибирующий эффект GcMAF на рост опухоли и увеличение продолжительности жизни животных (Toyohara et al., 2011).

Как известно, рост опухоли – это ангиогенез-зависимый процесс. Поэтому еще одним очень важным механизмом противоопухолевого действия GcMAF являются его анти-ангиогенная активность и способность ингибировать пролиферацию и миграцию эндотелиальных клеток, а также образование новых микрососудов в опухоли (Nonaka et al., 2012). В исследовании (Pacini et al., 2011, 2012a) с использованием анализа хориоаллантоической мембраны (chorioallantoic membrane [CAM] assay) было показано, что GcMAF в концентрации 1 нг/мл ингибирует стимулированный простагландин Е₁ ангиогенез у куриных эмбрионов.

Противоопухолевая активность GcMAF была продемонстрирована в различных экспериментальных моделях *in vivo*. Kisker с коллегами (2003) показали, что ежедневное интраперитонеальное введение GcMAF в дозе 4 нг/кг ингибирует рост графта рака поджелудочной железы человека (В × PC-3; 2.5 × 10⁶ клеток; п/к) у иммунодефицитных SCID мышей. При использовании GcMAF в дозе 4 мкг/кг наблюдалась полная регрессия опухолевого трансплантата. Гистологические исследования выявили выраженную инфильтрацию графта Mac-3⁺ макрофагами, меньшую

плотность капилляров и более высокий уровень апоптоза по сравнению с контролем (Kisker et al., 2003). Пятнадцатичасовая инкубация перитонеальных макрофагов ICR мышей с сывороткой человека, обогащенной GcMAF в дозе 10 нг по белку, приводила к значимому усилению их фагоцитарной активности в среднем на 73 %. Интраперитонеальное введение такой сыворотки (в дозе 1.552 мкг/кг/день в течение 7 дней) мышам с привитым асцитом карциномы Эрлиха (10×10^6 клеток/мышь) сопровождалось достоверным увеличением продолжительности жизни (Kuchiiike et al., 2013). В работе (Korbelik et al., 1997) на мышинной модели SCCVII (плоскоклеточная карцинома) было показано, что сочетание фотодинамической терапии (ФДТ) с адъювантной терапией GcMAF в виде комбинации внутривентральных и перитуморальных инъекций (по 50 и 0.5 нг/кг⁻¹ соответственно), вводимых в дни 0, 4, 8 и 12 после ФДТ, приводило к 100 % выживаемости животных против 25 % при использовании ФДТ в режиме монотерапии.

Анализ данных литературы позволяет заключить, что противоопухолевый эффект GcMAF реализуется не только через активацию опухоль-инфильтрирующих макрофагов (ОИМ), но и с участием других механизмов, связанных с прямым ингибиторным действием GcMAF на пролиферацию, миграцию и метастазирование опухолевых клеток, а также с подавлением неоангиогенеза в опухолевой ткани.

Опухолевая стратегия нейтрализации активности опухоль-инфильтрирующих макрофагов и GcMAF

Хорошо известно, что в процессе канцерогенеза трансформированные опухолевые клетки используют различные стратегии, позволяющие им как ускользнуть от иммунного надзора, так и подавлять развитие эффективного противоопухолевого иммунного ответа (Dunn et al., 2002; Kim et al., 2007). Одна из этих стратегий направлена, в частности, на нейтрализацию опухолецидной активности макрофагов в опухолевом микроокружении. Под действием TGF- β 1, секретируемого опухолевыми клетками, способность ОИМ к фагоцитозу снижается, и они приобретают фенотип макрофагов 2-го типа (M2-клетки), который ориентирован на поддержание роста опухоли (Tourtirais et al., 2003; Allavena et al., 2008). M2-клетки – одна из субпопуляций тканевых макрофагов, для которых характерна экспрессия аргиназы и секреция IL-10, простагландина E₂ и самого TGF- β 1, поддерживающего данный фенотип ОИМ через аутокринный механизм. M2-клетки отличаются также повышенной экспрессией фактора роста эндотелия сосудов (VEGF), активирующего ангиогенез и являющегося одним из факторов опухолевой иммуносупрессии (Martinez et al., 2006). Совместное воздействие TGF- β 1 и VEGF инициирует формирование опухолевой стромы, необходимой для обеспечения трофики опухолевых клеток, локальной иммуносупрессии и поддержки процесса метастазирования (Ohm, Carbone, 2001; Mantovani et al., 2002).

