

---

---

# СОДЕРЖАНИЕ

---

---

Том 58, номер 2, 2022

---

---

## Обзорные и теоретические статьи

- Молекулярно-генетические аспекты ответа картофеля на заражение фитофторозом  
*Т. С. Голубева, В. А. Черенко, О. И. Сеницына, А. В. Кочетов* 127
- 

## Молекулярная генетика

- Zvezda* – новое подсемейство *Tc1*-подобных транспозонов в геномах Asterozoa  
*Л. В. Пузакова, М. В. Пузаков* 137
- 

## Генетика микроорганизмов

- Двухступенчатая активация *lux*-регулона психрофильных люминесцирующих бактерий *Aliivibrio logei*  
*С. В. Баженов, Е. С. Шеглова, В. В. Фомин, Г. Б. Завильгельский, И. В. Манухов* 148
- Изменчивость генома эндосимбиотических бактерий *Wolbachia* в инвазивных популяциях азиатского тигрового комара *Aedes albopictus* юга России  
*Е. А. Конов, А. Н. Лукашев, Н. Ю. Оюн* 157
- 

## Генетика животных

- Распространение ассоциированных с болезнью хронического изнурения аллелей гена *PRNP* у диких и домашних северных оленей *Rangifer tarandus* на территории России  
*К. А. Курбаков, Е. А. Конов, М. Т. Семина, Ю. А. Столповский* 163
- Генетические маркеры овец карачаевской породы, выявленные методом полногеномного поиска ассоциаций  
*А. Ю. Криворучко, О. А. Яцык, А. В. Скокова, А. А. Каниболоцкая* 169
- Филогеография *Lacerta media* Lantz et Cyrén, 1920 (Lacertidae: Sauria) по результатам анализа митохондриального гена цитохрома *b*  
*М. А. Доронина, И. В. Доронин, С. А. Луконина, Л. Ф. Мазанаева, А. В. Барабанов* 177
- 

## Генетика человека

- Изучение роли полиморфизма генов репарации ДНК в формировании предрасположенности к развитию рака легкого у женщин  
*Р. А. Титов, В. И. Мина, А. В. Торгунакова, В. Ю. Буслаев, Е. Н. Воронина, А. Ю. Просеков, В. А. Титов, А. Н. Глушков* 187
- Анализ ассоциаций полиморфных локусов гена *LOXL1* с развитием первичной открытоугольной глаукомы у женщин Центрально-Черноземного региона России  
*Н. В. Елисеева, И. В. Пономаренко, М. И. Чурносков* 199
- Молекулярно-генетическая идентификация безымянных захоронений первой половины XVI в. из некрополя Вознесенского собора Московского Кремля  
*И. В. Корниенко, Т. Д. Панова, Т. Г. Фалеева, О. Ю. Арамова, Ю. С. Сидоренко* 206
- 

## Математические модели и методы

- Алгоритм фильтрации наборов данных ДНК-штрихкодирования природных сообществ  
*А. Ю. Краснопеев, Ю. С. Букин, С. А. Потапов, О. И. Бельх* 219
-

## Краткие сообщения

Филогенетическое положение сигов группы *Coregonus lavaretus* (L.) из Телецкого озера (Сибирь) на основе полногеномных данных митохондриальной ДНК

*М. М. Соловьев, Н. А. Бочкарев, Н. В. Орешкова,  
Е. Н. Кашинская, Е. П. Симонов*

232

Полиморфизм NBS-LRR генов устойчивости гибридов сахарной свеклы по данным NBS-профайлинга

*А. А. Трифонова, Е. Р. Парадня, К. В. Борис, А. М. Кудрявцев*

239

---

---

# Contents

---

---

Vol. 58, No. 2, 2022

---

---

## Reviews and Theoretical Articles

- Molecular and Genetic Aspects of Potato Response to Late Blight Infection  
*T. S. Golubeva, V. A. Cherenko, O. I. Sinitsyna, and A. V. Kochetov* 127
- 

## Molecular Genetics

- Zvezda* – A New Subfamily of *Tc I*-Like Transposons in Asterozoa Genomes  
*L. V. Puzakova and M. V. Puzakov* 137
- 

## Genetics of Microorganisms

- Two-Stage Activation of Lux-Regulon Psychrophilic Marine Luminescent Bacteria *Aliivibrio loei*  
*S. V. Bazhenov, E. S. Scheglova, V. V. Fomin, G. B. Zavilgelsky, and I. V. Manukhov* 148
- Genome Variation of Endosymbiotic *Wolbachia* in Introduced Populations of Asian Tiger Mosquito *Aedes albopictus* from Krasnodar Krai  
*E. A. Konorov, A. N. Lukashev, and N. Yu. Oyun* 157
- 

## Animal Genetics

- Associated with Chronic Wasting Disease Alleles of *PRNP* Gene Distribution in Wild and Domesticated Reindeer *Rangifer tarandus* in Russia  
*K. A. Kurbakov, E. A. Konorov, M. T. Semina, and Yu. A. Stolpovsky* 163
- Genetic Markers of Karachaevsky Sheep Identified by Genome-Wide Associations Study  
*A. Yu. Krivoruchko, O. A. Yatsyk, A. V. Skokova, and A. A. Kanibolotskaya* 169
- Cytochrome *b* Mitochondrial Gene Analysis-Based Phylogeography of a *Lacerta media* Lantz et Cyrén, 1920 (Lacertidae: Sauria)  
*M. A. Doronina, I. V. Doronin, S. A. Lukonina, L. F. Mazanaeva, and A. V. Barabanov* 177
- 

## Human Genetics

- Studying the Role of DNA Repair Genes Polymorphism in Formation of Predisposition to Lung Cancer Development in Women  
*R. A. Titov, V. I. Minina, A. V. Torgunakova, V. Yu. Buslaev, E. N. Voronina, A. Yu. Prosekov, V. A. Titov, and A. N. Glushkov* 187
- Analysis of Associations of Polymorphic Loci of the *LOXLI* Gene with the Development of Primary Open-Angle Glaucoma in Women of the Central Chernozem Region of Russia  
*N. V. Eliseeva, I. V. Ponomarenko, and M. I. Churnosov* 199
- DNA Identification of Nameless Burials of the First Half of the 16th Century from the Necropolis of the Ascension Cathedral of the Moscow Kremlin  
*I. V. Kornienko, T. D. Panova, T. G. Faleeva, O. Yu. Aramova, and Yu. S. Sidorenko* 206
- 

## Mathematical Models and Methods

- Data Filtering Algorithm for DNA Metabarcoding Studies of Environmental Communities  
*A. Yu. Krasnopeev, Yu. S. Bukin, S. A. Potapov, and O. I. Belykh* 219
-

## Short Communications

Phylogenetic Position of Whitefish *Coregonus lavaretus* (L.) from Teletskoye Lake (Siberia) Based on Complete Mitochondrial DNA

*M. M. Solovyev, N. A. Bochkarev, N. V. Oreshkova,  
E. N. Kashinskaya, and E. P. Simonov*

232

NBS-LRR Resistance Genes Polymorphism of Sugar Beet Hybrids according to NBS-Profiling Data

*A. A. Trifonova, E. R. Paradnya, K. V. Boris, and A. M. Kudryavtsev*

239

---

ОБЗОРНЫЕ  
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

УДК 632.4.01.08

МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИЕ АСПЕКТЫ ОТВЕТА КАРТОФЕЛЯ  
НА ЗАРАЖЕНИЕ ФИТОФТОРОЗОМ

© 2022 г. Т. С. Голубева<sup>1, 2, \*</sup>, В. А. Черенко<sup>1, 2</sup>, О. И. Синицына<sup>1, 2</sup>, А. В. Кочетов<sup>1, 2</sup>

<sup>1</sup>Федеральный исследовательский центр Институт цитологии и генетики  
Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

<sup>2</sup>Новосибирский государственный университет, Новосибирск, 630090 Россия

\*e-mail: frolova@bionet.nsc.ru

Поступила в редакцию 06.04.2021 г.

После доработки 08.06.2021 г.

Принята к публикации 23.06.2021 г.

Картофель (*Solanum tuberosum*) — одна из наиболее важных сельскохозяйственных культур, основным вредителем которой является фитофтора (*Phytophthora infestans*). Ежегодно производители несут большие потери урожая из-за различных фитофторозов. Данный обзор обобщает исследовательские статьи за достаточно большой период времени и суммирует накопленные экспериментальные данные о взаимодействии картофеля и фитофторы. Структура обзора включает четыре раздела: во вступительной части приводится исторический экскурс в понимание взаимодействия картофеля и фитофторы, здесь приведены самые первые и базовые исследования до 2010 г.; дальнейший период характеризуется увеличением количества исследований по теме обзора, которые сгруппированы в две части: первая содержит сведения о генетических аспектах устойчивости, во второй освещены химические подходы для усиления резистентности; четвертый раздел включает описание перспективных подходов для снижения патогенности фитофторы.

**Ключевые слова:** *Phytophthora infestans*, *Solanum tuberosum*, фитофтороз, устойчивость к патогену, R-ген, вторичные метаболиты, локусы количественной устойчивости.

**DOI:** 10.31857/S0016675822020059

Картофель, или паслен клубненосный (*Solanum tuberosum*), был одомашнен примерно 7000–10000 лет назад на территории современного Южного Перу. В Европу картофель был завезен испанцами относительно недавно, во второй половине XVI в. С тех пор картофель стал одной из основных сельскохозяйственных культур, занимающая четвертое место по валовому сбору выращивания после кукурузы, пшеницы и ржи. В связи с большими объемами производства весьма существенными являются потери урожая из-за различных патогенов, основным из которых является фитофтора (*Phytophthora infestans*), способная полностью уничтожить растения спустя несколько дней после появления первых симптомов заражения (рис. 1). Поэтому исследования механизмов взаимодействия картофеля и фитофторы на молекулярном и генетическом уровнях очень важны с целью дальнейшей разработки новых подходов для повышения резистентности картофеля.

Первые работы, посвященные взаимодействию растений картофеля и *P. infestans*, относятся к концу 1960-х гг. [1–3]. В работах проводилось сравнение восприимчивого (Majestic) и устойчивого (Orion) к фитофторе сортов картофеля. Ранее

было показано, что гены устойчивости (R-гены) экспрессируются только во фрагментах ткани толщиной более 10 клеток [4], поэтому исследование проводилось на клеточных культурах и тканевых агрегатах. В работе было сделано два ключевых вывода: 1) в тканях обоих сортов содержатся вещества, стимулирующие рост фитофторы, т.е. они являются нормальными метаболитами живых тканей, а не образуются в ответ на заражение; 2) тканевые агрегаты сорта Orion тормозили развитие фитофторы, но после заморозки и, как следствие, разрушения тканей это свойство пропадало, таким образом был сделан вывод, что резистентность — это свойство живых тканей, способных реагировать на патоген [1]. В дальнейшем исследовании было установлено, что картофель сорта Orion может быстро развивать постинфекционную токсичность, тормозящую рост зародышевых трубок фитофторы, тем самым претотвращая инфицирование. Ключевая роль в развитии этого ответа отводится R-гену [2].

Соединения, обуславливающие устойчивость растений к патогенам, ранее были названы фитоалексинами [5], высказывалось предположение, что именно они нарабатываются в тканях в ответ на инфицирование, что и приводит к инги-



Рис. 1. Лист картофеля с ранними симптомами фитофтороза.

бированию патогенов. Позднее была опубликована работа, свидетельствующая о накоплении в токсичных для фитофторы фракциях различных фенольных соединений [3]: салициловой, *n*-гидроксibenзойной и ванилиновой кислот. В работе высказывается предположение, что система устойчивости, основанная на *R*-гене, может быть связана с постинфекционной индукцией синтеза фитоалексинов. Примерно в это же время и значительно позже были описаны фитоалексины сесквитерпеноидной структуры: ришитин, любимин, фитуберин и солаветивон, выделенные из зараженных клубней картофеля [6–8].

В 1987 г. вышло подробное исследование временного накопления ришитина как наиболее показательного маркера иммунного ответа в клубнях резистентных и восприимчивых сортов в ответ на заражение *P. infestans* [9]: взаимодействия картофеля и *P. infestans*, при которых развивался фитофтороз, являются совместимыми, в случае же отсутствия каких-либо признаков патогенеза взаимодействие считается несовместимым. Обнаружено, что ришитин и некоторые его структурно родственные производные (любимин) быстро накапливались в клубнях при несовместимых взаимодействиях картофеля и фитофторы и достаточно медленно у совместимых, в листьях же подобный ответ отсутствовал. Подводится итог, что сесквитерпены могут быть полезны в развитии иммун-

ного ответа, но не являются обязательными компонентами устойчивости.

Интересная работа была проведена по устойчивости картофеля к *P. infestans* через получение соматических и половых гибридов культивируемого *S. tuberosum* с диким подвидом *S. circaeifolium* Bitter, устойчивость которого к фитофторозу является весьма привлекательным признаком для включения в генофонд культурного картофеля. Путем слияния клеток были получены тетраплоидные гибридные каллусы [10], выросшие из которых растения обладали полной устойчивостью к *P. infestans*. Исследователи сравнили содержание гликоалкалоидов у родительских растений и гибридов и обнаружили повышенное содержание гликозида томатидина, позаимствованное гибридом от дикого *S. circaeifolium*. Кроме того, был выявлен новый гликозид демиссидин, не обнаруженный ни у одного из родителей. Отмечается фертильность полученных женских растений и возможность их скрещивания с культивируемым *S. tuberosum* для закрепления приобретенной устойчивости. Группой исследователей были получены половые гибриды *S. tuberosum* и *S. circaeifolium* [11]. Среди гибридов наблюдались преимущественно триплоидные растения (83%), несущие двойной геном *S. circaeifolium*, что делает их более устойчивыми к патогену и, как следствие, более перспективными для внедрения в сельское хозяйство.

Половые гибриды также обладали повышенным содержанием гликозида томатыдина и появлением гликозида демиссидина, не обнаруженного у родительских растений.

### ГЕНЕТИЧЕСКИЕ АСПЕКТЫ УСТОЙЧИВОСТИ *S. tuberosum* К ФИТОФТОРОЗУ

В начале XXI в. развитие компьютерных технологий и статистических методов позволило проводить более масштабный анализ групп сцепления, выявлять локусы с генами количественных признаков и строить для них карты сцепления. Были получены карты сцепления для диплоидного картофеля, согласно которым локусы количественной устойчивости к фитофторе располагаются практически в каждой хромосоме, а на хромосомах III, IV, V и VI были сцеплены с поздней зрелостью [12, 13]. Позднее подобные результаты были получены и для тетраплоидного картофеля [14] при скрещивании устойчивого сорта (Stirling) и восприимчивого. Среди множества групп сцепления обнаружен перспективный локус в IV хромосоме, который не сцеплен с поздней зрелостью клубней. У устойчивого родительского растения *R*-ген, переданный потомству, был картирован в XI хромосоме.

Несмотря на обилие работ в это время, посвященных картированию *R*-генов в геноме картофеля [15–24], стоит отметить, что селекция устойчивых к фитофторе сортов картофеля на основании *R*-генов не особо выгодна с экономической точки зрения: ранее было отмечено, что уже через 5–10 лет сорт будет снова восприимчивым к новым расам *P. infestans* [25]. Распознавание патогена *R*-геном довольно легко обходится мутациями в соответствующем гене авирулентности (*Avr*) *P. infestans*, что позволяет патогену успешно проникать и колонизировать растение-хозяина при совместимом взаимодействии. В зависимости от генотипа растения-хозяина инвазия, рост и споруляция *P. infestans* прогрессируют с переменной эффективностью и скоростью. Эта естественная вариация совместимого взаимодействия представляет собой другой тип устойчивости, который контролируется множеством генетических факторов и факторов окружающей среды и называется количественной устойчивостью [26].

Биоинформатические методы были использованы не только для исследования устойчивости картофеля к фитофторозу, но и для изучения механизмов восприимчивости. Было показано, что карбоангидраза – фермент, обратимо конвертирующий диоксид углерода в бикарбонат, может играть большую роль во время несовместимых взаимодействий между патогеном и хозяином [21]. С помощью ДНК-микрочипов было исследовано временное изменение экспрессии генов: в первые часы

после заражения (6–12 ч) наблюдалась индукция экспрессии, но спустя некоторое время (48–72 ч) было репрессировано больше генов, чем индуцировано. Интересным выглядит подавление жасмонатного пути при восприимчивых взаимодействиях. Однако больше всего репрессированных после заражения генов связано с фотосинтезом: так, наиболее выраженное подавление обнаружено для пластидной карбоангидразы, которая обладает антиоксидантной активностью и способностью связывать салициловую кислоту [27]: в первые 12 ч наблюдалось существенное усиление экспрессии карбоангидразы у несовместимых взаимодействий, тогда как через 24 или 48 ч ее следы едва обнаруживались, что позволяло однозначно отличить устойчивые формы. Существенная роль салициловой кислоты именно в раннем ответе *S. tuberosum* на инфицирование была продемонстрирована с помощью трансгенных NahG растений, которые не способны накапливать салициловую кислоту, их восприимчивость к *P. infestans* была существенно выше, чем у дикой формы. Однако предварительная обработка препаратом салициловой кислоты практически уравнивала вероятность заражения [28].

Помимо единичных доминантных *R*-генов устойчивости, отвечающих за распознавание соответствующего гена авирулентности (*Avr*) *P. infestans* и запускающих защитный ответ, проявляющийся в локальной гибели клеток и тем самым останавливающий рост патогенных микроорганизмов, у растений существует группа генов с другим механизмом защиты – гены множественной устойчивости. Была исследована экспрессия четырех генов-транспортеров у картофеля, отвечающих за множественную устойчивость к препаратам: *StPDR1*, *StPDR2*, *StPDR3* и *StPDR4*, транскрипция которых регулировалась различными препаратами [29]. Среди них были выявлены и те, экспрессия которых существенно возрастает при заражении *P. infestans*, гены *StPDR1* и *StPDR2* экспрессировались активнее в 13 и 37 раз соответственно спустя 18 ч после заражения. Авторы полагают, что все исследованные ими гены (*StPDR1–4*) являются частью более сложного системного ответа растения на биотические и абиотические факторы.

В 2012 г. впервые был выполнен транскриптомный анализ для совместимых и несовместимых взаимодействий *P. infestans* и *S. tuberosum* [30]. Исследовалось пять изогенных линий, которые отличались только наличием и отсутствием *R*-гена (*R1*), чтобы исключить влияние генетического фона на транскриптом. Однако на деле оказалось, что изменения транскриптома между растениями одной группы различаются сильнее, чем средние транскриптомы групп между собой, что указывает на большое влияние как индивидуальных физиологических факторов, так и факторов окружающей среды на время защитной реакции. Одним

из них может быть движение во времени и пространстве защитных сигналов от места первоначального физического контакта между зооспорами *P. infestans* и клетками-хозяевами и соседними клетками. Анализ DeepSAGE позволил выявить некоторые интересные аспекты общей структуры транскриптома. Две трети (68%) унигагов были экспрессированы на низких уровнях в ткани листьев, из которых одна треть (36%) не соответствовала ни одному из известных транскриптов картофеля, что может быть объяснено немодельным объектом и, как следствие, немногочисленными данными о транскриптом *S. tuberosum*. С другой стороны, только 3% унигагов показали среднюю экспрессию, но составили 32% от общего транскриптома. Из очень частых транскриптов только 14% были неизвестны. Стоит отметить следующие транскрипты из этого исследования:

1. Наиболее частый маркер (*StET008016*) соответствует уникальному транскрипту *TC208859*, аннотированному как белок клеточной стенки, экспрессия этого гена составляла 4% всех транскриптов в тканях листьев. Белки этого семейства выполняют функцию каркаса в качестве агглютинирующих агентов для отложения компонентов клеточной стенки [31]. Также обнаружено их участие в защитной реакции растений против бактерий и грибов [32, 33].