Вирус-инфицированные, а также опухолевые клетки способны секретировать фермент нагалазу (экзо- и эндотипа), который также может участвовать в перепрограммировании ОИМ в M2-направлении (Mohamad et al., 2002a). Экзотип нагалазы в физиологических условиях

секретируется нормальными гепатоцитами и не дегликозилирует трисахарид DBP, тем самым не влияет на конвертацию этого предшественника в GcMAF (Ioannou et al., 1992; Nagasawa et al., 2005). В норме активность фермента в плазме крови варьирует от 0.35 до 0.65 нМ/мин/мг субстрата. При различных онкологических заболеваниях сывороточный уровень нагалазы значительно повышается и регистрируется в диапазоне от 0.63 до 5.21 нМ/мин/мг (Yamamoto et al., 1996; McCarty, 2013). Нагалаза экзо- и эндотипа, продуцируемая опухолевыми клетками, может полностью дегликозилировать как трисахарид DBP, так и остаточный сахар N-ацетилгалактозамин GcMAF (Yamamoto et al., 1996; Mohamad et al., 2002a; Matsuura et al., 2004). В опухолевом микроокружении возможны оба варианта тотального дегликозилирования, в результате которого либо блокируется конвертация DBP в GcMAF, либо инактивируется уже существующий GcMAF (Yamamoto et al., 1996; Saburi et al., 2017). Дегликозилированный GcMAF не в состоянии активировать опухолеинфильтрирующие макрофаги, что клинически проявляется развитием иммуносупрессии и прогрессией заболевания, связанной с потерей макрофагами противоопухолевой активности (Nagasawa et al., 2005; McCarty, 2013).

Результаты клинических испытаний GcMAF

Первые результаты открытых, пилотных клинических исследований препарата GcMAF 1-й генерации были опубликованы в 2008 г. (Yamamoto et al., 2008a–c). Следует отметить, что N. Yamamoto еще в 1991 г. разработал способ получения GcMAF 1-й генерации методом аффинной хроматографии на колонках, нагруженных 25-гидроксигалактозином D₃, с последующим дегликозилированием выделенного DBP растворимыми или иммобилизованными на носителе ферментами (β -галактозидазой и сиалидазой) (Yamamoto, Homma, 1991). Получаемый продукт характеризовался низкой концентрацией (100 нг/0.25 мл, 1 доза), нестабильностью при комнатной температуре и при отсутствии антиоксидантов в виде альбумина и мочевины. В своих пионерных работах N. Yamamoto с коллегами использовали GcMAF в дозе 100 нг, который вводили больным внутримышечно, 1 раз/нед, курсом от 3.5 до 6 мес. Было пролечено 16 больных раком простаты в возрасте от 46 до 76 лет (медиана, Me – 63.5), 16 больных раком молочной железы с метастазами в возрасте от 44 до 77 лет (Me – 62.5) и 8 больных колоректальным раком с метастазами в возрасте от 41 до 82 лет (Me – 65.5). Исследователи показали, что терапия GcMAF приводит к снижению до нормы активности сывороточной нагалазы и полному излечению пациентов. Продолжительность безрецидивного периода для рака молочной железы составила более 4 лет, для рака простаты и колоректального рака – 7 лет. Снижение нагалазы в процессе терапии свидетельствовало, по мнению авторов, об эрадикации опухоли, что для рака простаты было подтверждено томографическим сканированием.

В 2009 г. N. Yamamoto с коллегами опубликовали результаты иммунотерапии с использованием GcMAF 1-й генерации у ВИЧ-инфицированных больных ($n = 15$) (Yamamoto et al., 2009). Препарат вводили больным также внутримышечно, в дозе 100 нг, 1 раз/нед в течение

4.5 мес. После окончания курсового лечения у больных регистрировались снижение активности нагалазы в сыворотке крови, увеличение количества циркулирующих CD4+Т-лимфоцитов, нормализация индекса CD4/CD8. При этом нормальное содержание Т-клеток в кровотоке сохранялось у больных в течение 7 лет наблюдения.

Полученные группой Yamamoto результаты были настолько многообещающими, что послужили своего рода катализатором для дальнейших исследований клинической эффективности GcMAF. В 2013 г. были опубликованы данные по использованию препарата GcMAF 1-й генерации (в дозе 100 нг, в/м, 1 раз/нед) в группе 20 больных (9 мужчин и 11 женщин в возрасте от 42 до 76 лет) с продвинутыми стадиями рака различной локализации (мочевой пузырь, яичники, простата, молочная железа и т.д.). Было показано, что активность сывороточной нагалазы у больных значимо снижалась с исходного уровня 2.8 ± 0.26 до 1.6 ± 0.17 нМ/мин/мг после окончания иммунотерапии ($p < 0.01$), при этом интервал между обследованиями составлял в среднем 263 ± 45 дней (Thyer et al., 2013a).

Следует отметить, что в 2010 г. сотрудниками Университета Токушима и частной клиники Saisei Mirai (Япония) был разработан и запатентован способ получения GcMAF 2-й генерации (Uto et al., 2012). В этом случае GcMAF выделяют путем ферментативной обработки сыворотки конкретного больного, индивидуально без очистки аффинной хроматографией на 25-гидроксивитамин Д₃ колонках. GcMAF 2-й генерации отличается от препарата первого поколения более высокой концентрацией (1500 нг/0.5 мл/1 доза) и стабильностью. Было предложено использовать препарат GcMAF 2-й генерации в качестве одного из компонентов интегративной иммунотерапии рака (Inui et al., 2013). Был разработан протокол, включающий: 1) GcMAF 1500 нг/0.5 мл; в/м или п/к; 1–2 раза/нед; в течение всего курса иммунотерапии; 2) клеточная терапия с использованием Т- и NK-клеток (Natural Killer cells), 1 раз/нед, $n = 6$; 3) высокие дозы VitC (50–100 г); в/в; 2 раза/нед; 4) альфа-липоевая кислота (600 мг); per os; ежедневно; 5) VitD3 (5–10 000 МЕ); per os; ежедневно. По данной схеме с апреля 2011 г. по март 2013 г. на базе клиники Saisei Mirai было пролечено более 345 больных с различными нозологическими формами рака. Но в опубликованной работе авторы приводят только три клинических наблюдения, которые, по их мнению, наиболее убедительно свидетельствуют об эффективности предложенного подхода.

Пациент 1. М/71 год, диагноз: карцинома тимуса с метастазами в легкие. Интегративная иммунотерапия в течение 6 мес. Результат: отсутствие опухолевой прогрессии в течение 12 мес. после завершения терапии.

Пациент 2. М/74 года, диагноз: рак простаты с множественными метастазами в кости. Интегративная иммунотерапия в течение 3 мес. в сочетании с гипертермией. Результат: через 6 мес. после завершения терапии скинтиграммы костей в норме, метастазы не обнаруживаются.

Пациент 3. Ж/72 года, диагноз: метастатический рак печени после удаления сигмовидной кишки и двусторонней овариэктомии. Интегративная иммунотерапия в течение 6 мес. в сочетании с радиотерапией (55 Гр). Результат: в течение 6 мес. после завершения терапии

отсутствие рецидива по данным позитронной (ПЭТ) и компьютерной (КТ) томографии.

Накопленный опыт клинического применения препарата GcMAF свидетельствует о его безопасности, поскольку он нетоксичен и его длительное применение не приводит к появлению побочных нежелательных явлений. Основная концепция противоопухолевого эффекта GcMAF состоит в активации макрофагов, редукции опухоли вплоть до ее полной эрадикации, что проявляется снижением активности сывороточной нагалазы как диагностического маркера эффективности терапии.

Следует отметить еще одну область клинического применения GcMAF. Bradstreet с коллегами при обследовании когорты из 40 детей (32 мальчика и 8 девочек в возрасте от 1.5 до 21 лет; Me – 7 лет) с заболеваниями аутистического спектра (ASD) обнаружили, что активность сывороточной нагалазы выходит за верхнюю границу нормативного диапазона (> 0.95 нМ/мин/мг) в 95 % случаев (у 38/40 детей) и составляет в среднем по группе 1.93 нМ/мин/мг (Bradstreet et al., 2012). После курсового лечения в виде внутримышечного введения препарата GcMAF (1 раз/нед, в течение 3.5 мес., с постепенным увеличением дозы от 4 до 100 нг) активность нагалазы снизилась у всех больных ASD (за исключением одного ребенка) и составляла в среднем 1.03 нМ/мин/мг ($p < 0.0001$). На момент повторного тестирования у 24 из 40 детей (60 %) уровень нагалазы находился в границах нормативного диапазона (≤ 0.95 нМ/мин/мг). Оставшихся 16 пациентов (40 %), у которых полной нормализации активности сывороточной нагалазы достичь не удалось, авторы рассматривают в качестве кандидатов на продолжение курса GcMAF терапии. В процессе лечения не было отмечено развития каких-либо серьезных побочных осложнений, но при этом у значительной части детей регистрировалось улучшение разговорной речи, социализации и когнитивных функций.

Является ли GcMAF эффективным противоопухолевым препаратом?