2. Через день после инокуляции транскрипт *StET009643*, соответствующий гену трансальдозазы ToTAL2 (*TC196885*), специфически и временно повышался во время несовместимых взаимодействий. Трансальдозазы (EC 2.2.1.2) катализируют образование одного из предшественников шикимовой кислоты, который участвует в образовании фенолпропаноидов, алкалоидов и растительных гормонов ауксина и салициловой кислоты – важного компонента системы защиты растения.

3. Через три дня после инфицирования экспрессия транскрипта *StET010841*, соответствующего гену фибриллина 8 (FIB8) (*TC207935*), была строго и специфически подавлена. Растительные фибриллины представляют собой структурные липид-ассоциированные белки, расположенные в тилакоидных мембранах, которые, по-видимому, играют роль в реакциях биотического и абиотического стресса, роста и развития и в гормональной сигнализации.

4. Снижается уровень транскриптов карбоангидразы, что может быть использовано для повышения восприимчивости к *P. infestans* (уникальные транскрипты *TC209461*, *TC218724* или *TC221870* аннотированы как карбоангидразы).

Авторами настоящей статьи отмечается, что изменения транскрипции во время инфекции были гораздо более выраженными в совместимых, чем в несовместимых, взаимодействиях. Число последовательностей-мишеней, однако, было одинаково-

вым: 240 и 220 для несовместимых и совместимых взаимодействий соответственно. В целом транскриптомы несовместимых и совместимых взаимодействий показали больше различий, чем общности. Несовместимое взаимодействие приводит к запрограммированной гибели клеток небольшого числа клеток в месте первичного контакта с патогеном. Транскриптом этих клеток может претерпеть существенные изменения, тогда как отобранные соседние ткани остаются относительно нетронутыми. Наблюдаемые изменения могут быть причинами или следствиями сигналов системной приобретенной резистентности.

Итак, через некоторое время интерес к *R*-генам как перспективным для селекции кандидатам существенно уменьшился – устойчивость за счет них достаточно быстро сводилась на нет новыми расами фитофторы. Вместо поиска основного гена исследователи перешли к мнению, что резистентность более перспективно рассматривать как полигенный и, как следствие, количественный признак. Поэтому вектор исследований изменился на поиск новых генов-кандидатов и их картирования для эффективного сочетания аллелей повышенной резистентности в улучшенных сортах. Постулируется, что для преодоления полигенной устойчивости *P. infestans* вынуждена делать больше мутаций.

С помощью SNP-маркеров были выявлены как наиболее перспективные гены липоксигеназы (жасмонатный путь), 3-гидрокси-3-метил-глутарил-коэнзим А редуктаза (мевалонатный путь) и цитохром р450 (терпеновый биосинтез) [34].

С учетом предыдущих исследований [35–37] суммарно выявлено десять наиболее подходящих локусов для применения в качестве диагностических маркеров в селекционных программах. Эти гены кодируют ферменты, функционирующие в жасмонатном и оксипиноновом пути (*StAOS2*, *Plox1*), в биосинтезе липидов (*BCCP*, биотин-карбоксильный белок-носитель) и вторичных терпеновых метаболитов (*HMGCR*, *CYP71D11*). Есть гены с неизвестной функцией (*StGP28*) или с функцией распознавания патогенов (*Rpi-vnt1*) и транскрипционной регуляции (*TEF1*, *C3HL-TF*, *RBP50*). Их значение как генов-кандидатов важно для непосредственного участия в контроле количественной устойчивости к фитофторозу, которая не нарушается поздней зрелостью растений, для дальнейшей функциональной характеристики и для подтверждения диагностической способности в различных селекционных популяциях и средах.

Изопрен синтезируется в растениях двумя путями: ацетат/мевалонатным (ацетат/MVA) и дезоксилилозафосфат/метилэритритфосфатным. Ген-участник резистентности 1-дезоксид-ксилозула-5-фосфатсинтазы-1 (*StDXS1*) был обнару-

жен во втором пути [38]. Его экспрессия изменяется в ответ на заражение Р инф и коррелирует с накоплением 1-дезоксид-*D*-ксилоза-5-фосфатсинтазы-1 – фермента, катализирующего начальную стадию пластидиального 2С-метил-*D*-эритрит-4-фосфатного пути (DOXP-MEP), участвующего в биосинтезе изопреноидов, необходимых для оболочек хлоропластов. Изопреноиды также необходимы для синтеза каротиноидов и хлорофилла, абсцизовой и гиббереллиновой кислот, содержание которых соответственно увеличивается при увеличении экспрессии *DXS*.

Для изучения процесса взаимодействия *P. infestans* и *S. tuberosum* был выполнен протеомный анализ для SUMOилирования – процесса, который обеспечивают небольшие убиквитин-подобные модификаторы. Инвазивные растительные патогены развили способность модифицировать метаболизм своего хозяина, стимулируя метаболические процессы, которые способствуют росту патогена за счет общего хозяина, тем самым они вполне могут влиять не метаболические процессы в растительных клетках [39]. Действительно, было обнаружено, что во время процесса заражения содержание большинства известных конъюгатов SUMO *S. tuberosum* значительно изменяется, некоторые уменьшаются, но многие существенно увеличиваются. Выявлены белки-маркеры ответа на заражение со стороны картофеля – это патоген-связанный белок (StPR1) и пластидная карбоангидраза (CA). Со стороны *P. infestans* маркерами заражения были грибные целлюлозо-синтазы (CesA3) [40] и гаусторий-специфический мембранный белок (PiHmp1) [41]. Синтез StPR1 стимулируется салициловой кислотой, способность поддерживать его в высокой концентрации после заражения отличает резистентные растения [42]. Роль CA остается все еще неясной до конца, однако разница в уровне CA также позволяет различить устойчивые и неустойчивые к фитофторе сорта [21]. Гены белковых маркеров *P. infestans* сперва активно экспрессировались как в устойчивых, так и в неустойчивых сортах, но при несовместимом взаимодействии их активность начинала снижаться через 24 ч после заражения. Таким образом, были получены доказательства, что в восприимчивых сортах картофеля патогену удается ингибировать защитные механизмы растения и успешно инфицировать растения, в то время как в устойчивых сортах такой процесс не наблюдается.

Помимо генов семейства *PR1*, в ответ на грибное заражение также активируется ряд других семейств *PR*-генов [43]. Например, глюканазы ( $\beta$ -1,3-глюканаза картофеля), относящиеся к семейству *PR2*, а также хитиназы (*PR3*) участвуют в расщеплении полисахаридов грибной клеточной стенки и играют большую роль в иммунном ответе растений [43–45]. Гены семейства *PR5* кодируют тауматин-

подобные белки, обладающие противогрибковой активностью, и также активируются во время колонизации фитофторой [43, 45, 46].

Еще одна временная динамика при заражении была получена для гена *StPOTHRI* с использованием трансгенных растений, где целевой ген был включен за счет РНК-интерференции [47]. Продукт гена *StPOTHRI* локализуется на плазматической мембране клеток и существенно уменьшает степень колонизации, причем его сверхэкспрессия усиливает резистентность, а экспрессироваться этот ген начинает только после заражения в устойчивых сортах.

Масштабная попытка картирования локусов количественных признаков (QTL) у картофеля была вновь предпринята в 2018 г. Santa с соавт. [48]. Исследователи снова выбрали в качестве объекта тетраплоидный геном картофеля, отмечая его высокую важность для селекции и при этом существенные затруднения анализа в силу высокой гетерозиготности у автотетраплоидов. Авторам удалось обнаружить два новых QTL на III и VIII хромосомах. Отмечается, что один из аллелей первого локуса может опосредовать в среднем более высокую степень тяжести заболевания. Также этот локус включает транскрипционный фактор *Arf2*, связанный со старением листьев вследствие окислительного стресса, а также с передачей сигналов гиббереллина и брассиностероидных путей при взаимодействии растение–патоген [49–51]. QTL VIII хромосомы содержал аллель, отвечающий в среднем за более низкую степень тяжести заболевания. Этот маркер связан с геном, который кодирует фактор транскрипции спираль–петля–спираль (bHLH) *JAF13*, участвующий в биосинтезе флавоноидов у *Petunia* × *hybrida* [52].

Были получены трансгенные растения со сверхэкспрессией гена *D*-галактуронатредуктазы (*GalUR*) и, как следствие, с повышенным уровнем *L*-аскорбата [53]. После заражения размер некротических пятен у трансгенных растений был меньше, чем в контрольной группе, при этом отмечено увеличение экспрессии генов, участвующих в антиоксидантной активности. В результате отмечено снижение активных форм кислорода в клетках, что может объяснять меньший размер некротического пятна за счет уменьшения ответа гиперчувствительности, индуцируемого активными формами кислорода. В целом отмечалась большая устойчивость трансгенных растений к фитофторе, однако полной устойчивости не наблюдалось. *L*-аскорбат снизил содержание абсцизовой кислоты и увеличил гиббереллиновой, однако при этом отмечена существенная потеря в урожайности картофеля (ГМО пенальти), которая все же превышала урожай от зараженного.

## ЗАЩИТА *S. tuberosum* ОТ ФИТОФТОРОЗА С ПОМОЩЬЮ ХИМИЧЕСКИХ ПРЕПАРАТОВ

Были предприняты попытки усилить устойчивость картофеля к фитофторе за счет обработки химическими препаратами. Например, олигосахарид курдлан может служить активатором врожденного иммунитета у растений, что делает его безопасным и перспективным препаратом для сельского хозяйства [54]. Курдлан является элиситором, оказывающим активное влияние на реакции ранней и поздней защиты за счет индукции синтеза  $H_2O_2$  и салициловой кислоты. Как и другие защитные механизмы, вызванные экзогенными веществами, данные проявления со временем исчезали.

Несмотря на классификацию *P. infestans* как оомицета, она имеет черты схожести и с апикомплексами (*Apicomplexa*), соответственно препараты от малярии, к примеру, могут быть эффективными для подавления роста фитофторы. Одной из перспективных мишеней у споровиков является дегидрооротатдегидрогеназа (*DHODH*), которая входит в число 72 ключевых генов, необходимых для ее роста [55].

Исследователями были получены препараты ферментов *DHODH* из *P. infestans* и *S. tuberosum*, чувствительность которых была оценена к нескольким препаратам. В результате из 17 потенциальных ингибиторов был выбран один (терифлюномид), ингибирующий *DHODH* у *P. infestans*, но практически не влияющий на *S. tuberosum* [56].

Увенчались успехом эксперименты по обработке картофеля фосфитом [57]. Уже через 3 ч были обнаружены изменения в транскриптом и секретоме картофеля, которые держались менее 24 ч, но защитный эффект длился гораздо дольше. Транскрипты, связанные с защитой, ранением и окислительным стрессом, составили ядро реакции после обработки фосфитом. Также наблюдались изменения в первичном метаболизме и процессах, связанных с клеточной стенкой. Показано, что эти изменения не связаны с истощением уровня фосфора или закислением среды, вызванным обработкой фосфитом. Из всего множества транскриптов, регулируемых фосфитом, экспрессия 40% также изменялась с использованием  $\beta$ -аминомасляной кислоты (ВАВА) в качестве элиситора. Стоит отметить, что работа защитного гена *StPRI* была повышена только с помощью ВАВА, обработка фосфитом не изменяла его экспрессии. Также исследование линий трансгенного картофеля показало, что резистентность к фитофторозу, возникшая вследствие обработки фосфитом, не зависит от уровня салициловой и жасмоновой кислот.

## *P. infestans* КАК ОБЪЕКТ ИССЛЕДОВАНИЯ ПРИ ЗАРАЖЕНИИ КАРТОФЕЛЯ

Фитофтора также может быть объектом для воздействий при необходимости защитить картофель, хотя и гораздо более сложным в силу ряда

причин: скрытный образ жизни, который помогает *P. infestans* избежать систем защиты растений; эффекторы, которые подавляют защиту хозяина и способствуют восприимчивости; обильная споруляция с коротким латентным периодом, обеспечивающим быстрое распространение; а также структура генома, которая способствует адаптивной эволюции *P. infestans* путем стимулирования генетического разнообразия [58]. Этот патоген оказался весьма искусным в преодолении стратегий контроля, включая устойчивость хозяина и фунгициды.

Так, ооспоры фитофторы много лет могут сохраняться в почве, выжидая совместимого взаимодействия с картофелем или подходящих климатических условий. Наличие полового цикла в развитии позволяет генерировать рекомбинантные генотипы, устойчивые к химическим веществам или обладающие более патогенными свойствами [59]. Также выявлены закономерности чередования половой и бесполой форм у фитофторы в зависимости от колонизируемых сортов картофеля, что дает преимущества для адаптации патогена к хозяину. Установлено, что патоген переходит преимущественно на половое размножение на более восприимчивых и быстро погибающих сортах [60].

Самые масштабные исследования были проведены с 2011 по 2015 гг. в разных штатах Америки по распределению заболеваемости картофеля клональными линиями фитофторы. Исследования были комплексными, учитывали не только сезонные условия, но и сельскохозяйственные мероприятия и сроки их проведения. Подробно описаны морфологические и генетические признаки линий [61–66]. Данный цикл работ внес огромный вклад в эпидемиологическое исследование распространения возбудителя.

О генетических особенностях *P. infestans* известно очень мало, в основном из-за недостатка стабильных генетических маркеров. До 1980-х гг. в качестве маркеров были доступны только специфические (вирулентность, тип спаривания), однако сейчас установлено, что даже специфическая вирулентность является очень варибельным признаком [67].

## ЗАКЛЮЧЕНИЕ

Оомицет *P. infestans* является основным патогеном сельскохозяйственных культур семейства пасленовые. Учитывая, что картофель *S. tuberosum* является четвертой культурой в мире по масштабам выращивания, потери от фитофтороза огромны. Взаимодействие хозяина и паразита можно рассматривать с разных точек зрения: можно исследовать генетические аспекты картофеля в ответ на заражение, а можно — аспекты фитофторы при колонизации. Также можно рассматривать этот

процесс с химической точки зрения, исследуя вторичные метаболиты хозяина и патогена, а также разработку новых препаратов для защиты картофеля.

В данный момент генетических исследований защитных реакций со стороны картофеля существенно больше и основная их часть направлена на поиск и картирование локусов количественных признаков, отвечающих за резистентность. На заре исследований наибольший интерес вызывали *R*-гены, одного аллеля которых зачастую было достаточно для устойчивости к фитофторозу. Таким образом, резистентность рассматривалась преимущественно как качественный признак. Позже выяснилось, что созданные сорта невыгодны с экономической точки зрения — фитофтора легко преодолевает моногенную устойчивость за счет особенностей строения генома и, как следствие, быстрого возникновения мутаций. Это обстоятельство вынудило пересмотреть устоявшиеся взгляды и перейти к оценке резистентности как к количественному признаку, поэтому встал вопрос о поиске локусов количественных признаков, которые могут быть перспективными для селекции. В процессе исследования выяснилось, что у культурного картофеля основные локусы, дающие резистентность, зачастую сцеплены с негативными для сельского хозяйства качествами, например поздней зрелостью клубней. Поэтому поиск все новых и новых локусов по-прежнему остается очень актуальной задачей.

Достигнуты определенные успехи в поиске химических препаратов и вторичных метаболитов, стимулирующих защитные свойства картофеля и снижающих патогенность фитофторы. Как правило, при исследовании процесса заражения с химической точки зрения практически не рассматриваются изменения метаболизма, а в имеющихся работах сообщается о весьма коротком периоде ответа растения на препарат.

Несмотря на большой прогресс в понимании механизмов устойчивости и способности предсказывать вспышки фитофтороза, пандемические вспышки с колоссальным уроном все еще случаются в разных странах [68, 69], что свидетельствует о недостаточности полученных знаний для эффективной защиты сельскохозяйственных растений и необходимости новых исследований в этом направлении.

Исследование выполнено при поддержке Российского научного фонда (проект № 19-74-00067).

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Ingram D.S., Robertson N.F. Interaction between *Phytophthora infestans* and tissue cultures of *Solanum tuberosum* // J. Gen. Microbiol. 1965. V. 40. P. 431–437. <https://doi.org/10.1099/00221287-40-3-431>
2. Ingram D.S. The expression of R-gene resistance to *Phytophthora infestans* in tissue cultures of *Solanum tuberosum* // J. Gen. Microbiol. 1967. V. 49. P. 99–108. <https://doi.org/10.1099/00221287-49-1-99>
3. Robertson N.F., Friend J., Aveyard M. et al. The accumulation of phenolic acids in tissue culture pathogen combinations of *Solanum tuberosum* and *Phytophthora infestans* // J. Gen. Microbiol. 1968. V. 54. P. 261–268. <https://doi.org/10.1099/00221287-54-2-261>
4. Tomiyama K., Takakuwa M., Takase N. The metabolic activity in healthy tissue neighbouring the infected cells in relation to resistance to *Phytophthora infestans* in potatoes // Phytopathology. 1958. P. 237–250.
5. Müller K.O., Behr L. “Mechanism” of *Phytophthora*-resistance of potatoes // Nature. 1949. V. 163. P. 498–499. <https://doi.org/10.1038/163498a0>
6. Katsui N., Murai A., Takasugi M. et al. The structure of rishitin, a new antifungal compound from diseased potato tubers // Chem. Commun. 1968. P. 43–44. <https://doi.org/10.1039/C19680000043>
7. Kuc J. Phytoalexins from the Solanaceae // Phytoalexins. London: Blackie, 1982. P. 81–105.
8. Kuc J., Rush J.S. Phytoalexins // Arch. Biochem. Biophys. 1985. V. 236. P. 455–472.
9. Rohwer F., Fritzemeier K.H., Scheel D., Hahlbrock K. Biochemical reactions of different tissues of potato (*Solanum tuberosum*) to zoospores or elicitors from *Phytophthora infestans* — accumulation of sesquiterpenoid phytoalexins // Planta. 1987. V. 170. P. 556–561. <https://doi.org/10.1007/BF00402991>
10. Mattheij W.M., Eijlander R., de Koning J.R.A., Louwes K.M. Interspecific hybridization between the cultivated potato *Solanum tuberosum* subspecies *tuberosum* L. and the wild species *S. circaeifolium* subsp. *circaeifolium* Bitter exhibiting resistance to *Phytophthora infestans* (Mont.) de Bary and *Globodera pallida* (Stone) Behrens — 1. Somatic hybrids // Theor. Appl. Genet. 1992. V. 83. P. 459–466. <https://doi.org/10.1007/BF00226534>
11. Louwes K.M., Hoekstra R., Mattheij W.M. Interspecific hybridization between the cultivated potato *Solanum tuberosum* subspecies *tuberosum* L. and the wild species *S. circaeifolium* subsp. *circaeifolium* Bitter exhibiting resistance to *Phytophthora infestans* (Mont.) de Bary and *Globodera pallida* (Stone) Behrens — 2. Sexual hybrids // Theor. Appl. Genet. 1992. V. 84. P. 362–370. <https://doi.org/10.1007/BF00229495>
12. Simko I. Comparative analysis of quantitative trait loci for foliage resistance to *Phytophthora infestans* in tuber-bearing *Solanum* species // Am. J. Potato Res. 2002. V. 79. P. 125–132. <https://doi.org/10.1007/BF02881521>
13. Gebhardt C., Valkonen J.P.T. Organization of genes controlling disease resistance in the potato genome // Annu. Rev. Phytopathol. 2001. V. 39. P. 79–102. <https://doi.org/10.1146/annurev.phyto.39.1.79>