Следует сказать, что пионерные работы N. Yamamoto (Yamamoto et al., 2008a–c, 2009) не только послужили катализатором дальнейших исследований клинической эффективности GcMAF, но вызвали также определенную волну критических замечаний. Так, A. Ugarte с коллегами в своем письме редакторам журнала Cancer Immunology, Immunotherapy указывают на противоречивость данных группы Yamamoto и сомневаются в надежности полученных ими результатов (Ugarte et al., 2014). Высказывались следующие замечания: 1) нет базовой информации по больным (диагноза по TNM, стадии/тяжести рака, гистологии); 2) нет контрольной группы; 3) о наличии метастазов авторы судят по повышенному уровню нагалазы, что не является общепризнанным в онкологии; 4) авторы активно себя самоцитируют; 5) разрешительные документы на проведение клинических испытаний не достоверны; 6) упоминаемые в статье спонсоры на самом деле не поддерживали эти исследования; 7) в статье много неточностей и ошибок, в том числе некорректных ссылок. Сделано заключение: полученные результаты не являются научно подтвержденными. В результате по решению редакторских советов три статьи N. Yamamoto с соавто-

рами (2008а, с, 2009) были отозваны после публикации в журналах с формулировкой «некорректно проведенные клинические испытания и некорректно оформленная документация, сопровождающая эти испытания».

Существуют определенные факты, противоречащие рабочей гипотезе, на которой базируется GcMAF-терапия. В пионерных работах N. Yamamoto было показано, что у обследованных больных раком снижен эндогенный уровень GcMAF, поскольку в биологическом тесте обработка макрофагов сывороткой этих больных не стимулирует образование активных форм кислорода (Yamamoto et al., 2008а–с, 2009). В то же время D.S. Rehder с коллегами, используя масс-спектральный анализ, показали, что количество GalNAc DBP (что принято считать GcMAF) в сыворотке крови больных раком и здоровых доноров практически не различается (Rehder et al., 2009). Более того, по данным этих авторов, абсолютное количество GalNAc DBP у больных раком составляет в среднем 4–5 мг/л, что многократно превышает дозу экзогенно вводимого препарата GcMAF, которую N. Yamamoto использовал в качестве терапевтической (100 нг; 1 раз/нед). Кроме того, авторы не исключают вероятности, что поскольку у больных раком повышен уровень сывороточной гаглозазы, то вводимый препарат GcMAF может сразу инактивироваться вследствие тотального дегликозирования GalNAc. Не исключено также, что во фракции, содержащей GalNAc DBP, может присутствовать фракция белка, обладающая признаками GcMAF, но, по сути, таковой не являющаяся и не активирующая макрофаги. О такой возможности свидетельствовали тесты по связыванию лектина при отсутствии медиаторов воспаления (Kanan et al., 2000).

Для объяснения выявленных противоречий следует принять во внимание следующие факты. По всей видимости, в общей массе GalNAc DBP (GcMAF), определяемой в работе (Rehder et al., 2009), все-таки присутствует белок, который несет GalNAc после селективного дегликозирования β -галактозидазой и сиалидазой на активированных воспалением В- и Т-лимфоцитах. Одновременно это не какой-то аллельный вариант DBP1S, который доминирует в семействе Gc-белков сыворотки крови и поэтому хорошо детектируется масс-спектрометрией. По-видимому, существуют два типа GalNAc DBP, один из которых способен индуцировать макрофаги и является, таким образом, «истинным» GcMAF, а второй имеет аналогичное гликозилирование, но структурно организован иначе и не способен активировать макрофаги. Показано, что среди DBP1S белков, основная масса которых имеет трисахаридную группу в позиции T420, присутствует гликозилированная форма, имеющая карбогидрат в позиции T418. Именно DBP1S белок, гликозилированный в позиции T418, может быть вовлечен в формирование активного GcMAF. Поскольку T420 и T418 в структуре белковой молекулы расположены очень близко, а содержание этой формы GalNAc DBP в сыворотке чрезвычайно мало, его детекция при масс-спектральном пептидном картировании требует больших дополнительных усилий и понимания направления поиска.

Не исключена еще одна возможность. В отличие от других работ D.S. Rehder с коллегами выделяли DBP

белок методом аффинной хроматографии на колонках, нагруженных поликлональными антителами против Gc-белков, а не 25-(ОН) витамином D₃ (Rehder et al., 2009). Возможно, в этом случае выделяется дериват с дефектным витамин D₃-связывающим доменом, который необходим для активности GcMAF. И, следовательно, несмотря на все атрибуты GcMAF, выделенный таким образом белок, который в больших количествах содержится в сыворотке крови как больных раком, так и здоровых людей, не будет «истинным» GcMAF или будет содержать следовые концентрации функционально активного GcMAF.

Кроме того, GcMAF образуется при прямом контакте DBP с активированными Т- и В-клетками, инфильтрирующими опухоль. Здесь же присутствуют потенциальные клетки-мишени – тумор-инфильтрирующие макрофаги. Возможно, что активный GcMAF утилизируется *in situ*, поэтому его содержание в периферической крови как реального макрофаг-активирующего фактора снижено.

И наконец, возможно, что при различного рода выделениях на колонках сорбируется некая иная третичная конформация известного по сайту гликозирования Gc-белка и, таким образом, для получения активного GcMAF нужно использовать систему, где его активность (конкретного белка из конкретной пробирки) валидируется в биологических тестах, по крайней мере, до тех пор, пока не будет четко определена его молекулярная структура.

Заключение

Несмотря на критические замечания относительно позиционирования GcMAF в качестве перспективного противоопухолевого препарата, следует все-таки отметить, что в течение примерно двух десятилетий эффекты GcMAF независимо изучались несколькими исследовательскими группами как в витальных тестах, так и на различных экспериментальных моделях *in vivo*. Имеются также положительные результаты клинических испытаний, которые еще нуждаются в подтверждении на более высоком уровне доказательной медицины. Тем не менее уже сейчас препарат GcMAF как лекарственная форма разрешен для практического применения в Японии. К настоящему времени на базе клиники Saisei Mirai пролечено более 3000 больных различными формами рака с использованием интегративной иммунотерапии, одним из компонентов которой является GcMAF-терапия. В Израиле (компания Efranat Pharma) завершена I фаза клинических испытаний GcMAF (препарат EF-022, с эскалацией дозы от 100 до 1000 нг, в/м, еженедельно) в группе 24 больных с метастатическими формами солидных опухолей. Фармацевтическая компания планирует продолжение исследований и проведение клинических испытаний II–III фаз. В Европе и США также имеется широко не афишируемое лабораторное производство препарата GcMAF (Reno Integrative Medical Center, Nevada; Immuno Biotech Ltd). Кроме того, после некоторого затишья в литературе вновь стали появляться экспериментальные статьи, посвященные исследованию GcMAF. Все это свидетельствует, что фактор интересен в клинической перспективе, поэтому одна из актуальных задач – корректное определение молекулярной структуры белка, обладающего GcMAF активностью.