14. Bradshaw J.E., Pande B., Bryan G.J. et al. Interval mapping of quantitative trait loci for resistance to late blight [*Phytophthora infestans* (Mont.) de Bary], height and maturity in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*) // Genetics. 2004. V. 168. P. 983–995.  
<https://doi.org/10.1534/genetics.104.030056>
15. Bradshaw J.E., Bryan G.J., Lees A.K. et al. Mapping the R10 and R11 genes for resistance to late blight (*Phytophthora infestans*) present in the potato (*Solanum tuberosum*) R-gene differentials of Black // Theor. Appl. Genet. 2006. V. 112. P. 744–751.  
<https://doi.org/10.1007/s00122-005-0179-9>
16. Bradshaw J.E., Hackett C.A., Lowe R. et al. Detection of a quantitative trait locus for both foliage and tuber resistance to late blight [*Phytophthora infestans* (Mont.) de Bary] on chromosome 4 of a dihaploid potato clone (*Solanum tuberosum* subsp. *tuberosum*) // Theor. Appl. Genet. 2006. V. 113. P. 943–951.  
<https://doi.org/10.1007/s00122-006-0353-8>
17. Solomon-Blackburn R.M., Stewart H.E., Bradshaw J.E. Distinguishing major-gene from field resistance to late blight (*Phytophthora infestans*) of potato (*Solanum tuberosum*) and selecting for high levels of field resistance // Theor. Appl. Genet. 2007. V. 115. P. 141–149.  
<https://doi.org/10.1007/s00122-007-0550-0>
18. Ballvora A., Ercolano M.R., Weiß J. et al. The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes // Plant J. 2002. V. 30. P. 361–371.  
<https://doi.org/10.1046/j.1365-313X.2001.01292.x>
19. Van Der Vossen E., Sikkema A., Te Lintel Hekkert B. et al. An ancient R-gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato // Plant J. 2003. V. 36. P. 867–882.  
<https://doi.org/10.1046/j.1365-313X.2003.01934.x>
20. Park T.H., Vleeshouwers V.G.A.A., Huigen D.J. et al. Characterization and high-resolution mapping of a late blight resistance locus similar to R2 in potato // Theor. Appl. Genet. 2005. V. 111. P. 591–597.  
<https://doi.org/10.1007/s00122-005-2050-4>
21. Restrepo S., Myers K.L., Del Pozo O. et al. Gene profiling of a compatible interaction between *Phytophthora infestans* and *Solanum tuberosum* suggests a role for carbonic anhydrase // Mol. Plant-Microbe Interact. 2005. V. 18. P. 913–922.  
<https://doi.org/10.1094/MPMI-18-0913>
22. Tan M.Y.A., Hutten R.C.B., Celis C. et al. The RPi-mcd1 locus from *Solanum microdontum* involved in resistance to *Phytophthora infestans*, causing a delay in infection, maps on potato chromosome 4 in a cluster of NBS-LRR genes // Mol. Plant-Microbe Interact. 2008. V. 21. P. 909–918.  
<https://doi.org/10.1094/MPMI-21-7-0909>
23. Brugmans B., Wouters D., Van Os H. et al. Genetic mapping and transcription analyses of resistance gene loci in potato using NBS profiling // Theor. Appl. Genet. 2008. V. 117. P. 1379–1388.  
<https://doi.org/10.1007/s00122-008-0871-7>
24. Rauscher G., Simko I., Mayton H. et al. Quantitative resistance to late blight from *Solanum berthaultii* cosegregates with RPi-ber: Insights in stability through isolates and environment // Theor. Appl. Genet. 2010. V. 121. P. 1553–1567.  
<https://doi.org/10.1007/s00122-010-1410-x>
25. Stewart H.E., Bradshaw J.E., Pande B. The effect of the presence of R-genes for resistance to late blight (*Phytophthora infestans*) of potato (*Solanum tuberosum*) on the underlying level of field resistance // Plant Pathol. 2003. V. 52. P. 193–198.  
<https://doi.org/10.1046/j.1365-3059.2003.00811.x>
26. Poland J.A., Balint-Kurti P.J., Wisser R.J. et al. Shades of gray: The world of quantitative disease resistance // Trends Plant Sci. 2009. V. 14. P. 21–29.
27. Slaymaker D.H., Navarre D.A., Clark D. et al. The tobacco salicylic acid-binding protein 3 (SABP3) is the chloroplast carbonic anhydrase, which exhibits antioxidant activity and plays a role in the hypersensitive defense response // Proc. Natl Acad. Sci. USA. 2002. V. 99. P. 11640–11645.  
<https://doi.org/10.1073/pnas.182427699>
28. Halim V.A., Eschen-Lippold L., Altmann S. et al. Salicylic acid is important for basal defense of *Solanum tuberosum* against *Phytophthora infestans* // Mol. Plant-Microbe Interact. 2007. V. 20. P. 1346–1352.  
<https://doi.org/10.1094/MPMI-20-11-1346>
29. Ruocco M., Ambrosino P., Lanzuise S. et al. Four potato (*Solanum tuberosum*) ABCG transporters and their expression in response to abiotic factors and *Phytophthora infestans* infection // J. Plant Physiol. 2011. V. 168. P. 2225–2233.  
<https://doi.org/10.1016/j.jplph.2011.07.008>
30. Gyetvai G., Sønderkær M., Göbel U. et al. The transcriptome of compatible and incompatible interactions of potato (*Solanum tuberosum*) with *Phytophthora infestans* revealed by DeepSAGE analysis // PLoS One. 2012. V. 7. e31526.  
<https://doi.org/10.1371/journal.pone.0031526>
31. Mangeon A., Junqueira R.M., Sabetto-Martins G. Functional diversity of the plant glycine-rich proteins superfamily // Plant Signal. Behav. 2010. V. 5. P. 99–104.  
<https://doi.org/10.4161/psb.5.2.10336>
32. Fu Z.Q., Guo M., Jeong B.R. et al. A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity // Nature. 2007. V. 447. P. 284–288.  
<https://doi.org/10.1038/nature05737>
33. Park C.J., Park C.B., Hong S.S. et al. Characterization and cDNA cloning of two glycine- and histidine-rich antimicrobial peptides from the roots of shepherd's purse, *Capsella bursa-pastoris* // Plant Mol. Biol. 2000. V. 44. P. 187–197.  
<https://doi.org/10.1023/a:1006431320677>
34. Mosquera T., Alvarez M.F., Jiménez-Gómez J.M. et al. Targeted and untargeted approaches unravel novel candidate genes and diagnostic SNPs for quantitative resistance of the potato (*Solanum tuberosum* L.) to *Phytophthora infestans* causing the late blight disease // PLoS One. 2016. V. 11. e0156254.  
<https://doi.org/10.1371/journal.pone.0156254>
35. Pajeroska-Mukhtar K., Stich B., Achenbach U. et al. Single nucleotide polymorphisms in the allene oxide synthase 2 gene are associated with field resistance to late blight in populations of tetraploid potato cultivars // Ge-

- netics. 2009. V. 181. P. 1115–1127.  
<https://doi.org/10.1534/genetics.108.094268>
36. Odeny D.A., Stich B., Gebhardt C. Physical organization of mixed protease inhibitor gene clusters, coordinated expression and association with resistance to late blight at the StKI locus on potato chromosome III // *Plant. Cell Environ.* 2010. V. 33. P. 2149–2161.  
<https://doi.org/10.1111/j.1365-3040.2010.02213.x>
  37. Muktar M.S., Lübeck J., Strahwald J., Gebhardt C. Selection and validation of potato candidate genes for maturity corrected resistance to *Phytophthora infestans* based on differential expression combined with SNP association and linkage mapping // *Front. Genet.* 2015. V. 6. P. 294.  
<https://doi.org/10.3389/fgene.2015.00294>
  38. Henriquez M.A., Soliman A., Li G. et al. Molecular cloning, functional characterization and expression of potato (*Solanum tuberosum*) 1-deoxy-d-xylulose 5-phosphate synthase 1 (StDXS1) in response to *Phytophthora infestans* // *Plant Sci.* 2016. V. 243. P. 71–83.  
<https://doi.org/10.1016/j.plantsci.2015.12.001>
  39. Colignon B., Dieu M., Demazy C. et al. Proteomic study of SUMOylation during *Solanum tuberosum*–*Phytophthora infestans* interactions // *Mol. Plant-Microbe Interact.* 2017. V. 30. P. 855–865.  
<https://doi.org/10.1094/MPMI-05-17-0104-R>
  40. Grenville-Briggs L.J., Anderson V.L., Fugelstad J. et al. Cellulose synthesis in *Phytophthora infestans* is required for normal appressorium formation and successful infection of potato // *Plant Cell.* 2008. V. 20. P. 720–738.  
<https://doi.org/10.1105/tpc.107.052043>
  41. Avrova A.O., Boevink P.C., Young V. et al. A novel *Phytophthora infestans* haustorium-specific membrane protein is required for infection of potato // *Cell. Microbiol.* 2008. V. 10. P. 2271–2284.  
<https://doi.org/10.1111/j.1462-5822.2008.01206.x>
  42. Eschen-Lippold L., Landgraf R., Smolka U. et al. Activation of defense against *Phytophthora infestans* in potato by down-regulation of syntaxin gene expression // *New Phytol.* 2012. V. 193. P. 985–996.  
<https://doi.org/10.1111/j.1469-8137.2011.04024.x>
  43. Ali S., Ganai B.A., Kamili A.N. et al. Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance // *Microbiol. Res.* 2018. V. 212–213. P. 29–37.
  44. Chandrasekaran M., Chun S.C. Expression of PR-protein genes and induction of defense-related enzymes by *Bacillus subtilis* CBR05 in tomato (*Solanum lycopersicum*) plants challenged with *Erwinia carotovora* subsp. *carotovora* // *Biosci. Biotechnol. Biochem.* 2016. V. 80. P. 2277–2283.  
<https://doi.org/10.1080/09168451.2016.1206811>
  45. Xiao C., Huang M., Gao J. et al. Comparative proteomics of three Chinese potato cultivars to improve understanding of potato molecular response to late blight disease // *BMC Genomics.* 2020. P. 21.  
<https://doi.org/10.1186/s12864-020-07286-3>
  46. Xiao C., Gao J., Zhang Y. et al. Quantitative proteomics of potato leaves infected with *Phytophthora infestans* provides insights into coordinated and altered protein expression during early and late disease stages // *Int. J. Mol. Sci.* 2019. P. 20.  
<https://doi.org/10.3390/ijms20010136>
  47. Chen Q., Tian Z., Jiang R. et al. StPOTHR1, a NDR1/HIN1-like gene in *Solanum tuberosum*, enhances resistance against *Phytophthora infestans* // *Biochem. Biophys. Res. Commun.* 2018. V. 496. P. 1155–1161.  
<https://doi.org/10.1016/j.bbrc.2018.01.162>
  48. Santa J.D., Berdugo-Cely J., Cely-Pardo L. et al. QTL analysis reveals quantitative resistant loci for *Phytophthora infestans* and *Tecia solanivora* in tetraploid potato (*Solanum tuberosum* L.) // *PLoS One.* 2018. V. 13. e0199716.  
<https://doi.org/10.1371/journal.pone.0199716>
  49. Lim P.O., Lee I.C., Kim J. et al. Auxin response factor 2 (ARF2) plays a major role in regulating auxin-mediated leaf longevity // *J. Exp. Bot.* 2010. V. 61. P. 1419–1430.  
<https://doi.org/10.1093/jxb/erq010>
  50. Vert G., Walcher C.L., Chory J., Nemhauser J.L. Integration of auxin and brassinosteroid pathways by Auxin Response Factor 2 // *Proc. Natl Acad. Sci. USA.* 2008. V. 105. P. 9829–9834.  
<https://doi.org/10.1073/pnas.0803996105>
  51. Koch A., Biedenkopf D., Furch A. et al. An RNAi-based control of *Fusarium graminearum* infections through spraying of long dsRNAs involves a plant passage and is controlled by the fungal silencing machinery // *PLoS Pathog.* 2016. V. 12. e1005901.  
<https://doi.org/10.1371/journal.ppat.1005901>
  52. Quattrocchio F., Verweij W., Kroon A. et al. PH4 of petunia is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway // *Plant Cell.* 2006. V. 18. P. 1274–1291.  
<https://doi.org/10.1105/tpc.105.034041>
  53. Chung I.M., Venkidasamy B., Upadhyaya C.P. et al. Alleviation of *Phytophthora infestans* mediated necrotic stress in the transgenic potato (*Solanum tuberosum* L.) with enhanced ascorbic acid accumulation // *Plants.* 2019. V. 8.  
<https://doi.org/10.3390/plants8100365>
  54. Li J., Zhu L., Lu G. et al. Curdian  $\beta$ -1,3-glucooligosaccharides induce the defense responses against *Phytophthora infestans* infection of potato (*Solanum tuberosum* L. cv. McCain G1) leaf cells // *PLoS One.* 2014. V. 9. e97197.  
<https://doi.org/10.1371/journal.pone.0097197>
  55. Rodenburg S.Y.A., Seidl M.F., de Ridder D., Govers F. Genome-wide characterization of *Phytophthora infestans* metabolism: A systems biology approach // *Mol. Plant Pathol.* 2018. V. 19. P. 1403–1413.  
<https://doi.org/10.1111/mpp.12623>
  56. Garavito M.F., Narvaez-Ortiz H.Y., Pulido D.C. et al. *Phytophthora infestans* dihydroorotate dehydrogenase is a potential target for chemical control – a comparison with the enzyme from *Solanum tuberosum* // *Front. Microbiol.* 2019. V. 10. P. 1479.  
<https://doi.org/10.3389/fmicb.2019.01479>
  57. Burra D.D., Berkowitz O., Hedley P.E. et al. Phosphite-induced changes of the transcriptome and secretome in *Solanum tuberosum* leading to resistance against *Phytophthora infestans* // *BMC Plant Biol.* 2014. V. 14. P. 254.  
<https://doi.org/10.1186/s12870-014-0254-y>

58. Leesutthiphonchai W., Vu A.L., Ah-Fong A.M.V., Judelson H.S. How does *Phytophthora infestans* evade control efforts? Modern insight into the late blight disease // *Phytopathology*. 2018. V. 108. P. 916–924. <https://doi.org/10.1094/PHTO-04-18-0130-IA>
59. López Orona C.A., Martínez A.R., Arteaga T.T. et al. First report of homothallic isolates of *Phytophthora infestans* in commercial potato crops (*Solanum tuberosum*) in the Toluca Valley, Mexico // *Plant Dis.* 2013. V. 97. P. 1112.
60. Clément J.A.J., Magalon H., Pellé R. et al. Alteration of pathogenicity-linked life-history traits by resistance of its host *Solanum tuberosum* impacts sexual reproduction of the plant pathogenic oomycete *Phytophthora infestans* // *J. Evol. Biol.* 2010. V. 23. P. 2668–2676. <https://doi.org/10.1111/j.1420-9101.2010.02150.x>
61. Wijekoon C.P., Peters R.D., Al-Mughrabi K.I., Kawchuk L.M. First report of late blight caused by *Phytophthora infestans* clonal lineage US-23 on tomato and potato in Atlantic Canada // *Plant Dis.* 2014. V. 3. P. 426. <https://doi.org/10.1094/PDIS-08-13-0807-PDN>
62. Cárdenas M.E., Medina E., Tabima J. et al. First report of *Phytophthora infestans* causing late blight on *Solanum viarum* in Colombia // *Plant Dis.* 2011. V. 95. P. 875. <https://doi.org/10.1094/pdis-11-10-0853>
63. Gevens A.J., Seidl A.C. First report of late blight caused by *Phytophthora infestans* clonal lineage US-24 on potato (*Solanum tuberosum*) in Wisconsin // *Plant Dis.* 2013. V. 97. P. 152. <https://doi.org/10.1094/PDIS-09-12-0825-PDN>
64. Wharton P.S., Nolte P., Kirk W.W. et al. First report of late blight caused by *Phytophthora infestans* clonal lineage US-23 on potato in Idaho // *Plant Dis.* 2015. V. 99. P. 417. <https://doi.org/10.1094/PDIS-02-14-0196-PDN>
65. Gevens A.J., Seidl A.C. First report of late blight caused by *Phytophthora infestans* clonal lineage US-22 on tomato and potato in Wisconsin United States // *Plant Dis.* 2013. V. 97. P. 423. <https://doi.org/10.1094/PDIS-08-12-0807-PDN>
66. Gevens A.J., Seidl A.C. First report of late blight caused by *Phytophthora infestans* clonal lineage US-23 on tomato and potato in Wisconsin, United States // *Plant Dis.* 2013. V. 97. P. 839. <https://doi.org/10.1094/PDIS-09-12-0821-PDN>
67. Fry W.E. *Phytophthora infestans*: New tools (and old ones) lead to new understanding and precision management // *Annu. Rev. Phytopathol.* 2016. V. 54. P. 529–547. <https://doi.org/10.1146/annurev-phyto-080615-095951>
68. Chowdappa P., Nirmal Kumar B.J., Madhura S. et al. Severe outbreaks of late blight on potato and tomato in South India caused by recent changes in the *Phytophthora infestans* population // *Plant Pathol.* 2015. V. 64. P. 191–199. <https://doi.org/10.1111/ppa.12228>
69. Fry W.E., McGrath M.T., Seaman A. et al. The 2009 late blight pandemic in the eastern United States – causes and results // *Plant Dis.* 2013. V. 97. P. 296–306. <https://doi.org/10.1094/PDIS-08-12-0791-FE>

## Molecular and Genetic Aspects of Potato Response to Late Blight Infection

T. S. Golubeva<sup>a, b, \*</sup>, V. A. Cherenko<sup>a, b</sup>, O. I. Sinitsyna<sup>a, b</sup>, and A. V. Kochetov<sup>a, b</sup>

<sup>a</sup>Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia

<sup>b</sup>Novosibirsk State University, Novosibirsk, 630090 Russia

\*e-mail: frolova@bionet.nsc.ru

Potato (*Solanum tuberosum*) is one of the most significant agricultural crops with the main pest – *Phytophthora infestans* – annually producers sustain big crop losses due to late blight varieties. The review generalizes the research articles since the 1960s and summarizes the basic experimental data about *S. tuberosum* and *P. infestans* interaction. The structure of the review includes 4 parts: in the introduction there is a historical digression for the comprehension of the *S. tuberosum* and *P. infestans* interaction with the very first and basic researches up to 2010. The further period is characterized by increasing of the number of researches on the review's issue which are grouped into 2 parts: the second part contains data about the genetic aspects of resistance; in the third part chemical approaches for resistance increasing are covered. The fourth part includes a description of promising approaches for reducing late blight's pathogenicity.

**Keywords:** *Phytophthora infestans*, *Solanum tuberosum*, late blight, pathogen resistance, R-gene, secondary metabolites, quantitative trait locus.

## ***Zvezda* – НОВОЕ ПОДСЕМЕЙСТВО *Tc1*-ПОДОБНЫХ ТРАНСПОЗОНОВ В ГЕНОМАХ ASTEROZOA**

© 2022 г. Л. В. Пузакова<sup>1</sup>, \*, М. В. Пузаков<sup>1, 2</sup>

<sup>1</sup>Федеральный исследовательский центр Институт биологии южных морей им. А.О. Ковалевского  
Российской академии наук, Севастополь, 299011 Россия

<sup>2</sup>Севастопольский государственный университет, Севастополь, 299053 Россия

\*e-mail: [kyluda@yandex.ru](mailto:kyluda@yandex.ru)

Поступила в редакцию 29.04.2021 г.

После доработки 02.06.2021 г.

Принята к публикации 23.06.2021 г.

Мобильные генетические элементы оказывают существенное влияние на эволюцию геномов живых организмов. Эукариотические МГЭ подразделяют на два класса – ретротранспозоны и ДНК-транспозоны. *ITm*-транспозоны являются одной из широко распространенных групп ДНК-транспозонов и обнаруживаются практически у всех организмов. В настоящей работе мы исследовали распространенность, структуру и эволюцию элементов с необычно длинным для *ITm*-транспозонов каталитическим доменом DD46E. Данные элементы были обнаружены только в подтипе *Asterozoa*. Была изучена их копияность, структура, возможная функциональность транспозазы, а также филогенетические взаимоотношения с другими представителями *ITm*-транспозонов. Показано, что обнаруженная нами группа транспозонов, которую мы назвали *Zvezda*, является подсемейством *Tc1*-подобных транспозонов.

**Ключевые слова:** мобильные генетические элементы, ДНК-транспозоны, DD46E-домен, *Asterozoa*, *Zvezda*.

**DOI:** 10.31857/S001667582201009X

Мобильные генетические элементы (МГЭ) являются неотъемлемой частью геномов как эукариотических, так и прокариотических организмов. Они представляют собой фрагменты ДНК, способные к перемещениям внутри хозяйского генома. В результате транспозиций МГЭ могут увеличивать число собственных копий, а также оказывать существенное влияние на структуру и функционирование всего генетического аппарата [1]. Более поздние исследования свидетельствуют о том, что мобильные генетические элементы могут служить источником новых генов [1].

Неоднократно показано, что активность МГЭ может возрасти в ответ на воздействие физических, химических и биологических факторов окружающей среды [2–4]. Активизация МГЭ в ответ на стресс вызывает дестабилизацию генома и всевозможные мутации, которые могут стать “сырьем” для действия движущего отбора. Таким образом, МГЭ играют немаловажную роль в адаптивных и эволюционных процессах [2–4].

В основе классификации МГЭ лежат различия в их структурно-функциональных особенностях. На сегодняшний день все эукариотические МГЭ подразделяются на два класса – ретротранспозо-

ны и ДНК-транспозоны [1, 5]. Ретротранспозоны, или класс I, кодируют обратную транскриптазу и перемещаются посредством создания собственной копии и ее вставки в какой-либо участок генома. Такой тип перемещения называется “копирование–вставка” и он позволяет ретротранспозонам быстро увеличивать число копий. ДНК-транспозоны, или класс II, кодируют транспозазу. Этот тип элементов перемещается путем вырезания собственной копии и ее последующей вставки в какой-либо другой участок генома. Такой тип перемещения называется “вырезание–вставка”. Несмотря на то что механизм увеличения количества копий транспозонов не описан, все же ДНК-транспозоны способны довольно эффективно “окупируют” хозяйские геномы.

Так, например, одна из распространенных групп ДНК-транспозонов – *IS630/Tc1/mariner* (*ITm*) насчитывает в геномах до нескольких сотен и даже тысяч копий [6–8]. Представители *ITm*-транспозонов присутствуют практически во всех живых организмах [9]. Длина элементов колеблется от 1 до 3 тыс. пар нуклеотидов (тпн), но может достигать и 6000 пн. *ITm*-элементы ограниче-

ны концевыми инвертированными повторами (КИП), протяженность которых очень вариabельна и колеблется от 20 до 1900 пн. У некоторых *ITm*-транспозонов имеются еще и субконцевые инвертированные повторы (СИП), длиной от 175 до 1403 пн [10]. Неразрывная открытая рамка считывания (ОРС) кодирует фермент транспозазу, протяженность которого варьирует в среднем от 350 до 650 аминокислотных остатков (а.о.). *ITm*-транспозаза имеет ДНК-связывающийся домен (PAIRED) в N-концевой части и каталитический (DDE/D) домен в C-концевой части [11]. Домен PAIRED состоит из шести  $\alpha$ -спиралей. Первая триада  $\alpha$ -спиралей носит название PAI-субдомена, вторая триада – RED-субдомена. Домен PAIRED обеспечивает сайт-специфическое связывание с ДНК-мишенью и с КИП. Между PAI- и RED-субдоменами расположен GRPR-подобный мотив, функцией которого, как предполагается, является связывание PAIRED-домена с малой бороздкой ДНК дублируемого сайта встраивания TA [12]. DDE/D-домен обладает эндонуклеазной и лигирующей активностью и обеспечивает вырезание и вставку МГЭ. Также для *ITm*-транспозазы характерно наличие сигнала ядерной локализации (NLS-сигнала), который, как предполагается, обеспечивает проникновение транспозазы из цитоплазмы в ядро [13, 14].

Классификация *ITm*-транспозонов достаточно сложная, в некоторых моментах противоречивая, и в последние годы претерпела значительные дополнения и изменения [15, 16]. Обобщая последние данные, можно выделить несколько основных групп: *Tc1*-подобные элементы (*TLE/DD34-38E*), *mariner*-подобные элементы (*MLE/DD34D*), *maT/DD37D*, *Visitor/DD41D*, *Guest/DD39D*, *mosquito/DD37E*, *pogo/DDxD* и *IS630/DDxE* [8, 10, 11, 16–18]. Группу элементов *pogo/DDxD* определяют как отдельное надсемейство [16]. Оставшиеся перечисленные группы (за исключением бактериальных транспозонов *IS630/DDxE*) объединяют в надсемейство *Tc1/mariner*. Еще выделяют четыре малых семейства: *Tec/DD34E*, *HvSm/DD34E*, *L31/DD37E* и *TBE/DD34E*, которые не входят в вышеперечисленные группы [15, 19].

В настоящей работе мы изучали распространенность, структуру и эволюцию ДНК-транспозонов надсемейства *Tc1/mariner* с необычно длинным каталитическим доменом DD46E, которые были названы нами *Zvezda*.

## МАТЕРИАЛЫ И МЕТОДЫ

### Поиск DD46E-транспозонов

Для поиска ДНК-транспозонов с каталитическим доменом DD46E был использован tBLASTn со стандартными настройками (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). В качестве матрицы для

поиска были взяты аминокислотные последовательности транспозаз *Mariner-18\_CGi\_p (Repbase)* и *Zvezda-1\_ARub/CABPRM03*. Полногеномные последовательности ДНК иглокожих были взяты из базы данных NCBI (<https://www.ncbi.nlm.nih.gov>) (табл. 1). Для того чтобы выявить полные нуклеотидные последовательности МГЭ, гомологичные транспозазе последовательности с наивысшей идентичностью к матрице были взяты из соответствующих скаффолдов вместе с фланкирующими областями протяженностью 3000 пн. Полноразмерная последовательность каждого выявленного элемента использовалась для подсчета присутствующих в геноме копий. Копии протяженностью менее 10% от длины полноразмерного МГЭ не учитывали при подсчете. Копии протяженностью от 95 до 100% от длины полноразмерного МГЭ считались полноразмерными. Копии протяженностью от 10 до 100% от длины полноразмерного МГЭ подсчитывались как общее количество копий.

### Анализ последовательностей

Границы предполагаемых ОРС определяли с помощью ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) и далее уточняли визуально. Последовательность сигнала ядерной локализации (NLS) выявляли с помощью PSORT (<https://www.genscript.com>). ДНК-связывающий мотив PAIRED определяли, анализируя вторичную структуру, предсказанную с помощью PSIPRED v4.0 [20]. Мотив GRPR-типа, а также DDE/D-домен идентифицировали визуально.

### Филогенетический анализ

Для филогенетического анализа были взяты аминокислотные последовательности транспозаз, относящиеся к разным группам *ITm*-транспозонов (табл. 2), и последовательности, принадлежащие транспозонам *Zvezda* (табл. 3). Множественное выравнивание аминокислотных последовательностей было выполнено с помощью MUSCLE [21] с использованием стандартных настроек. Филогенетический анализ проводили с использованием пакета программ MEGA X [21].

### Выявление случаев горизонтального переноса

Для выявления случаев горизонтального переноса у многоклеточных была использована аминокислотная последовательность *Zvezda-1\_ARub/CABPRM03*. Для поиска использовали программу BLASTn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Поскольку никаких соответствий с транспозонами *Zvezda* найдено не было, то и дальнейший поиск в соответствии с рекомендованным алгоритмом [22] не осуществлялся.

Таблица 1. Характеристика сборок геномных последовательностей, использованных для поиска транспозонов *Zvezda*

Класс	Вид	Место получения образца	Дата получения образца	Название геномного проекта	Идентификатор сборки	Уровень сборки	N50	Покрывтие генома	Размер сборки, пн	
Asteroidea	<i>Asterias rubens</i>	—	2018-02-23	SABPRM03	GCA_902459465	Хромосомы	20558067	103×	417601740	
		—	2018-02-23	SABPRO01	GCA_902459445	Скаффолды	161267	103×	390768305	
	<i>Pisaster ochraceus</i>	USA: California	2016-11-12	JAAF001	GCA_010994315	Хромосомы	20188303	212.8×	401943971	
		USA: California	2003	AKZP01	GCA_000285935	Скаффолды	52614	<b>15.0×</b> <b>454;</b> 70× Illumina	811028858	
	<i>Acanthaster planci</i>	USA: California	2019-09	JADOBP01	GCA_015706575	Скаффолды	23093800	150.0×	608344308	
		Japan: Okinawa, Motobu	2013-05-28	BDGF01	GCA_001949145	Скаффолды	1521119	46×	383860178	
	<i>Patiriella regularis</i>	Australia: Great Barrier Reef	2013-02-04	BDGH01	GCA_001949165	Скаффолды	916880	40×	383525304	
		Australia: Great Barrier Reef	2013-02-04	BBNW01*	GCA_000950615	Скаффолды	1389596	50×	<b>2584918</b>	
	Ophiuroidea	<i>Ophioneis fasciata</i>	—	—	CYSQ01	GCA_900067625	Скаффолды	<b>557</b>	<b>23×</b>	949333185
			—	—	CZLG01	GCA_900067615	Скаффолды	<b>484</b>	<b>13×</b>	1184528790
<i>Ophiothrix spiculata</i>		USA: California	—	JXSR01	GCA_000969725	Скаффолды	72780	275.0×	2764315159	

Примечание. Геномы со статистическим параметром N50 менее 10000, с геномным покрытием менее 25× и размером сборки менее 300 млн пн выделены жирным шрифтом. Звездочкой обозначена сборка, в которой не обнаружен транспозон *Zvezda*.

Таблица 2. ДНК-транспозоны IS630/Tc I/mariner, используемые в филогенетическом анализе

Семейство	Элемент	Источник	Семейство	Элемент	Источник
Guest DD39D	<i>Soymar1</i>	AF078934	<i>Visitor</i> DD41D	<i>Crmar2.5</i>	AAK61417
	<i>Br-oleracea</i>	XP013589454		<i>Apismar4.1</i>	ABLF02014333
	<i>Ca-sativa</i>	XP010462775		<i>Mariner-12_CGi</i>	Repbase
	<i>Phyllostachys edulis</i>	ADP24264		<i>Lsra_Ap</i>	[11]
	<i>Pisum sativum</i>	AAx51974		<i>rosa_Ae</i>	[11]
<i>maT</i> DD37D	<i>Bnmar1</i>	U47917	<i>Mosquito</i> DD37E	<i>Ae-atropalpus1</i>	AF377999
	<i>Bnmar6</i>	AF461149		<i>An-gambiae1</i>	AF378002
	<i>Cemar6</i>	LK928390		<i>PrDD37E1</i>	DQ138288
	<i>CbmaT4</i>	AC084524			
<i>Traveler</i> DD35E	<i>TR-Habu</i>	[27]	<i>L18</i> DD37E	<i>Mariner-18_CGi</i>	Repbase
	<i>TR-Xetr</i>	[27]		<i>L18-NVec</i>	NW_001834331
	<i>TR-Tajfu</i>	[27]		<i>L18-APar</i>	JXUT01105677
	<i>TR-Onmy</i>	[27]		<i>L18-HMag</i>	EQ256867
	<i>TR-Xihe</i>	[27]		<i>L18-AOce</i>	NXFZ01003443
<i>Intruder</i> DD38E	<i>IT_Cf</i>	[29]	<i>Incomer</i> DD36E	<i>In-MHra</i>	[28]
	<i>IT_Sl</i>	[29]		<i>In-Sale</i>	[28]
	<i>IT_Sp</i>	[29]		<i>In-MLuc</i>	[28]
	<i>IT_Rm</i>	[29]		<i>In-CVar</i>	[28]
	<i>IT_At</i>	[29]		<i>In-RMar</i>	[28]
<i>TcI</i> DD34E	<i>TcI</i>	X01005	<i>TRT</i> DD37E	<i>DrTRT</i>	[30]
	<i>Passport</i>	CAB5137		<i>CmpTRT</i>	[30]
	<i>Quetzal</i>	AA02109		<i>TfTRT</i>	[30]
	<i>Frog_Prince</i>	AAP49009		<i>SsTRT</i>	[30]
	<i>Sleeping_Beauty</i>	AFR53956		<i>HbTRT</i>	[30]
<i>cmTLE</i> DD34E	<i>Mariner-5_CGi</i>	Repbase	<i>TLEW1</i> DD36E	<i>TLEW1-1_CGi</i>	[26]
	<i>Mariner-23_CGi</i>	Repbase		<i>TLEW1-2_CGi</i>	[26]
	<i>Mariner-8_CGi</i>	Repbase		<i>TLEW1-2_MPh</i>	[26]
	<i>Mariner-14_CGi</i>	Repbase		<i>TLEW1-4_MYe</i>	[26]
<i>MLE</i> DD34D	<i>Dnmar1</i>	AAA28678	<i>IS630</i> DDxE	<i>RS(alpha)</i>	X02581
	<i>Hsmar1</i>	AAC52010		<i>IS630Ss</i>	X05955
	<i>Famar1</i>	AAO12863		<i>IS630Se</i>	NP_073225
	<i>Bytmar1</i>	CAD45367			
	<i>Tymar1</i>	AAP45328			

Таблица 3. ДНК-транспозоны *Zvezda* у *Asterozoa*

Вид	Название транспозона	Длина элемента, пн	Длина КИП	Длина СИП	Транспозаза, а.о.	Общее число копий	Число полноразмерных копий	Число потенциально функциональных копий
<i>Patiria miniata</i>	<i>Zvezda-1_PMin/AKZP01</i>	3742	12/12	н/о	340	59	1	1
	<i>Zvezda-1_PMin/JADOBP01</i>	757	н/о	н/о	248	12	0	0
<i>Pisaster ochraceus</i>	<i>Zvezda-1_POch</i>	3977	17/17	н/о	339	18	1	0
<i>Acanthaster planci</i>	<i>Zvezda-1_APla/BDGF01</i>	1056	н/о	н/о	334	32	0	0
	<i>Zvezda-1_APla/BDGH01</i>	1785	38/38	н/о	355	32	1	0
<i>Patiriella regularis</i>	<i>Zvezda-1_PReg</i>	1811	26/26	52/52	353	15	1	1
<i>Asterias rubens</i>	<i>Zvezda-1_ARub/CABPRM03</i>	1777	31/31	46/46	327	3	1	1
	<i>Zvezda-1_ARub/CABPRO01</i>	1772	31/31	46/46	289	3	2	0
<i>Ophioneis fasciata</i>	<i>Zvezda-1_OFas</i>	1086	н/о	н/о	361	2	0	0
<i>Ophiothrix spiculata</i>	<i>Zvezda-1.1_OSpi</i>	1542	н/о	156/–	366	493	1	0
	<i>Zvezda-1.2_OSpi</i>	3445	36/36	156/159	358			

Примечание. н/о – не обнаружены; пн – пары нуклеотидов; а.о. – аминокислотные остатки.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

*Распространенность транспозона  
Zvezda у Metazoa*

При исследовании разнообразия *ITm*-транспозонов у *Asterias rubens* нами был обнаружен МГЭ с необычно длинным каталитическим доменом — DD46E. Для того чтобы выяснить, единственный ли это мутантный вариант или представитель уникальной эволюционной группы *ITm*-транспозонов, мы изучили распространенность данного МГЭ. *ITm*-транспозоны с DD46E-доменом (названные нами *Zvezda*) обнаружались у семи организмов (т.е. у всех, имеющих в NCBI полные последовательности ДНК), принадлежащих к двум классам — морским звездам (*Asteroidea*) и офиурам (*Ophiuroidea*) (табл. 3). Эти два таксона объединены в подтип *Asterozoa*, который отделился от остальных иглокожих примерно 541 млн лет назад (по разным оценкам 458–625 млн лет назад, <http://www.timetree.org/>) [23]. При исследовании всех *Metazoa* на предмет гомологии к DD46E-транспозону *A. rubens* ни у кого, кроме *Asterozoa*, подобный транспозон обнаружен не был. Для объяснения данного феномена можно предложить две версии — либо ДНК-транспозон *Zvezda* возник непосредственно в геноме предка *Asterozoa*, либо встроился в геном предка *Asterozoa* в результате горизонтального переноса от неизвестного нам организма, который на сегодняшний день либо потерял его в процессе эволюции, либо вымер, либо его полная геномная последовательность ДНК еще не определена. Так или иначе, но ДНК-транспозон *Zvezda* не получил распространения по древу жизни и присутствует лишь в узкой филогенетической группе, внутри которой распространился, по всей видимости, в результате дивергенции видов. Отсутствие ДНК-транспозона *Zvezda* у каких-либо групп организмов, кроме *Asterozoa*, свидетельствует об отсутствии случаев горизонтального переноса.