Список литературы / References

- Сахно Л.В., Шевела Е.Я., Тихонова М.А., Останин А.А., Черных Е.Р. Молекулярные механизмы иммуносупрессорной активности М2-макрофагов. *Иммунология*. 2016;37(6):311-315. [Sakhno L.V., Shevela E.Ya., Tikhonova M.A., Ostanin A.A., Chernykh E.R. Molecular mechanisms of M2 macrophage immunosuppressive activity. *Immunologia = Immunology*. 2016;37(6):311-315. DOI 10.18821/0206-49522016-37-6-311-315. (in Russian)]
- Allavena P., Sica A., Garlanda C., Mantovani A. The Yin-Yang of tumor-associated macrophages in neoplastic progression and immune surveillance. *Immunol. Rev.* 2008;222(1):155-161. DOI 10.1111/j.1600-065X.2008.00607.x.
- Bradstreet J.J., Vogelaar E., Thyer L. Initial observations of elevated alpha-N-acetylgalactosaminidase activity associated with autism and observed reductions from Gcprotein-macrophage activating factor injections. *Autism Insights*. 2012;4:31-38. DOI 10.4137/AUI.S10485.
- Delanghe J.R., Speeckaert R., Speeckaert M.M. Behind the scenes of vitamin D binding protein: more than vitamin D binding. *Best Pract. Res. Clin. Endocrinol. Metab.* 2015;29(5):773-786. DOI 10.1016/j.beem.2015.06.006.
- Dunn G.P., Bruce A.T., Ikeda H., Old L.J., Schreiber R.D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat. Immunol.* 2002;3(11):991-998. DOI 10.1038/ni1102-991.
- Gregory K.J., Zhao B., Bielenberg D.R., Dridi S., Wu J., Jiang W., Huang B., Pirie-Shepherd S., Fannon M. Vitamin D binding protein-macrophage activating factor directly inhibits proliferation, migration, and uPAR expression of prostate cancer cells. *PLoS One*. 2010;5(10):e13428. DOI 10.1371/journal.pone.0013428.
- Haddad J.G., Hu Y.Z., Kowalski M.A., Laramore C., Ray K., Robzyk P., Cooke N.E. Identification of the sterol- and actin-binding domains of plasma vitamin D binding protein (Gc-globulin). *Biochemistry*. 1992;31(31):7174-7181. DOI 10.1021/bi00146a021.
- Homma S., Yamamoto M., Yamamoto N. Vitamin D-binding protein (group-specific component) is the sole serum protein required for macrophage activation after treatment of peritoneal cells with lysophosphatidylcholine. *Immunol. Cell Biol.* 1993;71(Pt. 4):249-257. DOI 10.1038/icb.1993.29.
- Inui T., Kuchiike D., Kubo K., Mette M., Uto Y., Hori H., Sakamoto N. Clinical experience of integrative cancer immunotherapy with GcMAF. *Anticancer Res.* 2013;33(7):2917-2919.
- Ioannou Y.A., Bishop D.F., Desnick R.J. Overexpression of human alpha-galactosidase A results in its intracellular aggregation, crystallization in lysosomes, and selective secretion. *J. Cell Biol.* 1992;119:1137-1150. DOI 10.1083/jcb.119.5.1137.
- Ishikawa M., Inoue T., Inui T., Kuchiike D., Kubo K., Uto Y., Nishikata T. A novel assay system for macrophage-activating factor activity using a human U937 cell line. *Anticancer Res.* 2014;34(8):4577-4581.
- Kanan R.M., Cook D.B., Datta H.K. Lectin immunoassay for macrophage-activating factor (Gc-MAF) produced by deglycosylation of Gc-globulevidence for noninducible generation of Gc-MAF. *Clin. Chem.* 2000;46:412-414.
- Kim R., Emi M., Tanabe K. Cancer immunoediting from immune surveillance to immune escape. *Immunology*. 2007;121(1):1-14. DOI 10.1111/j.1365-2567.2007.02587.x.
- Kisker O., Onizuka S., Becker C.M., Fannon M., Flynn E., D'Amato R., Zetter B., Folkman J., Ray R., Swamy N., Pirie-Shepherd S. Vitamin D binding protein-macrophage activating factor (DBP-maf) inhibits angiogenesis and tumor growth in mice. *Neoplasia*. 2003;5(1):32-40. DOI 10.1016/S1476-5586(03)80015-5.
- Korbelik M., Naraparaju V.R., Yamamoto N. Macrophage-directed immunotherapy as adjuvant to photodynamic therapy of cancer. *Br. J. Cancer*. 1997;75(2):202-207. DOI 10.1038/bjc.1997.34.
- Kuchiike D., Uto Y., Mukai H., Ishiyama N., Abe C., Tanaka D., Kawai T., Kubo K., Mette M., Inui T., Endo Y., Hori H. Degalactosylated/desialylated human serum containing GcMAF induces macrophage phagocytic activity and *in vivo* antitumor activity. *Anticancer Res.* 2013;33(7):2881-2885.
- Malik S., Fu L., Juras D.J., Karmali M., Wong B.Y., Gozdzik A., Cole D.E. Common variants of the vitamin D binding protein gene and adverse health outcomes. *Crit. Rev. Clin. Lab. Sci.* 2013;50(1):1-22. DOI 10.3109/10408363.2012.750262.
- Mantovani A., Sozzani S., Locati M., Allavena P., Sica A. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol.* 2002;23(11):549-555. DOI 10.1016/S1471-4906(02)02302-5.
- Martinez F.O., Gordon S., Locati M., Mantovani A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *J. Immunol.* 2006;177(10):7303-7311. DOI 10.4049/jimmunol.177.10.7303.
- Matsuura T., Uematsu T., Yamaoka M., Furusawa K. Effect of salivary gland adenocarcinoma cell-derived alpha-N-acetylgalactosaminidase on the bioactivity of macrophage activating factor. *Int. J. Oncol.* 2004;24(3):521-528. DOI 10.3892/ijo.24.3.521.
- McCarty F. Overview of macrophage activating factor and the nagalase assay – potential for control of micrometastatic or early primary cancer. 2013. Available at <https://pdfs.semanticscholar.org/8c6d/d28ae1280f52d857145bbd7b14d4a6146e2d.pdf>.
- Mohamad S.B., Nagasawa H., Uto Y., Hori H. Tumor cell alpha-N-acetylgalactosaminidase activity and its involvement in GcMAF-related macrophage activation. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* 2002a;132(1):1-8. DOI 10.1016/S1095-6433(02)00190-3.
- Mohamad S.B., Nagasawa H., Uto Y., Hori H. Preparation of Gc protein-derived macrophage activating factor (GcMAF) and its structural characterization and biological activities. *Anticancer Res.* 2002b;22(6C):4297-4300.
- Morales E.M. GcMAF: a polemic or a highly promising molecule? *World Scientific News*. 2017;65:20-36.
- Nagasawa H., Uto Y., Sasaki H., Okamura N., Murakami A., Kubo S., Kirk K.L., Hori H. Gc protein (vitamin D-binding protein): Gc genotyping and GcMAF precursor activity. *Anticancer Res.* 2005;25:3689-3696.
- Nonaka K., Onizuka S., Ishibashi H., Uto Y., Hori H., Nakayama T., Matsuura N., Kanematsu T., Fujioka H. Vitamin D binding protein-macrophage activating factor inhibits HCC in SCID mice. *J. Surg. Res.* 2012;172(1):116-122. DOI 10.1016/j.jss.2010.07.057.
- Ohm J.E., Carbone D.P. VEGF as a mediator of tumor-associated immunodeficiency. *Immunol. Res.* 2001;23(2-3):263-272. DOI 10.1385/IR.23.2-3:263.
- Otterbein L.R., Cosio C., Graceffa P., Dominguez R. Crystal structures of the vitamin D-binding protein and its complex with actin: structural basis of the actin-scavenger system. *Proc. Natl. Acad. Sci. USA*. 2002;99:8003-8008. DOI 10.1073/pnas.122126299.
- Pacini S., Morucci G., Punzi T., Gulisano M., Ruggiero M. Gc protein-derived macrophage-activating factor (GcMAF) stimulates cAMP formation in human mononuclear cells and inhibits angiogenesis in chick embryo chorioallantoic membrane assay. *Cancer Immunol. Immunother.* 2011;60(4):479-485. DOI 10.1007/s00262-010-0953-7.
- Pacini S., Morucci G., Punzi T., Gulisano M., Ruggiero M., Amato M., Aterini S. Effect of paricalcitol and GcMAF on angiogenesis and human peripheral blood mononuclear cell proliferation and signaling. *J. Nephrol.* 2012a;25(4):577-581. DOI 10.5301/jn.5000035.
- Pacini S., Punzi T., Morucci G., Gulisano M., Ruggiero M. Effects of vitamin D-binding protein-derived macrophage-activating factor on human breast cancer cells. *Anticancer Res.* 2012b;32(1):45-52.
- Rehder D.S., Nelson R.W., Borges C.R. Glycosylation status of vitamin D binding protein in cancer patients. *Protein Sci.* 2009;18(10):2036-2042. DOI 10.1002/pro.214.
- Saburi E., Saburi A., Ghanei M. Promising role for Gc-MAF in cancer immunotherapy: from bench to bedside. *Caspian J. Intern. Med.* 2017;8(4):228-238. DOI 10.22088/cjim.8.4.228.
- Song Y.H., Naumova A.K., Liebhaver S.A., Cooke N.E. Physical and meiotic mapping of the region of human chromosome 4q11-q13

- encompassing the vitamin D binding protein DBP/Gc-globulin and albumin multigene cluster. *Genome Res.* 1999;9(6):581-587.
- Thyer L., Ward E., Smith R., Branca J.J., Morucci G., Gulisano M., Noakes D., Eslinger R., Pacini S. GC protein-derived macrophage-activating factor decreases α -N-acetylgalactosaminidase levels in advanced cancer patients. *Oncoimmunology.* 2013a;2(8):e25769. DOI 10.4161/onci.25769.
- Thyer L., Ward E., Smith R., Fiore M.G., Magherini S., Branca J.J., Morucci G., Gulisano M., Ruggiero M., Pacini S. A novel role for a major component of the vitamin D axis: vitamin D binding protein-derived macrophage activating factor induces human breast cancer cell apoptosis through stimulation of macrophages. *Nutrients.* 2013b;5(7):2577-2589. DOI 10.3390/nu5072577.
- Toutirais O., Chartier P., Dubois D., Bouet F., Leveque J., Catros-Quemener V., Genetet N. Constitutive expression of TGF- β 1, interleukin-6 and interleukin-8 by tumor cells as a major component of immune escape in human ovarian carcinoma. *Eur. Cytokine Netw.* 2003;14(4):246-255.
- Toyohara Y., Hashitani S., Kishimoto H., Noguchi K., Yamamoto N., Urade M. Inhibitory effect of vitamin D-binding protein-derived macrophage activating factor on DMBA-induced hamster cheek pouch carcinogenesis and its derived carcinoma cell line. *Oncol. Lett.* 2011;2(4):685-691. DOI 10.3892/ol.2011.306.
- Ugarte A., Bouche G., Meheus L. Inconsistencies and questionable reliability of the publication "Immunotherapy of metastatic colorectal cancer with vitamin D-binding protein-derived macrophages-activating, GcMAF" by Yamamoto et al. *Cancer Immunol. Immunother.* 2014;63(12):1347-1348. DOI 10.1007/s00262-014-1587-y.
- Uto Y., Hori H., Kubo K., Ichihashi M., Sakamoto N., Mette M., Inui T. GcMAF: our next-generation immunotherapy. *Nature.* 2012;485: S67-S70.
- Verboven C., Rabijns A., De Maeyer M., Van Baelen H., Bouillon R., De Ranter C. A structural basis for the unique binding features of the human vitamin D-binding protein. *Nat. Struct. Biol.* 2002;9(2):131-136. DOI 10.1038/nsb754.
- Yamamoto N., Homma S. Vitamin D3 binding protein (group-specific component) is a precursor for the macrophage-activating signal factor from lysophosphatidylcholine-treated lymphocytes. *Proc. Natl. Acad. Sci. USA.* 1991;88(19):8539-8543. DOI 10.1073/pnas.88.19.8539.
- Yamamoto N., Kumashiro R. Conversion of vitamin D3 binding protein (group-specific component) to a macrophage activating factor by the stepwise action of beta-galactosidase of B cells and sialidase of T cells. *J. Immunol.* 1993;151(5):2794-2802.
- Yamamoto N., Naraparaju V.R., Asbell S.O. Deglycosylation of serum vitamin D3-binding protein leads to immunosuppression in cancer patients. *Cancer Res.* 1996;56(12):2827-2831.
- Yamamoto N., Suyama H., Nakazato H., Yamamoto N., Koga Y. Immunotherapy of metastatic colorectal cancer with vitamin D-binding protein-derived macrophage-activating factor, GcMAF. *Cancer Immunol. Immunother.* 2008a;57:1007-1016. DOI 10.1007/s00262-007-0431-z.
- Yamamoto N., Suyama H., Yamamoto N. Immunotherapy for prostate cancer with Gc protein-derived macrophage-activating factor, GcMAF. *Transl. Oncol.* 2008b;1(2):65-72. DOI 10.1593/tlo.08106.
- Yamamoto N., Suyama H., Yamamoto N., Ushijima N. Immunotherapy of metastatic breast cancer patients with vitamin D-binding protein-derived macrophage activating factor (GcMAF). *Int. J. Cancer.* 2008c;122:461-467. DOI 10.1002/ijc.23107.
- Yamamoto N., Ushijima N., Koga Y. Immunotherapy of HIV-infected patients with Gc protein-derived macrophage activating factor (GcMAF). *J. Med. Virol.* 2009;81:16-26. DOI 10.1002/jmv.21376.