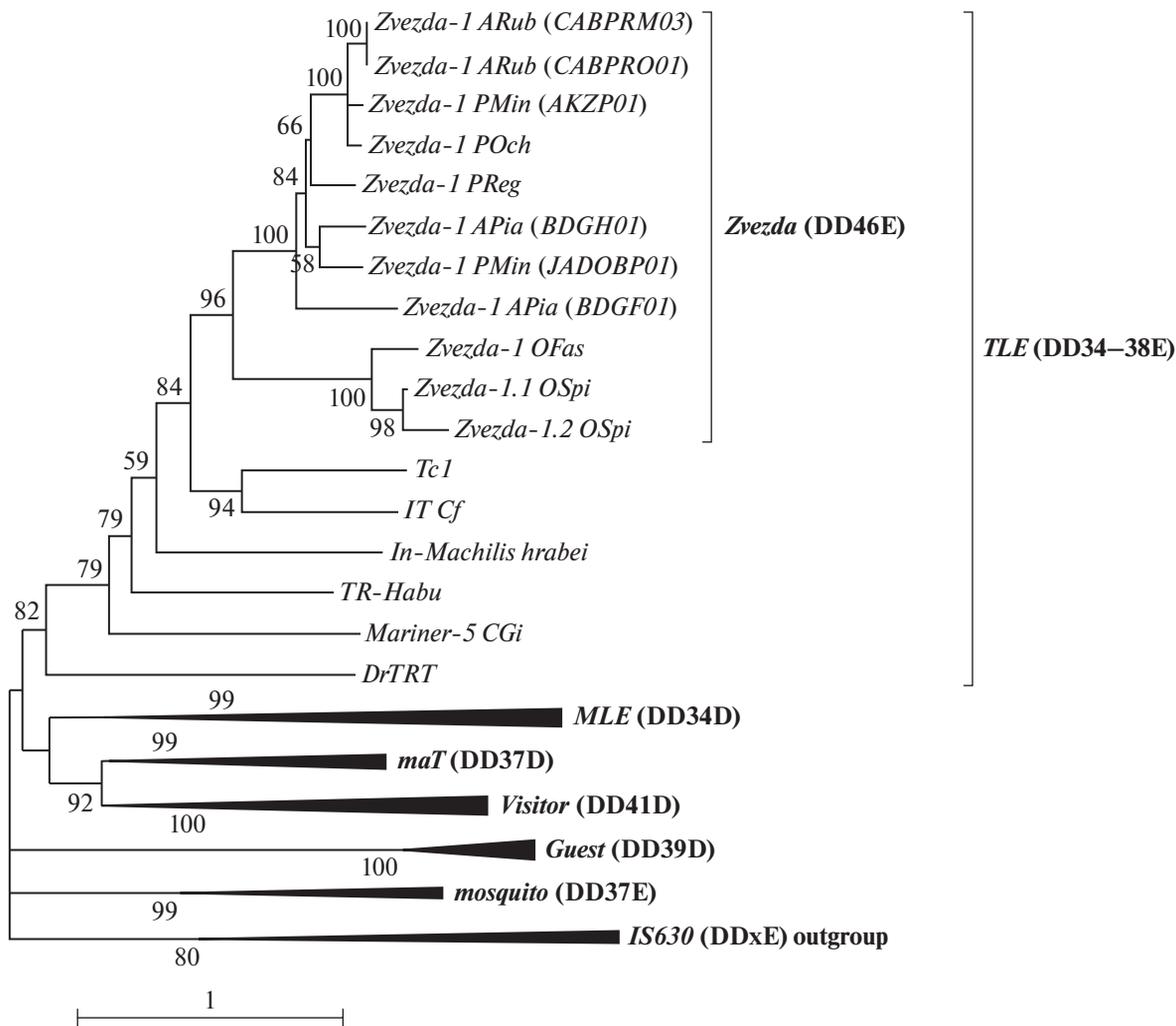
Анализ числа копий ДНК-транспозона *Zvezda* у *Asterozoa* показал значительную неоднородность среди изученных организмов (табл. 3). У двух видов число копий крайне низкое (2–3 копии). Четыре вида имели от 15 до 60 копий. И только один вид — *Ophiothrix spiculata* — имеет большое число копий — 493, при этом подавляющая часть этих копий сохранила КИП при делегированной центральной части. Согласно модели жизненного цикла, увеличение числа копий происходит на стадии амплификации в жизненном цикле МГЭ [24]. Однако мы не обнаружили в геноме *O. spiculata* ни одной потенциально функциональной копии элемента *Zvezda*, следовательно в настоящее время данный транспозон не может находиться на стадии амплификации. Очевидно, транспозон *Zvezda* был активен у предка *Astero-*

зоа, поэтому все *Asterozoa* в настоящий момент имеют малое число копий, большинство из которых повреждены и нефункциональны вследствие постепенной деградации. Вероятно, *O. spiculata* имеет на сегодняшний день так много копий по причине еще одного всплеска активности, произошедшего позже. Причиной второй волны активности мог быть рестарт жизненного цикла — одна из копий могла вновь стать активной вследствие случайных мутаций и вызвать “взрыв” перемещений. Также нельзя исключить и возможность кросс-индукции. Транспозаза какого-либо активного МГЭ, проникшего в геном *O. spiculata*, могла воздействовать и на копии элемента *Zvezda*, индуцируя их активность. Явление кросс-мобилизации уже было описано ранее [25]. Было обнаружено, что *hobo*-транспозаза способна взаимодействовать с терминальными последовательностями не только элемента *hobo*, но и *Hermes*. Хотя у *ITm*-транспозонов явление кросс-мобилизации не известно, такую возможность нельзя исключить, тем более что КИП, необходимые для перемещения, присутствуют практически во всех копиях.

*Zvezda — новое подсемейство  
Tc1-подобных транспозонов*

Для того чтобы определить эволюционные отношения элементов *Zvezda/DD46E* и *ITm*-транспозонов мы провели филогенетический анализ, в который были включены 11 транспозаз *Zvezda/DD46E* (с сохранившимся DDE-доменом) и 45 МГЭ, представляющих различные группы *ITm*-транспозонов (табл. 2). В результате было установлено, что элементы *Zvezda/DD46E* формируют отдельную ветвь, которая располагается в одной кладе с *Tc1*-подобными элементами (рис. 1). Таким образом, элементы *Zvezda/DD46E* являются новым подсемейством *Tc1*-подобных элементов.

На филогенетическом дереве видно, что транспозоны *Zvezda* объединяются в одну монофилетическую группу с высокой бутстрэп-поддержкой. В последние годы семейство *Tc1*-подобных элементов хорошо изучено и в нем найден ряд семейств, которые имеют домен, отличающийся от классического домена DD34E. Это, например, *TLEWI/DD36E*, *Traveler/DD35E*, *Incomer/DD36E*, *Intruder/DD38E*, *TRT/DD37E* [26–30]. Все они имеют не столь существенные отличия в размере домена от классического DD34E, в то время как у ДНК-транспозона *Zvezda* протяженность последовательности между вторым D и E составляет 46 а.о. (DD46E). Филогенетический анализ, в который были включены все известные на данный момент *Tc1*-подобные элементы, подтвердил принадлежность транспозонов *Zvezda* к отдельному подсемейству (рис. 2).



**Рис. 1.** Эволюционные взаимоотношения транспозонов *Zvezda* и ДНК-транспозонов *ITm*. Филогенетический анализ выполнен в программе MEGA X с помощью метода максимального правдоподобия. Используемая модель WAG + G + I + F. Бутстрэп 1000.

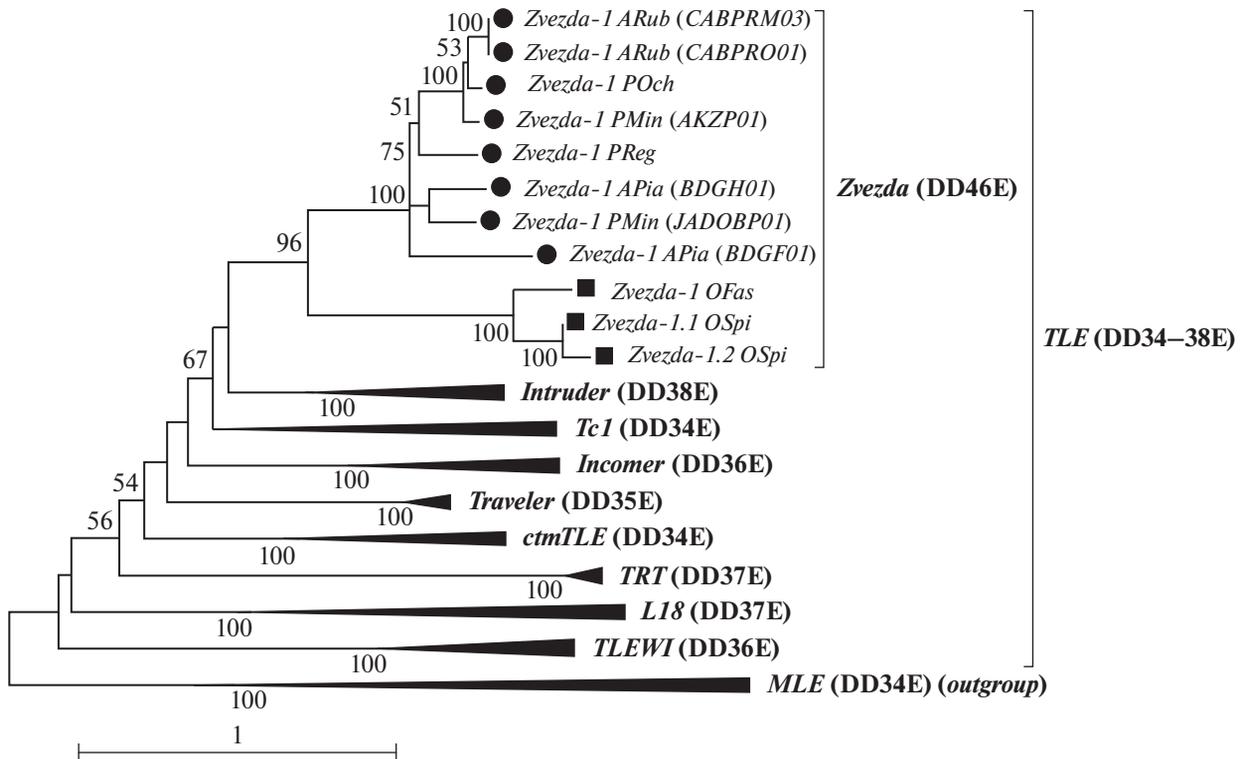
*Структурные особенности ДНК-транспозона Zvezda*

У большинства видов полноразмерные копии элемента *Zvezda* были обнаружены в единственном числе (табл. 3). Полноразмерными мы считали копии, обладающие обоими КИП и неповрежденной транспозазой. У *A. rubens* в геномной сборке CABPRO01 было обнаружено два полноразмерных элемента. У *Ophionereis fasciata* и *Ophiothrix spiculata* полноразмерных копий не было обнаружено.

В среднем длина полноразмерных элементов *Zvezda* колеблется от 1700 до 1800 пн (табл. 3), что характерно для *Tc1*-подобных элементов. Однако есть случаи, выбивающиеся из общей тенденции. У *O. spiculata* длина элемента – 3445 пн, у *Patiria miniate* (AKZP01) – 3742 пн, у *Pisaster ochraceus* – 3977 пн. В последних двух случаях КИП находятся далеко от гена транспозазы, вследствие этого

общая длина элемента сильно увеличена. У *O. spiculata* внутри гена транспозазы присутствуют вставки, увеличивающие длину элемента, а также один из КИП находится значительно дальше от гена транспозазы, чем обычно (табл. 3).

КИП были найдены у шести видов Asterozoa, СИП – у трех видов (табл. 3). Длина КИП колебалась от 26 до 38 пн. У *Patiria miniate* (AKZP01) и *Pisaster ochraceus* КИП были нетипично короткие, 12/12 и 17/17 пн соответственно. При этом они отстояли от гена транспозазы довольно далеко, в связи с чем общая длина элемента у этих видов превышает среднюю. СИП имели длину 46/46, 52/52, а в случае *O. spiculata* СИП были необычно длинными – 156/159 пн. Оба исследованных вида офиур имеют длинные СИП. У некоторых видов морских звезд также есть СИП, но значительно короче. Возможно, морские звезды потеряли СИП в процессе эволюции.



**Рис. 2.** Эволюционные взаимоотношения транспозонов *Zvezda* и *Tc1*-подобных транспозонов. Транспозоны морских звезд отмечены черным кругом, транспозоны офиур — черным квадратом. Филогенетический анализ выполнен в программе MEGA X с помощью метода максимального правдоподобия. Используемая модель LG + G + I. Бутстрэп 1000.

Длина полноразмерных транспозаз оказалась обычной для этого семейства транспозонов, а именно от 327 до 366 а.о. (табл. 3). Неповрежденные транспозазы имеются у четырех видов. У трех видов транспозазы имеют повреждения — стоп-кодны и делеции.

Чтобы лучше понять, являются ли транспозазы потенциально функциональными, мы изучили такие их особенности, как PAIRED-домен, АТ-крюк, каталитический домен DDE/D и NLS (рис. 3). Во всех транспозазах был обнаружен PAIRED-домен. Однако у МГЭ *Zvezda-1\_APla/BDGF01* и *Zvezda-1\_POch* первая альфа-спираль PAI-субдомена отсутствует. Также у *Zvezda-1\_APla/BDGF01* третья альфа-спираль RED-домена отсутствует. *ARub/CABPRO01* и *Zvezda-1\_PMin/JADOBP01* вообще не имеют PAI-субдомена.

Последовательность NLS не была обнаружена нами ни в одной из транспозаз *Zvezda*, что вполне закономерно, учитывая работы коллег, в которых показана редкая встречаемость NLS в *ITm*-транспозонах.

АТ-крюк обнаружен у всех исследованных видов, за исключением *Zvezda-1\_PMin/JADOBP01*. Он расположен на своем месте, между PAI- и RED-субдоменами.

Каталитический домен DD46E был обнаружен во всех транспозазах, но в аминокислотной последовательности транспозазы элемента *Zvezda-1\_APla/BDGF01* первый аспарат DDE/D-домена был замещен на аспарагин (N); кроме того, промежуток между вторым D (аспарат) и E (глутамат) составляет 47 аминокислотных остатков, а не 46, как во всех остальных транспозазах. Кроме того, данный транспозон имеет стоп-кодны и сдвиги рамки считывания. Так же у *Zvezda-1\_PMin/JADOBP01* каталитический домен был изменен (DD42E). Однако в других сборках этих же видов у мобильных элементов *Zvezda-1\_APla/BDGH01* и *Zvezda-1\_PMin/AKZP01* присутствует классический DD46E-домен.

Подытоживая результаты проведенного анализа, потенциально функциональными можно назвать только три элемента — *Zvezda-1\_PMin/AKZP01*, *Zvezda-1\_PReg* и *Zvezda-1\_ARub/CABPRM03* (рис. 3). Сравнивая полученные данные с имеющимися данными по другим *ITm*-транспозонам, можно сделать вывод, что только незначительное их количество имеют функциональную транспозазу и обладают транспозиционной активностью [31–34]. Этот факт связывают с тем, что пик активности *ITm*-транспозонов был в далеком прошлом, а следовательно “жизненный цикл” элементов проходит свою завершающую стадию [24]. По

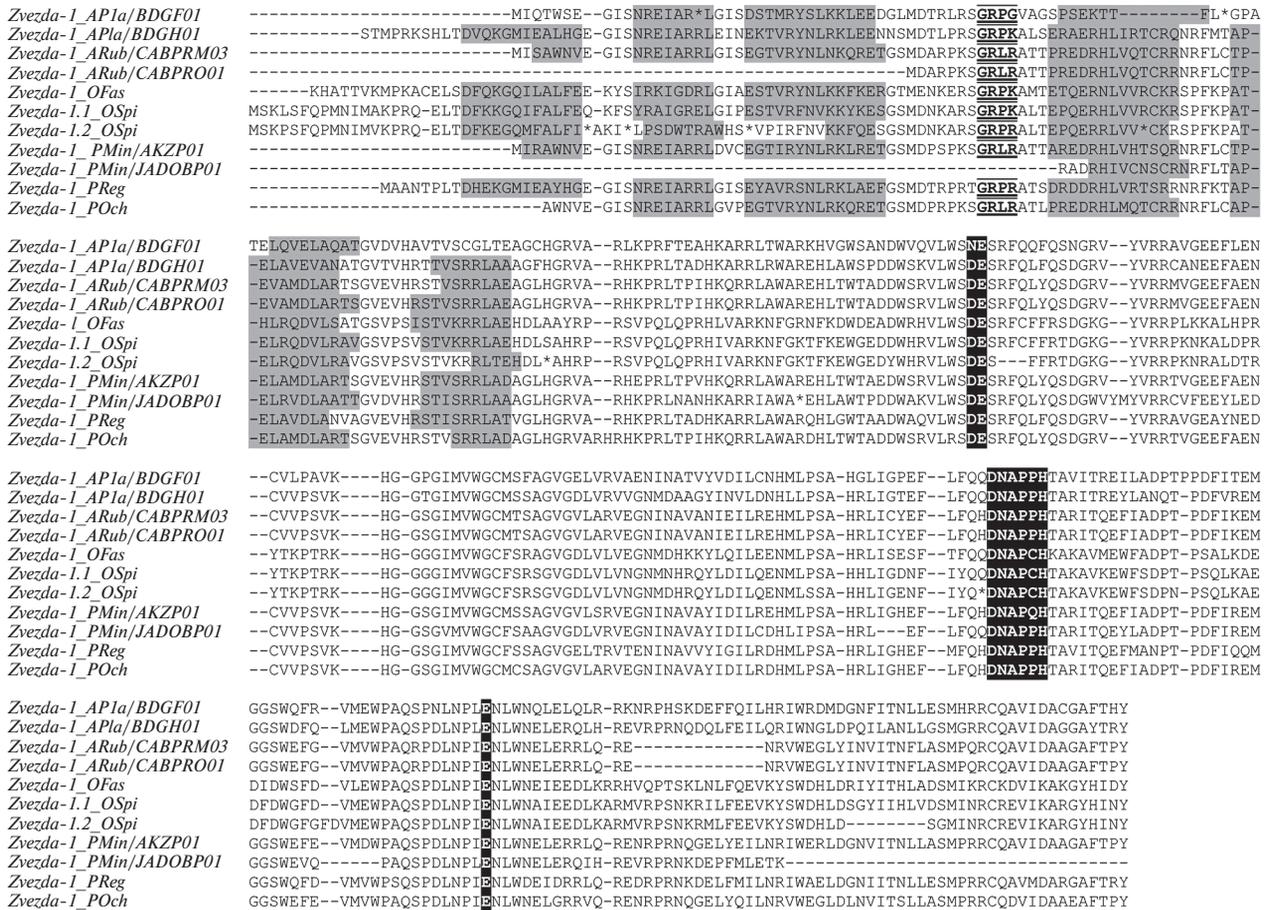


Рис. 3. Множественное выравнивание транспозаз элемента *Zvezda*. Серым цветом выделен PAIRED-домен, состоящий из шести α-спиралей, жирным шрифтом с подчеркиванием выделен GRPR-подобный мотив, белым шрифтом на черном фоне – DDE-домен.

этой причине у большинства *ITm*-транспозонов последовательности, кодирующие транспозазу, нарушены делециями, стоп-кодонами и сдвигами ОРС. Наличие одиночных полноразмерных/потенциально функциональных копий элемента *Zvezda* при наличии делетированных копий в геномах морских звезд и офиур может свидетельствовать о древности его происхождения.

Внутривидовое разнообразие

Были исследованы семь видов подтипа *Asterozoa*, полногеномные последовательности которых на момент исследования были представлены в NCBI (табл. 3). У всех семи видов был обнаружен мобильный элемент с каталитическим доменом DD46E. У *Acanthaster planci* имелось три сборки, одна из которых (GCA\_000950615.1) не имела сходства с матрицей, в качестве которой мы использовали аминокислотную последовательность элемента *Mariner-18 CGi*, обнаруженного у *Crasostrea gigas* [19]. У *Asterias rubens* и *Patiria miniata* имелось по две сборки.

Имея в доступе по две сборки видов *Acanthaster planci*, *Patiria miniata* и *Asterias rubens*, мы получили интересную возможность провести небольшое сравнение внутривидового разнообразия, хоть и на столь малом количестве особей. Мы сравнивали общее число транспозонов и их структуру.