ORCID ID

A.A. Ostanin orcid.org/0000-0001-6895-938X
S.S. Kirikovich orcid.org/0000-0002-3426-4501
E.V. Dolgova orcid.org/0000-0002-5543-248X
A.S. Proskurina orcid.org/0000-0002-7650-4331
E.R. Chernykh orcid.org/0000-0003-2346-6279
S.S. Bogachev orcid.org/0000-0002-2019-9382

Acknowledgements. This work was supported by State Budgeted Project 0324-2019-0042. The authors also acknowledge the financial support provided by I.N. Zaytseva, BA-Farma Company.

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

Received March 27, 2019. Revised May 6, 2019. Accepted May 6, 2019.

Прием статей через электронную редакцию на сайте <http://vavilov.elpub.ru/index.php/jour>
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции»/“Vavilov Journal of Genetics and Breeding”
до 2011 г. выходил под названием «Информационный вестник ВОГиС»/
“The Herald of Vavilov Society for Geneticists and Breeding Scientists”.

Регистрационное свидетельство ПИ № ФС77-45870 выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 20 июля 2011 г.

«Вавиловский журнал генетики и селекции» включен ВАК Минобрнауки России в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, Российский индекс научного цитирования, ВИНТИ, базы данных Emerging Sources Citation Index (Web of Science), Zoological Record (Web of Science), Scopus, Ebsco, DOAJ, Ulrich's Periodicals Directory, Google Scholar, Russian Science Citation Index на платформе Web of Science, каталог научных ресурсов открытого доступа ROAD.

Открытый доступ к полным текстам:
на сайте ИЦиГ СО РАН – bionet.nsc.ru/vogis/
платформе Elpub – vavilov.elpub.ru/index.php/jour
платформе Научной электронной библиотеки – elibrary.ru/title_about.asp?id=32440

Подписку на «Вавиловский журнал генетики и селекции» можно оформить в любом почтовом отделении России. Индекс издания 42153 по каталогу «Пресса России».

При перепечатке материалов ссылка на журнал обязательна.

✉ e-mail: vavilov_journal@bionet.nsc.ru

Издатель: Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»,
проспект Академика Лаврентьева, 10, Новосибирск, 630090.

Адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090.

Секретарь по организационным вопросам С.В. Зубова. Тел.: (383)3634977.

Издание подготовлено информационно-издательским отделом ИЦиГ СО РАН. Тел.: (383)3634963*5218.

Начальник отдела: Т.Ф. Чалкова. Редакторы: В.Д. Ахметова, И.Ю. Ануфриева. Дизайн: А.В. Харкевич.

Компьютерная графика и верстка: Т.Б. Коняхина, О.Н. Савватеева.

.....
Подписано в печать 19.08.2019. Выход в свет 30.08.2019. Формат 60 × 84 1/8. Усл. печ. л. 15.11.

Уч.-изд. л. 17.9. Тираж 150 экз. Заказ № 163. Цена свободная.

.....
Отпечатано в типографии ФГУП «Издательство СО РАН», Морской проспект, 2, Новосибирск, 630090.