Так, у *Asterias rubens* в двух разных сборках оказались удивительно сходные результаты – общее число транспозонов в сборке CABPRM03 и сборке CABPRO01 оказалось 3, полноразмерных копий 1 и 2 соответственно, терминальные и субтерминальные повторы оказались в обеих сборках абсолютно идентичными по длине и составили 31 и 46 пн соответственно; общая длина элементов отличалась несущественно – 1777 и 1772 пн, соответственно. Поскольку дата получения обоих образцов указана одна и та же, по всей видимости образцы были взяты в одном и том же месте, из одной популяции (табл. 1). Вероятно, данная популяция довольно гомогенна по количеству и структурным особенностям элемента *Zvezda*, что свидетельствует о том, что в данный момент этот транспозон не активен. Это подтверждается и ма-

лым количеством копий транспозона *Zvezda* в геноме *Asterias rubens*.

У *Acanthaster planci* в обеих сборках оказалось по 32 копии транспозона *Zvezda*. Образцы были взяты из разных регионов — *A. planci* (BDGF01) собран в Японии, а *A. planci* (BDGH01) в Австралии в одном и том же году. Равное число копий можно считать как совпадением, так и признаком однородности популяций. Сравнить строение транспозона *Zvezda* у *A. planci* оказалось невозможным, так как в одной из сборок (BDGF01) не было полноразмерных копий транспозона *Zvezda* (табл. 3). В базе данных присутствует и третья сборка *A. planci*, но результат поиска транспозона *Zvezda* в программе BLAST отрицательный, по всей видимости из-за частичной представленности генома (табл. 1).

Сходная ситуация и у *Patiria miniata*, также представленной двумя сборками, — в одной из двух сборок ОРС транспозона *Zvezda* также отсутствовали полноразмерные копии, следовательно структуру транспозонов сравнить нам не удалось. Количество копий *Zvezda* в этих двух геномах *P. miniata* — 59 и 12 (табл. 3). Образцы были взяты с промежутком в 16 лет, местом сбора указана Калифорния. Разное число копий в геномах образцов *P. miniata* может быть связано с внутривидовым разнообразием, а также и с тем, что сборка генома *Patiria miniata*/JADOBP01 неполная, поскольку общий размер сборки *Patiria miniata*/JADOBP01 на 25% меньше, чем у *Patiria miniata*/AKZP01.

Изучение МГЭ *Zvezda* проведено в рамках Государственного задания ФИЦ ИнБЮМ “Функциональные, метаболические и токсикологические аспекты существования гидробионтов и их популяций в биотопах с различным физико-химическим режимом”, номер гос. регистрации 121041400077-1.

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Bourque G., Burns K.H., Gehring M. et al. Ten things you should know about transposable elements // *Genome Biol.* 2018. V. 19. P. 199. <https://doi.org/10.1186/s13059-018-1577-z>
2. Юрченко Н.Н., Коваленко Л.В., Захаров И.К. Мобильные генетические элементы: нестабильность генов и геномов // Вавилов. журн. генетики и селекции. 2011. Т. 15. № 2. С. 261–270.
3. Piacentini L., Fanti L., Specchia V. et al. Transposons, environmental changes, and heritable induced phenotypic variability // *Chromosoma.* 2014. V. 123. P. 345–354.
4. Auvinet J., Graça P., Belkadi L. et al. Mobilization of retrotransposons as a cause of chromosomal diversification and rapid speciation: The case for the Antarctic teleost genus *Trematomus* // *BMC Genom.* 2018. V. 19. P. 339.
5. Kojima K.K. Structural and sequence diversity of eukaryotic transposable elements // *Genes Genet Syst.* 2020. V. 94. P. 233–252. Epub 2018 Nov 9. PMID: 30416149. <https://doi.org/10.1266/ggs.18-00024>
6. Lee C.C., Wang J. Rapid expansion of a highly germline-expressed *Mariner* element acquired by horizontal transfer in the fire ant genome // *Genome Biol. Evol.* 2018. V. 10. P. 3262–3278. <https://doi.org/10.1093/gbe/evy220>
7. Xie L.-Q., Wang P.-L., Jiang S.-H. et al. Genome-wide identification and evolution of *TC1/Mariner* in the silkworm (*Bombyx mori*) genome // *Genes Genomics.* 2018. V. 40. P. 485–495. <https://doi.org/10.1007/s13258-018-0648-6>
8. Shen D., Gao B., Miskey C. et al. Multiple invasions of *Visitor*, a DD41D family of *Tc1/mariner* transposons, throughout the evolution of vertebrates // *Genome Biol. Evol.* 2020. V. 12. P. 1060–1073. <https://doi.org/10.1093/gbe/evaa135>
9. Yuan Y.W., Wessler S.R. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies // *Proc. Natl Acad. Sci. USA.* 2011. V. 108. P. 7884–7889. <https://doi.org/10.1073/pnas.1104208108>
10. Zhang H.H., Shen Y.H., Xiong X.M. et al. Identification and evolutionary history of the DD41D transposons in insects // *Genes Genom.* 2016. V. 38. P. 109–117. <https://doi.org/10.1007/s13258-015-0356-4>
11. Tellier M., Claeys Bouuaert C., Chalmers R. *Mariner* and the *ITm* superfamily of transposons // *Microbiol. Spectrum.* 2015. V. 3. <https://doi.org/10.1128/microbiolspec.MDNA3-0033-2014>
12. Ivics Z., Izsvák Z. *Sleeping beauty* transposition // *Microbiol. Spectrum.* 2015. V. 3. <https://doi.org/10.1128/microbiolspec.MDNA3-0042-2014>
13. Ivics Z., Hackett P.B., Plasterk R.H., Izsvák Z. Molecular reconstruction of *Sleeping Beauty*, a *Tc1*-like transposon from fish, and its transposition in human cells // *Cell.* 1997. V. 91. P. 501–510.
14. Plasterk R.H., Izsvák Z., Ivics Z. Resident aliens the *Tc1/mariner* superfamily of transposable elements // *Trends in Genetics.* 1999. V. 15. P. 527–538.
15. Dupeyron M., Baril T., Bass C., Hayward A. Phylogenetic analysis of the *Tc1/mariner* superfamily reveals the unexplored diversity of *pogo*-like elements // *Mobile DNA.* 2020. V. 11. P. 21. <https://doi.org/10.1186/s13100-020-00212-0>
16. Gao B., Wang Y., Diaby M. Evolution of *pogo*, a separate superfamily of *IS630-Tc1-mariner* transposons, revealing recurrent domestication events in vertebrates // *Mobile DNA.* 2020. V. 11. P. 25. <https://doi.org/10.1186/s13100-020-00220-0>
17. Shao H., Tu Z. Expanding the diversity of the *IS630-Tc1-mariner* superfamily: discovery of a unique DD37E transposon and reclassification of the DD37D and

- DD39D transposons // Genetics. 2001. V. 159. P. 1103–1115.
18. Wang S., Diaby M., Puzakov M. et al. Divergent evolution profiles of DD37D and DD39D families of *Tc1/mariner* transposons in eukaryotes // Mol. Phylogenet. Evol. 2021. V. 161. A. 107143. <https://doi.org/10.1016/j.ympev.2021.107143>
  19. Puzakov M.V., Puzakova L.V., Cheresiz S.V. An analysis of IS630/*Tc1/mariner* transposons in the genome of a Pacific oyster, *Crassostrea gigas* // J. Mol. Evol. 2018. V. 86. P. 566–580. <https://doi.org/10.1007/s00239-018-9868-2>
  20. Buchan D.W.A., Jones D.T. The PSIPRED protein analysis workbench: 20 years on // Nucl. Acids Res. 2019. V. 47. P. 402–407. <https://doi.org/10.1093/nar/gkz297>
  21. Kumar S., Stecher G., Li M. et al. MEGA X: Molecular evolutionary genetics analysis across computing platforms // Mol. Biol. Evol. 2018. V. 35. P. 1547–1549.
  22. Filée J., Rouault J.D., Harry M., Hua-Van A. Mariner transposons are sailing in the genome of the blood sucking bug *Rhodnius prolixus* // BMC Genomics. 2015. V. 16. P. 1061. <https://doi.org/10.1186/s12864-015-2060-9>
  23. Kumar S., Stecher G., Suleski M., Hedges S.B. TimeTree: a resource for timelines, timetrees, and divergence times // Mol. Biol. Evol. 2017. V. 34. P. 1812–1819. <https://doi.org/10.1093/molbev/msx116>
  24. Schaack S., Gilbert C., Feschotte C. Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution // Trends Ecol. Evol. 2010. V. 25. P. 537–546.
  25. Sundararajan P., Atkinson P.W., O'Brochta D.A. Transposable element interactions in insects: crossmobilization of *hobo* and *Hermes* // Insect Mol. Biol. 1999. V. 8. № 3. P. 359–368.
  26. Puzakov M.V., Puzakova L.V., Cheresiz S.V. The *Tc1*-like elements with the spliceosomal introns in mollusk genomes // Mol. Genet. Genomics. 2020. V. 295. P. 621–633.
  27. Zong W., Gao B., Diaby M. et al. Traveler, a new DD35E family of *Tc1/mariner* transposons, invaded vertebrates very recently // Genome Biol. Evol. 2020. V. 12. P. 66–76.
  28. Sang Y., Gao B., Diaby M. et al. Incomer, a DD36E family of *Tc1/mariner* transposons newly discovered in animals // Mobile DNA. 2019. V. 10. P. 45.
  29. Gao B., Zong W., Miskey C. et al. Intruder (DD38E), a recently evolved sibling family of DD34E/*Tc1* transposons in animals // Mobile DNA. 2020. V. 11. P. 32.
  30. Zhang H.H., Li G.Y., Xiong X.M. et al. TRT, a vertebrate and protozoan *Tc1*-like transposon: current activity and horizontal transfer // Genome Biol. Evol. 2016. V. 8. P. 2994–3005.
  31. Emmons S.W., Yesner L., Ruan K., Katzenberg D. Evidence for a transposon in *Caenorhabditis elegans* // Cell. 1983. V. 32. P. 55–65.
  32. Franz G., Savakis C. Minos, a new transposable element from *Drosophila hydei*, is a member of the *Tc1*-like family of transposons // Nucl. Acids Res. 1991. V. 19. P. 6646.
  33. Langin T., Capy P., Daboussi M.J. The transposable element *impala*, a fungal member of the *Tc1*-mariner superfamily // Mol. Gen. Genet. 1995. V. 246. P. 19–28.
  34. Clark K.J., Carlson D.F., Leaver M.J. et al. Passport, a native *Tc1* transposon from flatfish, is functionally active in vertebrate cells // Nucl. Acids Res. 2009. V. 37. P. 1239–1247.

## Zvezda – A New Subfamily of *Tc1*-Like Transposons in Asterozoa Genomes

L. V. Puzakova<sup>a, \*</sup> and M. V. Puzakov<sup>a, b</sup>

<sup>a</sup>Kovalevsky Institute of Biology of the Southern Seas of Russian Academy of Science, Sevastopol, 299011 Russia

<sup>b</sup>Sevastopol State University, Sevastopol, 299053 Russia

\*e-mail: [kvluda@yandex.ru](mailto:kvluda@yandex.ru)

Mobile genetic elements have a significant impact on the evolution of the genomes of living organisms. They are divided into two classes – retrotransposons and DNA transposons. *ITm* transposons are one of the widespread groups of DNA transposons and are found in almost all organisms. In this work, we investigated the abundance, structure, and evolution of elements with the DD46E catalytic domain that is unusually long for *ITm* transposons. These elements were found only in the subtype Asterozoa. Their copy number, structure, possible functionality of transposase, as well as phylogenetic relationships with other representatives of *ITm* transposons were studied. It was shown that the group of transposons that we discovered, which we named *Zvezda*, is a subfamily of *Tc1*-like transposons.

**Keywords:** transposable elements, DNA transposons, DD46E domain, Asterozoa, *Zvezda*.

ДВУХСТУПЕНЧАТАЯ АКТИВАЦИЯ *lux*-РЕГУЛОНА ПСИХРОФИЛЬНЫХ  
ЛЮМИНЕСЦИРУЮЩИХ БАКТЕРИЙ *Aliivibrio logei*© 2022 г. С. В. Баженов<sup>1, \*</sup>, Е. С. Щеглова<sup>1</sup>, В. В. Фомин<sup>1</sup>,  
Г. Б. Завильгельский<sup>2</sup>, И. В. Манухов<sup>1, 2, \*\*</sup><sup>1</sup>Московский физико-технический институт, Московская область, Долгопрудный, 141701 Россия<sup>2</sup>Государственный научно-исследовательский институт генетики и селекции промышленных микроорганизмов  
Национального исследовательского центра “Курчатовский институт”, Москва, 117545 Россия

\*e-mail: bazhenov1994@gmail.com

\*\*e-mail: manukhovi@mail.ru

Поступила в редакцию 30.04.2021 г.

После доработки 01.07.2021 г.

Принята к публикации 12.07.2021 г.

*Lux*-регулон психрофильных морских люминесцентных бактерий *Aliivibrio logei* контролируется “quorum sensing”-системой типа LuxI/LuxR. В данной работе определен характер аутоиндуктор-зависимой регуляции промоторов *luxI* и *luxCDABEG*-генов как по скорости синтеза аутоиндуктора и интенсивности биолюминесценции клеток *A. logei*, так и в гетерологичной системе в клетках *Escherichia coli* с репортерными генами *luxCDABE Photorhabdus luminescens*, находящимися под контролем исследуемых промоторов. Показано, что существуют различия в аутоиндуктор-зависимой активации оперона *luxCDABEG* и отдельно расположенного гена *luxI A. logei*. Экспрессия *luxI* усиливается в присутствии N-3-оксо-гексаноил-L-гомосерин лактона в концентрациях порядка 10 нМ, а экспрессия генов *luxCDABEG* усиливается при значительно больших концентрациях – от 1 мкМ и выше. Это определяется как последовательностью сайта посадки регуляторных белков *lux*-бокс2 и *lux*-бокс1 в составе соответствующих промоторов, так и свойствами двух белков LuxR1 и LuxR2, различающихся по способности связываться с аутоиндукторами.

**Ключевые слова:** чувство кворума, люминесценция, аутоиндуктор, регуляция, *Aliivibrio logei*, *luxR*.

DOI: 10.31857/S0016675822020023

Люминесценция морских бактерий *Aliivibrio fischeri* и *Aliivibrio logei* управляется по механизму “quorum sensing” (QS), позволяющему за счет синтеза специальных сигнальных веществ, аутоиндукторов (АИ) и детекции их с помощью регуляторных белков типа LuxR осуществлять скоординированную регуляцию экспрессии генов с использованием N-3-оксо-гексаноил-L-гомосерин лактона (ЗОС6-HSL) в качестве АИ [1–5]. У бактерий вида *A. fischeri* гены *luxCDABEG*, определяющие светимость клетки, объединены в один оперон с геном *luxI*, отвечающим за синтез АИ, и регулируются совместно, а регуляторный ген *luxR* расположен рядом с промотором оперона, но ориентирован в противоположном направлении [6]. Психрофильные бактерии *A. logei* значительно отличаются от мезофильных бактерий вида *A. fischeri* по архитектуре *lux*-оперона: ген *luxI* и касета *luxCDABEG* транскрибируются с отдельных промоторов и регулируются двумя гомологичными генами *luxR2* и *luxR1*, расположенными

на хромосоме непосредственно рядом с регулируемым промоторами [5, 7].

Ранее в гетерологичной системе в клетках *Escherichia coli* было проведено сравнение регуляции промоторов генов *luxI* и *luxCDABEG A. logei* [8]. Исследуемые промоторы клонировались вместе с ближайшими регуляторными генами:  $P_{luxI}$  вместе с *luxR2* и  $P_{luxCDABEG}$  вместе с *luxR1*. Было показано, что комбинация  $P_{luxI}$  с *luxR2* обеспечивает большую чувствительность к АИ и большую амплитуду индукции, чем комбинация  $P_{luxCDABEG}$  с *luxR1*. В работе [9] в экспериментах в гетерологичной системе клеток *E. coli* было показано, что внесение гена *luxR2 A. logei* в *trans*-положение увеличивает чувствительность к АИ и амплитуду ответа промотора  $P_{luxCDABEG}$ , клонированного в *cis*-положении с геном *luxR1*.

В настоящей работе была исследована регуляция *luxI* и *luxCDABEG* промоторов в клетках *A. logei* и проведено сравнение результатов с данными, полученными в гетерологичных системах.

**Таблица 1.** Плазмиды, использованные в работе

Плаزمида	Описание	Ссылка на источник
pIVA	Вектор pDEW201, в котором <i>luxR1 A. lojei</i> под контролем P <sub>luxR1</sub> <i>A. lojei</i> , <i>luxCDABE P. luminescens</i> под контролем P <sub>luxCDABEG</sub> <i>A. lojei</i> , Ap <sup>r</sup>	[8]
pSV16	pDEW201, в котором <i>luxR2 A. lojei</i> под контролем P <sub>luxR2</sub> <i>A. lojei</i> , <i>luxCDABE P. luminescens</i> под контролем P <sub>lux1</sub> <i>A. lojei</i> , Ap <sup>r</sup>	[13]
pIV3	Вектор pACYC184, в котором <i>luxR1 A. lojei</i> под контролем собственного промотора, встроенный в <i>Bam</i> HI сайт, Cm <sup>r</sup>	[8]
pIV2	pACYC184, в котором <i>luxR2 A. lojei</i> под контролем собственного промотора, встроенный в <i>Bam</i> HI сайт, Cm <sup>r</sup>	[8]
pD-lb1	pDEW201, в котором <i>luxCDABE P. luminescens</i> под контролем P <sub>luxCDABEG</sub> <i>A. lojei</i> с нативной последовательностью <i>lux</i> -бокс1: CTCTGTAAAGTTATACAGGT	Настоящая работа
pD-lb2	pDEW201, в котором <i>luxCDABE P. luminescens</i> под контролем P <sub>luxCDABEG</sub> <i>A. lojei</i> с заменой последовательности <i>lux</i> -бокс1 на <i>lux</i> -бокс2: TCCTGTAATATTGTACAGGT	Настоящая работа
pVFR1	pDEW201, в котором <i>luxR A. fischeri</i> под контролем собственного промотора, <i>luxCDABE P. luminescens</i> под контролем P <sub>luxCDABEG</sub> <i>A. fischeri</i> , Ap <sup>r</sup>	[14]
pSVRAF	pACYC184, в котором <i>luxR A. fischeri</i> под контролем собственного промотора, встроенный в <i>Bam</i> HI сайт, Cm <sup>r</sup>	[15]

## МАТЕРИАЛЫ И МЕТОДЫ

### Бактериальные штаммы и плазмиды

В работе использовались штаммы *A. lojei* K18-44 и *A. lojei* KCh1 дикого типа [3, 10] и *E. coli* MG1655 [11], трансформированный гибридными плазмидами. Для исследования регуляции отдельных промоторов использовались гибридные плазмиды на основе вектора pDEW201 [12], в котором перед генами *luxCDABE Photorhabdus luminescens* встраивались исследуемые промоторы (см. табл. 1).

### Условия культивирования

Клетки *A. lojei* выращивались при температуре 12–16°C в жидкой среде SWT (морская соль 15 г/л, триптон 5 г/л, др. экстракт 2.5 г/л, глицерин 3 г/л) с постоянным перемешиванием (150 об./мин) или на чашках с агаризованной (15 г/л) SWT. Клетки *E. coli* выращивали в среде LB при температуре 37°C.

### Конструирование гибридных плазмид

Для получения pD-lb2 фрагмент хромосомальной ДНК *A. lojei* KCh1 был амплифицирован с использованием праймеров 5'-GATTTCCTGTAATATTGTACAGGTTTACCTAAATAATTACCCTGCTA-3' и 5'-GACACCGCCGATGATAATTGGA-3', клонирован в pTZ57R/T векторе, а затем перенесен по сайтам рестрикции *Eco*RI/*Bam*HI в вектор pDEW201. В последовательности праймера

подчеркиванием выделена последовательность *lux*-бокс. Для получения pD-lb1 фрагмент хромосомальной ДНК *A. lojei* KCh1 был амплифицирован с использованием праймеров 5'-GGATCCGATCTCTGTAAAGTTATACAGGTTTACCTA-3' и 5'-GACACCGCCGATGATAATTGGA-3', клонирован в pAL2-T векторе, а затем перенесен по ориентации вставки так, чтобы гены *luxCDABE P. luminescens* были под контролем встроенного промотора *A. lojei*.

### Реактивы

Сборка гибридных плазмид проводилась с использованием ДНК-полимеразы Taq (Евроген, Россия), эндонуклеаз рестрикции *Eco*RI и *Bam*HI (Promega, США), фосфатазы и лигазы T4 (Thermo Scientific, США). Аутоиндуктор 3OC6-HSL приобретен у SigmaAldrich (США). Для приготовления SWT использовали морскую соль производства RedSea (Гватемала). Триптон, дрожжевой экстракт, агар-агар и хлорид натрия были приобретены у компании “Диа-М” (Россия), глицерин – AppliChem (США).

### Измерение биолюминесценции клеток

Биолюминесценцию клеток измеряли с использованием планшетного люминометра Synergy HT (Biotek, США) и высокочувствительного

кюветного люминометра “Биотокс-7ВМ” (Био-ФизТех, Россия) при комнатной температуре.

В экспериментах с биосенсорными клетками *E. coli* культуру клеток выращивали до OD 0.1–0.2, затем разделяли на аликвоты по 180 мкл, к клеткам добавляли контрольный раствор, раствор аутоиндуктора или исследуемые аутоиндуктор-содержащие среды в различных разведениях (по 20 мкл), после чего проводили периодические измерения светимости. В эксперименте с клетками *A. logei* суспензию клеток по 200 мкл переносили в кювету непосредственно перед измерением.

#### Определение концентрации АИ в образцах надосадочной культуры клеток *A. logei*

Определение концентрации АИ в растворе проводили с использованием цельноклеточного *lux*-биосенсора *E. coli* MG1655 pVFR1 pSVRAF (за счет повышенного содержания регуляторного гена *luxR* *A. fischeri* достигается повышенная чувствительность к АИ [16]), люминесценция которого возрастает в дозозависимой манере после добавления АИ в концентрациях от 0.03 до 100 нМ. Калибровку проводили по ЗОС6-HSL – основному АИ в QS типа LuxR/LuxI в бактериях рода *Aliivibrio* [4, 17].

#### Определение скорости синтеза АИ в зависимости от концентрации АИ в среде

Прирост концентрации АИ определяли следующим образом: 1 – измерение  $[AI]_{t1}$ ; 2 – инкубация определенным промежутком времени; 3 – измерение концентрации АИ ( $[AI]_{t2}$ ); 4 – расчет прироста концентрации АИ по формуле

$$\Delta[AИ] = [AI]_{t2} - [AI]_{t1}, \quad (1)$$

где  $t1$  и  $t2$  – время забора проб для измерения концентрации АИ.

Скорость синтеза АИ в зависимости от концентрации АИ в среде определяли следующим образом: 1 – измерение  $[AI]_{t1}$ ; 2 – экзогенное добавление АИ от 1 нМ до 10 мкМ ( $[AI]_{ex}$ ), инкубация 4 ч; 3 – измерение концентрации АИ  $[AI]_{t2}$ ; 4 – расчет прироста концентрации АИ по формуле

$$\Delta[AИ] = [AI]_{t2} - [AI]_{ex} - [AI]_{t1}. \quad (2)$$

При определении удельной скорости синтеза АИ значение  $\Delta[AИ]$  делили на площадь под графиком зависимости оптической плотности культуры клеток *A. logei* от времени:

$$\left(\frac{d}{dt}[AI]_{уд}\right) = \Delta[AИ] / \int_{t1}^{t2} OD(t) dt, \quad (3)$$

где  $OD(t)$  – функция зависимости оптической плотности культуры клеток *A. logei* от времени.

#### Статистическая обработка результатов

Все эксперименты с культурами клеток *A. logei* и *E. coli* проводились в трех повторах. Погрешность при измерении люминесценции и скорости синтеза АИ вычислялась по формуле стандартного отклонения на основе трех повторов, ее значение на графиках отражено планками ошибок.

## РЕЗУЛЬТАТЫ

#### Люминесценция и синтез АИ клетками *A. logei*

Для изучения работы QS-системы типа LuxI/LuxR в клетках бактерий *A. logei*, в частности регуляции генов *luxCDABEG* и *luxI*, проводилось культивирование клеток *A. logei* K18-44 в жидкой среде с регулярным измерением оптической плотности, люминесценции и концентрации АИ в среде (рис. 1,а). На основе полученных данных были рассчитаны удельная светимость и удельная скорость синтеза АИ в расчете на единицу оптической плотности культуры клеток (рис. 1,б).

В ранних фазах роста культуры клетки делятся с постоянной скоростью, и зависимость оптической плотности культуры приближенно описывается следующей формулой:

$$OD(t) = OD_0 e^{at}, \quad (4)$$

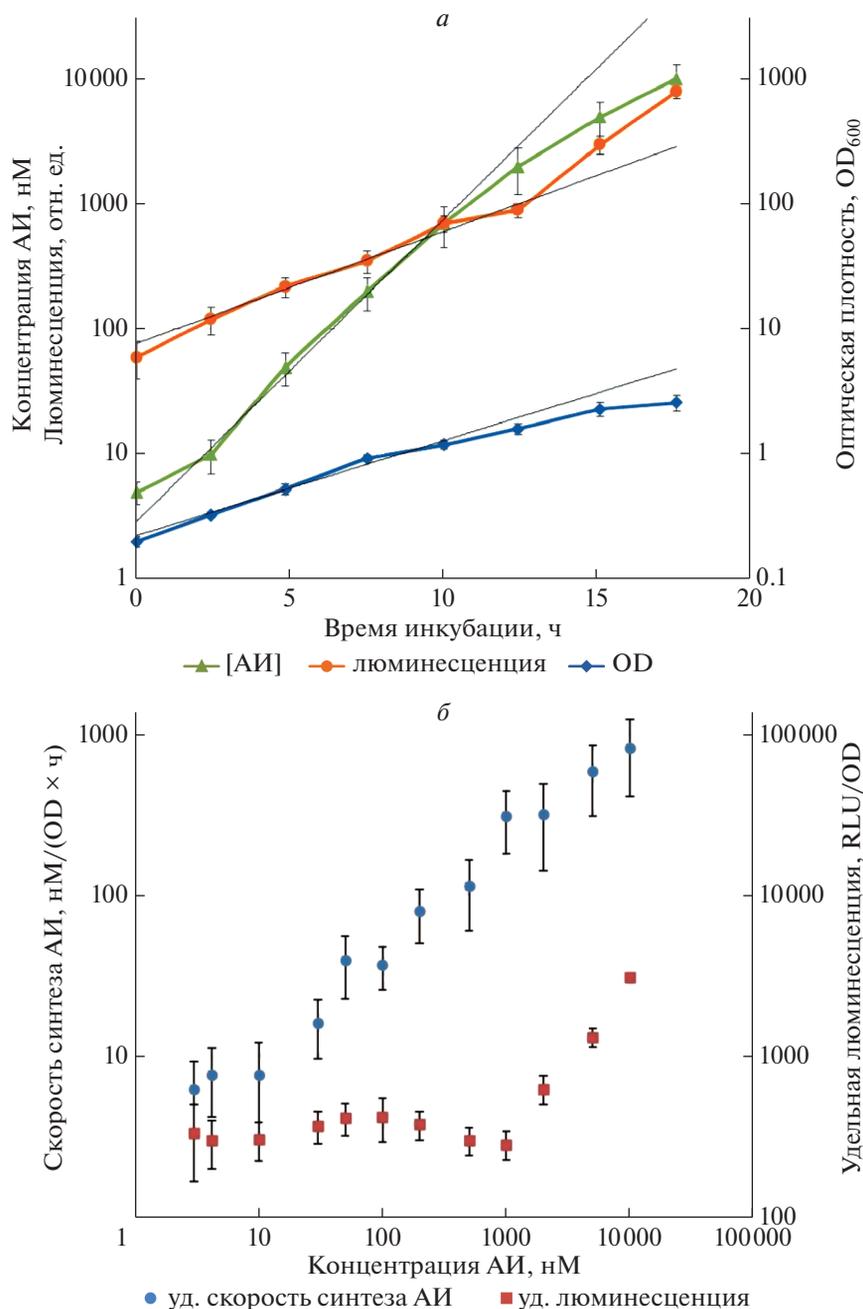
где  $OD_0$  – это оптическая плотность культуры в начале эксперимента, параметр  $a$  определяет темпы роста и зависит от множества факторов эксперимента, таких как штамм, температура, среда, скорость перемешивания и т.п., а  $t$  – это переменная времени. Если пренебречь деградацией АИ и считать, что каждая клетка синтезирует его с одинаковой скоростью в определенном интервале времени, то скорость синтеза АИ культурой клеток в целом будет зависеть от времени также экспоненциально:

$$\frac{d[AИ]}{dt} = COD(t) = COD_0 e^{at}, \quad (5)$$

где  $C$  – это скорость синтеза АИ в расчете на единицу OD в единицу времени. Для получения зависимости концентрации АИ от времени необходимо проинтегрировать уравнение (5):

$$[AI] = [AI]_0 + \frac{1}{a} COD_0 e^{at}, \quad (6)$$

где  $[AI]_0$  – это концентрация АИ в начале эксперимента. Из формул (4) и (6) явным образом следует, что при постоянстве скорости синтеза АИ каждой отдельной клеткой кривые концентрации АИ и оптической плотности культуры должны приближаться экспонентами с одинаковыми показателями  $a$ . Как видно из графика на рис. 1,а, эти кривые аппроксимируются экспонентами со значительно отличающимися показателями, причем показатель экспоненты кривой накопления АИ в среде

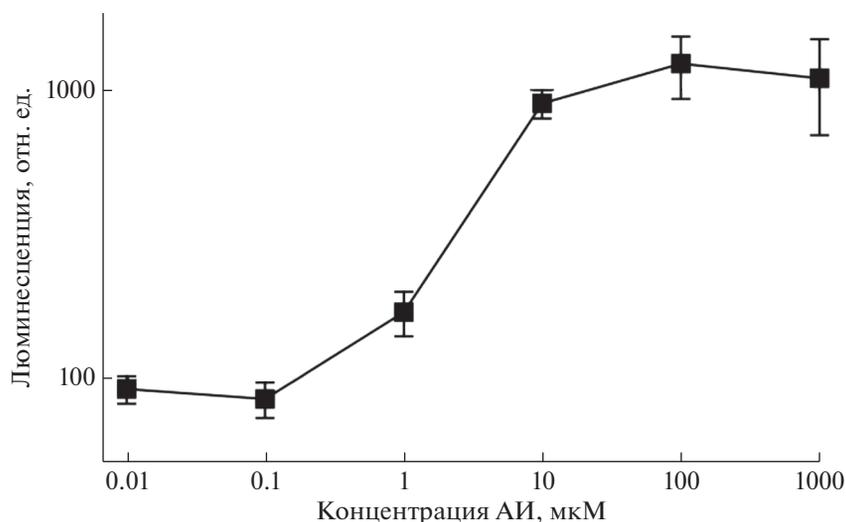


**Рис. 1.** Изменение люминесценции и скорости синтеза АИ в процессе культивирования клеток *A. logei* K18-44 и влияние концентрации АИ на эти процессы. *а* — график зависимости оптической плотности, концентрации АИ и люминесценции культуры клеток *A. logei* K18-44 от времени культивирования. На графике добавлены линии тренда, характеризующие показатель экспоненты для каждой из измеряемых величин; *б* — график изменения удельной светимости и синтеза АИ в расчете на 1 OD в зависимости от концентрации АИ в среде.

выше. Кривые люминесценции и оптической плотности на графике (рис. 1,а) практически параллельны, что означает одинаковую светимость каждой отдельной клетки, до отметки 12.5 ч инкубации. После 12.5 ч наблюдается резкое возрастание люминесценции в расчете на клетку. Вышеописанный резкий рост удельной светимости культуры происходит в позднелогарифмической

фазе роста культуры, когда концентрация АИ достигает значений порядка 1 мкМ.

Удельная скорость прироста концентрации АИ в расчете на клетку ( $\Delta[\text{АИ}]_{\text{уд}}$ ) начинает резко возрастать при концентрации АИ выше 10 нМ (рис. 1,б). Это свидетельствует об активации экспрессии гена *luxI* и росте скорости синтеза АИ в



**Рис. 2.** Зависимость люминесценции культур клеток *A. logei* K18-44 от концентрации экзогенного АИ. Измерение проводилось через 4 ч после добавления АИ.

расчете на одну клетку при концентрациях АИ 10 нМ и выше.

Таким образом, мы наблюдаем двухэтапную активацию QS-системы *A. logei*, где при низких концентрациях активируется промотор гена АИ-синтезазы *luxI* и лишь при высоких концентрациях АИ активируется экспрессия генов люминесценции *luxCDABEG*.

#### Люминесценция и синтез АИ клетками *A. logei* при экзогенном добавлении АИ

Рост люминесценции *A. logei* в зависимости от АИ был проверен с помощью добавления к культуре с неактивированной QS-системой химически синтезированного АИ в различных концентрациях. Была взята культура *A. logei* K18-44, выращенная до  $OD = 0.5$ , отмыта от АИ стерильной средой SWT и затем разделена на аликвоты, к которым добавлялись различные концентрации экзогенного АИ от 10 нМ до 1 мМ. После инкубации в течение четырех часов измерялась люминесценция образцов (рис. 2).

Как видим из данных, приведенных на рис. 2, рост люминесценции начинается только при высоких концентрациях экзогенно внесенного АИ — от 1 мкМ и выше.

Изменение интенсивности синтеза АИ клетками *A. logei* в зависимости от концентрации АИ было определено с помощью добавления к культуре с неактивированной QS-системой химически синтезированного АИ в различных концентрациях. Была взята культура *A. logei* K18-44, выращенная до  $OD = 0.2$ , отмыта от АИ стерильной средой SWT и затем разделена на аликвоты. Измерения АИ в пробе после отмытки и концентрирования

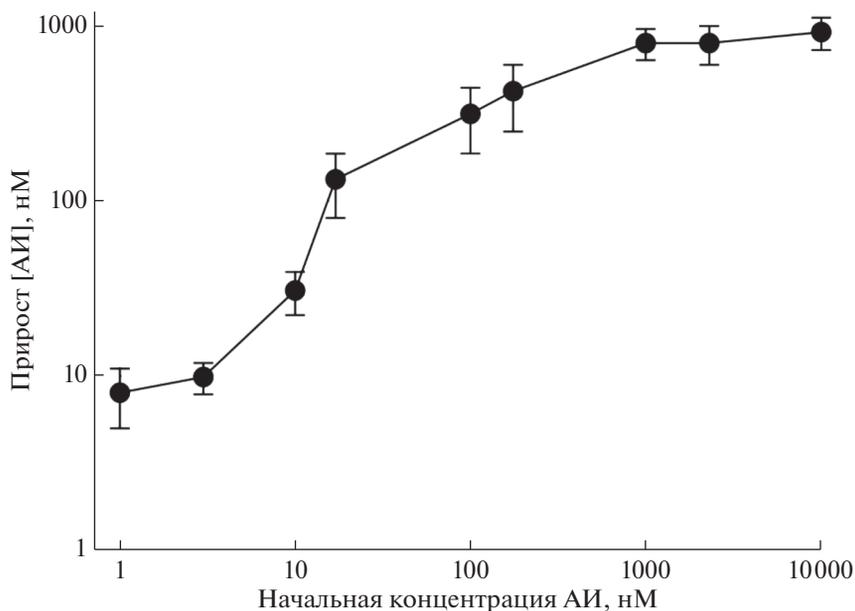
показали, что к началу эксперимента аликвоты содержали не более 1 нМ АИ. К аликвотам добавлялись различные концентрации экзогенного АИ от 1 нМ до 10 мкМ. После инкубации в течение четырех часов измерялась концентрация АИ в образцах (рис. 3).

Как можно видеть по графику на рис. 3, синтез АИ клетками *A. logei* значительно зависит от его концентрации в среде — при концентрациях выше 10 нМ АИ в среде активируется промотор  $P_{luxI}$  и увеличивается экспрессия гена аутоиндуктор-синтезазы *luxI*. При этом в неактивированном состоянии клетки *A. logei* синтезируют АИ со скоростью порядка 10 нМ/( $OD \times \text{час}$ ).

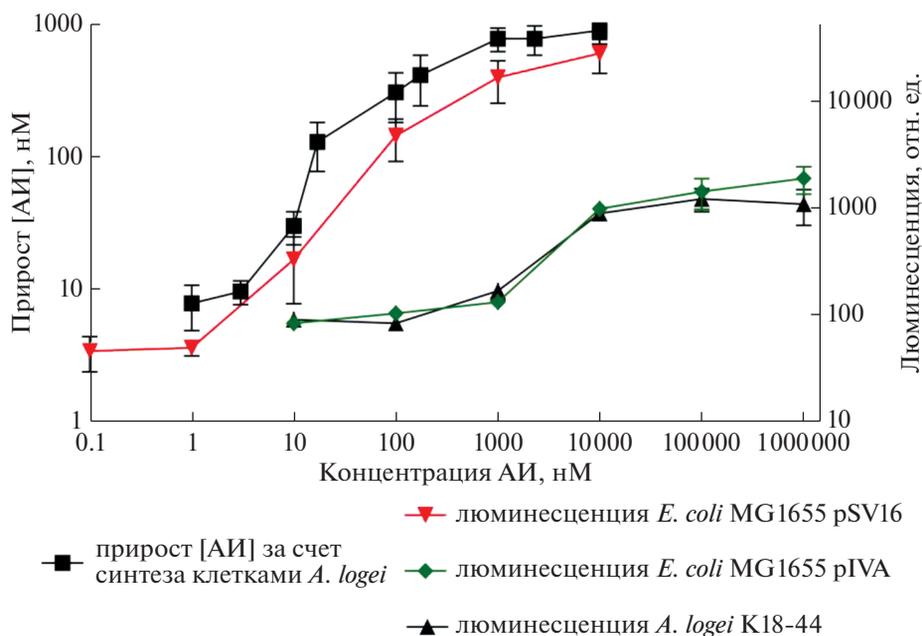
#### Определение АИ-зависимой регуляции $P_{luxI}$ и $P_{luxCDABEG}$ *A. logei* в гетерологичной системе *E. coli*

Для подтверждения различия АИ-зависимой регуляции генов *luxCDABEG* и *luxI* клеток *A. logei* был проведен сравнительный анализ их регуляции в гетерологичной системе клеток *E. coli* (рис. 4). Были использованы клетки *E. coli* MG1655, трансформированные плазмидой pSV16, в которой имеется ген *luxR2*, а гены *luxCDABEG* *P. luminescens* поставлены под контроль  $P_{luxI}$ , или плазмидой pIVA с *luxRI* геном и *luxCDABEG* *P. luminescens* под контролем  $P_{luxCDABEG}$  (табл. 1). Активность промоторов оценивалась по люминесценции соответствующих культур клеток *E. coli*, измеренной через 2 ч после экзогенного добавления 3OC6-HSL.

Результаты определения характера АИ-зависимой активации промоторов  $P_{luxI}$  и  $P_{luxCDABEG}$  *A. logei* в клетках *A. logei* и *E. coli* хорошо согласуются. Очевидны различия исследуемых промоторов по амплитуде активации и по пороговым концентрациям



**Рис. 3.** Зависимость прироста концентрации АИ за четыре часа в культуре клеток *A. logei* от концентрации АИ в начале временного интервала.



**Рис. 4.** Сопоставление результатов исследования АИ-зависимой регуляции промоторов  $P_{luxI}$  и  $P_{luxCDABEG}$  *A. logei* по скорости синтеза АИ и люминесценции клеток *A. logei* и по люминесценции биосенсорных клеток *E. coli* MG1655 pSV16 (промотор  $P_{luxI}$ ) и *E. coli* MG1655 pIVA (промотор  $P_{luxCDABEG}$ ).

АИ.  $P_{luxI}$ , а следовательно и синтез АИ в *A. logei* активируются при концентрациях АИ от 10 нМ.  $P_{luxCDABEG}$  *A. logei* как в клетках *A. logei*, так и в гетерологической системе *E. coli* активируется при концентрациях АИ от 1 мкМ.

*Роль последовательностей lux-боксов в последовательной активации промоторов  $P_{luxI}$  и  $P_{luxCDABEG}$*

Промоторы  $P_{luxI}$  и  $P_{luxCDABEG}$  различаются по последовательности связывания LuxR-белков с

		0	1	2	3	4	5	6	7	8	9	9	8	7	6	5	4	3	2	1	0
AF	<i>lux</i> -бокс	AC	CT	GT	AG	GA	GG	AA	TT	CG	TAC	AG	GG	T							
AL	<i>lux</i> -бокс1	CT	CT	GT	TAA	AG	TT	AT	TAC	AG	GG	T									
AL	<i>lux</i> -бокс2	TC	CT	GT	TA	ATA	TT	GT	TAC	AG	GG	T									

**Рис. 5.** Сравнение последовательностей *lux*-боксов промоторных областей *luxR-luxI A. fischeri* (AF), *luxR1-luxC* и *luxR2-luxI A. logei* (AL). Выделение нуклеотидов: синим – нуклеотиды критичные для связывания LuxR согласно [18]; зеленым – расположенные симметрично в *lux*-боксе; полужирным шрифтом – консервативные позиции.

ДНК (*lux*-бокс2 и *lux*-бокс1 соответственно). Было проведено сравнение последовательностей *lux*-боксов промоторных областей *luxR-luxI A. fischeri*, *luxR1-luxC* и *luxR2-luxI A. logei* (рис. 5).

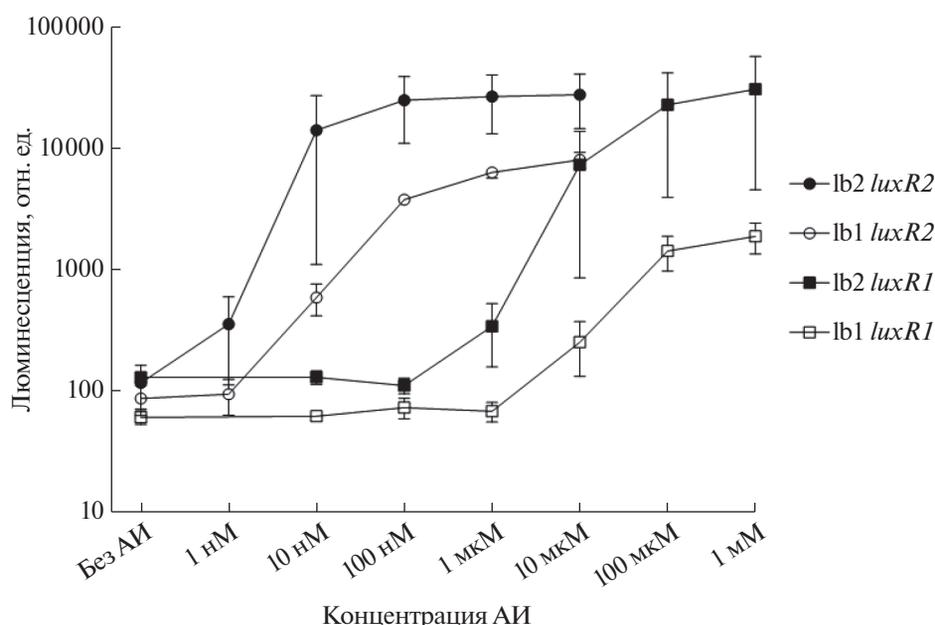
Большинство нуклеотидов в *lux*-боксов-последовательности являются консервативными и совпадают даже в геномах бактерий разных видов (*A. fischeri*, *A. logei* и *A. salmonicida*). Среди консервативных позиций есть описанные в [18] – CTG-CAG-, ключевые для связывания LuxR с сайтом посадки. Также сразу видна разница в степени симметричности разных *lux*-боксов: наиболее симметричный в промоторе  $P_{luxI}$  *A. fischeri*, менее симметричный в промоторе  $P_{luxI}$  *A. logei* и наименее симметричный в промоторе  $P_{luxCDABEG}$  *A. logei*.

Чтобы исследовать роль последовательности *lux*-боксов в наблюдаемых различиях в регуляции промоторов  $P_{luxI}$  и  $P_{luxCDABEG}$ , были сконструиро-

ваны плазмиды pD-lb1 и pD-lb2, в которых гены *luxCDABE P. luminescens* находятся под контролем промотора  $P_{luxCDABEG}$  *A. logei* с разными *lux*-боксов-последовательностями – *lux*-бокс1 и *lux*-бокс2 соответственно (рис. 5). Плазмиды pD-lb1 и pD-lb2 не содержат гены *luxR1* или *luxR2* в отличие от pSV16 и pIVA, эти гены вносились в клетки на отдельных плаزمиды pIV3 и pIV2 (табл. 1) соответственно. Такой подход позволил сравнить специфичность каждого из LuxR-белков к каждому из сайтов связывания и изучить влияние замены *lux*-боксов на регуляцию промотора при прочих равных условиях. Клетки *E. coli* MG1655 трансформировали различными комбинациями плазмид pD-lb1/pD-lb2 и pIV3/pIV2, после чего исследовали зависимость люминесценции полученных клеток от концентрации АИ в среде (рис. 6).

Индукция промотора, содержащего нативный *lux*-боксов промотора  $P_{luxCDABEG}$  (lb1), происходит при больших концентрациях АИ и имеет меньшую амплитуду по сравнению с таковыми для промотора, содержащего *lux*-боксов промотора  $P_{luxI}$  (lb2). Эта закономерность наблюдается в случае с обоими регуляторными генами *luxR1* и *luxR2*, откуда следует, что природная последовательность *lux*-боксов из  $P_{luxCDABEG}$  *A. logei* обладает значительно меньшим сродством к обоим гомологам: LuxR1 и LuxR2.

В этом же эксперименте впервые однозначно было показано, что сам ген *luxR2* обеспечивает значительно большую чувствительность клетки к АИ, чем ген *luxR1*. Этот эффект наблюдается для обеих последовательностей *lux*-боксов в промотор-



**Рис. 6.** Зависимость люминесценции клеток *E. coli* MG1655, несущих гены *luxCDABE P. luminescens* под контролем  $P_{luxCDABE}$  *A. logei* с нативным *lux*-боксом и *lux*-боксом из  $P_{luxI}$  *A. logei* в комбинации с геном *luxR1 A. logei* или *luxR2 A. logei*, от концентрации АИ в среде.

ной области (сравнение кривых с закрашенными символами и кривых с пустыми символами, рис. 6). LuxR2 активирует обе версии промотора при концентрациях АИ 1 и 10 нМ в зависимости от последовательности сайта связывания, а LuxR1 – при концентрациях 1 и 10 мкМ соответственно.

### ОБСУЖДЕНИЕ

Несмотря на то что бактерии *A. logei* были обнаружены и описаны более 40 лет назад [19], а структура их *lux*-оперона известна с 2011 г. [13], двухэтапная активация *lux*-оперона бактерий *A. logei* с последовательной активацией экспрессии генов *luxI* и *luxCDABEG* (рис. 1, б) была впервые показана в настоящей работе. Эти результаты проливают свет на значение отличий в структуре *lux*-оперонов *A. fischeri* и *A. logei*.

Промоторы гена *luxI* у бактерий *A. logei* и *A. fischeri* обладают очень схожими характеристиками по амплитуде индукции и пороговым концентрациям АИ, необходимым для их индукции (порядка 10 нМ), что было показано в гетерологической системе клеток *E. coli* [13]. Но в клетках *A. fischeri* этот промотор регулирует одновременно два процесса – синтез АИ и люминесценцию [20], а у *A. logei* только синтез АИ. При этом активация промотора генов *luxCDABEG A. logei* происходит лишь при высоких концентрациях АИ (порядка 1–10 мкМ, рис. 2). Концентрация АИ 10–20 мкМ в среде является предельной, более высокой концентрации АИ не удавалось достичь при культивировании клеток *A. logei* в жидкой среде SWT при температурах от 4 до 20°C (наши неопубликованные данные), это согласуется с результатами для *A. salmonicida* [4].

Различия в регуляции промоторов генов *luxI* и *luxCDABEG* были описаны в гетерологической системе клеток *E. coli* с использованием биосенсорных плазмид [21] и в клетках *A. logei* по скорости синтеза АИ и люминесценции при различных концентрациях АИ (настоящая работа), а полученные результаты с высокой точностью согласуются друг с другом (рис. 4).

Таким образом, наблюдаются два этапа активации *lux*-оперона: при концентрациях АИ от 10 нМ усиливается его синтез, а при достижении концентрации 1–10 мкМ начинает расти светимость клеток. У психрофильных бактерий стадия роста при промежуточных концентрациях АИ может быть растянута до нескольких суток, в то время как для мезофильных бактерий рост культуры и активация QS-системы происходят быстро, например при 28°C *lux*-оперон *A. fischeri* полностью активируется в течение часа или даже быстрее [14]. Полученные результаты позволяют высказать гипотезу: двухстадийный механизм регуляции QS-системы позволяет психрофильным бактериям экономить ресурсы и не использовать восстановленные эквиваленты для люминесценции клеток в процессе накопления АИ до тех пор,

пока плотность популяции не достигнет значений, обеспечивающих видимость биолоуминесценции невооруженным глазом.

Исследования, проводимые С.В. Баженовым, включающие разработку идеи, постановку основных экспериментов, обработку результатов и написание манускрипта, выполнены при финансовой поддержке РФФИ в рамках научного проекта № 19-34-90020.

Работа И.В. Манухова, включающая разработку идеи, плана экспериментов и интерпретацию результатов, финансировалась в рамках проекта FSMG-2020-0003, соглашение № 075-00337-20-03, заказчик Министерство науки и высшего образования РФ.

Работа В.В. Фомина, включающая эксперименты с различными *lux*-боксами, финансировалась в рамках соглашения № 075-15-2019-1672 от 31.10.2019 г., заказчик Министерство науки и высшего образования РФ.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

1. *Fuqua W.C., Winans S.C., Greenberg E.P.* Quorum sensing in bacteria: The LuxR-LuxI family of cell density-responsive transcriptional regulators // *J. Bacteriol.* 1994. V. 176. № 2. P. 269–275. <https://doi.org/10.1128/jb.176.2.269-275.1994>
2. *Eberhard A., Burlingame A.L., Eberhard C. et al.* Structural identification of autoinducer of *Photobacterium fischeri* luciferase // *Biochemistry.* 1981. V. 20. № 9. P. 2444–2449. <https://doi.org/10.1021/bi00512a013>
3. *Khrulnova S.A., Manukhov I.V., Zarubina A.P., Zavilgelsky G.B.* *Aliivibrio logei* KCh1 (Kamchatka Isolate): Biochemical and bioluminescence characteristics and cloning of the *lux* operon // *Microbiology.* 2010. № 3. P. 349–355. <https://doi.org/10.1134/S0026261710030112>
4. *Hansen H., Purohit A.A., Leiros H.K.S. et al.* The autoinducer synthases LuxI and AinS are responsible for temperature-dependent AHL production in the fish pathogen *Aliivibrio salmonicida* // *BMC Microbiol.* 2015. V. 15. № 1. P. 1–13. <https://doi.org/10.1186/s12866-015-0402-z>
5. *Manukhov I.V., Khrul'nova S.A., Baranova A., Zavilgelsky G.B.* Comparative analysis of the *lux* operons in *Aliivibrio logei* KCh1 (a Kamchatka Isolate) and *Aliivibrio salmonicida* // *J. Bacteriol.* 2011. V. 193. № 15. P. 3998–4001. <https://doi.org/10.1128/JB.05320-11>
6. *Devine J.H., Shadel G.S., Baldwin T.O. et al.* Identification of the operator of the *lux* regulon from the *Vibrio fischeri* strain ATCC7744 (bioluminescence/regulation/autoinduction/repression/activator) // *Proc. Natl Acad. Sci. USA.* 1989. V. 86. P. 5688–5692.

7. *Fidopiastis P.M., Sørum H., Ruby E.G.* Cryptic luminescence in the cold-water fish pathogen *Vibrio salmonicida* // Arch. Microbiol. 1999. V. 171. № 3. P. 205–209. <https://doi.org/10.1007/s002030050700>
8. *Khrulnova S.A., Baranova A., Bazhenov S.V. et al.* Lux-operon of the marine psychrophilic bacterium *Aliivibrio logei*: A comparative analysis of the LuxR1/LuxR2 regulatory activity in *Escherichia coli* cells // Microbiology. 2016. V. 162. № 4. P. 717–724. <https://doi.org/10.1099/mic.0.000253>
9. *Konopleva M.N., Khrulnova S.A., Baranova A. et al.* A combination of *luxR1* and *luxR2* genes activates Pr-promoters of psychrophilic *Aliivibrio logei* lux-operon independently of chaperonin GroEL/ES and protease Lon at high concentrations of autoinducer // Biochem. Biophys. Res. Commun. 2016. V. 473. № 4. P. 1158–1162. <https://doi.org/10.1016/j.bbrc.2016.04.032>
10. *Bazhenov S.V., Khrulnova S.A., Konopleva M.N., Manukhov I.V.* Seasonal changes in luminescent intestinal microflora of the fish inhabiting the Bering and Okhotsk seas // FEMS Microbiol. Lett. 2019. V. 366. № 4. <https://doi.org/10.1093/femsle/fnz040>
11. *Guyer M.S., Reed R.R., Steitz J.A., Low K.B.* Identification of a sex-factor-affinity site in *E. coli* as gamma delta // Cold Spring Harb. Symp. Quant. Biol. 1981. V. 45. Pt 1. P. 135–140. <https://doi.org/10.1101/sqb.1981.045.01.022>
12. *Van Dyk T.K., Rosson R.A.* *Photorhabdus luminescens* luxCDABE promoter probe vectors // Methods Mol. Biol. 1998. V. 102. P. 85–95. <https://doi.org/10.1385/0-89603-520-4:85>
13. *Хрульнова С.А., Манухов И.В., Завильгельский Г.Б.* “Quorum sensing” регуляция экспрессии lux-генов и структура lux-оперона у морских бактерий *Aliivibrio logei* // Генетика. 2011. Т. 47. № 12. С. 1596–1603.
14. *Манухов И.В., Котова В.Ю., Завильгельский Г.Б.* Внутриклеточные факторы регуляции экспрессии lux-оперона *Vibrio fischeri* в клетках *Escherichia coli* // Микробиология. 2006. Т. 75. № 4. С. 525–531.
15. *Melkina O.E., Goryanin I.I., Bazhenov S.V. et al.* Comparative analysis of *Aliivibrio logei* luxR1 and luxR2 genes regulation in *Escherichia coli* cells // Arch. Microbiol. 2019. V. 201. № 10. P. 1415–1425. <https://doi.org/10.1007/s00203-019-01691-3>
16. *Bazhenov S., Novoyatlova U., Scheglova E. et al.* Influence of the luxR regulatory gene dosage and expression level on the sensitivity of the whole-cell biosensor to Acyl-Homoserine Lactone // Biosensors. 2021. V. 11. № 6. P. 166. <https://doi.org/10.3390/bios11060166>
17. *Colton D.M., Stabb E.V., Hagen S.J.* Modeling analysis of signal sensitivity and specificity by *Vibrio fischeri* LuxR variants // PLoS One. 2015. V. 10. № 5. P. 1–21. <https://doi.org/10.1371/journal.pone.0126474>
18. *Antunes L.C.M., Ferreira R.B.R., Lostroh C.P., Greenberg E.P.* A mutational analysis defines *Vibrio fischeri* LuxR binding sites // J. Bacteriol. 2008. V. 190. № 13. P. 4392–4397. <https://doi.org/10.1128/JB.01443-07>
19. *Bang S.S., Baumann P., Nealson K.H.* Phenotypic characterization of *Photobacterium logei* (sp. nov.), a species related to *P. fischeri* // Curr. Microbiol. 1978. V. 1. № 5. P. 285–288. <https://doi.org/10.1007/BF02601683>
20. *Engebrecht J.A., Silverman M.* Identification of genes and gene products necessary for bacterial bioluminescence // Proc. Natl Acad. Sci. USA. 1984. V. 81. № 13. P. 4154–4158. <https://doi.org/10.1073/pnas.81.13.4154>
21. *Манухов И.В.* Структура lux-оперона и механизмы регуляции типа “quorum sensing” у морских бактерий: Дис. ... д-ра биол. наук. 2011. 151 с.

## Two-Stage Activation of Lux-Regulon Psychrophilic Marine Luminescent Bacteria *Aliivibrio logei*

S. V. Bazhenov<sup>a, \*</sup>, E. S. Scheglova<sup>a</sup>, V. V. Fomin<sup>a</sup>, G. B. Zavilgelsky<sup>b</sup>, and I. V. Manukhov<sup>a, b, \*\*</sup>

<sup>a</sup>Moscow Institute of Physics and Technology, Moscow oblast, Dolgoprudny, 141701 Russia

<sup>b</sup>State Research Institute of Genetics and Selection of Industrial Microorganisms of the National Research Center “Kurchatov Institute”, Moscow, 117545 Russia

\*e-mail: bazhenov1994@gmail.com

\*\*e-mail: manukhovi@mail.ru

The lux-regulon of the psychrophilic marine luminescent bacteria *Aliivibrio logei* is controlled by the “quorum sensing” LuxI/LuxR type system. In this work, the properties of the autoinducer-dependent regulation of the promoters of *luxI* and *luxCDABEG* genes were determined both by the measurement of the rate of autoinducer (AI) synthesis and the intensity of bioluminescence of *A. logei* cells and by the reporter genes activity in the heterologous system in *Escherichia coli* cells. The reporter genes were *luxCDABEG* of *Photorhabdus luminescens* under the control of the promoters of the interest from *A. logei*. It was shown that there are differences in the AI-dependent activation of the *luxCDABEG* operon and the separately located *luxI* gene in the *A. logei* cells. The expression of *luxI* is enhanced in the presence of 3-oxo-hexanoyl-homoserine lactone at concentrations of the order of 10 nM, while the expression of *luxCDABEG* genes is enhanced at significantly higher concentrations – 1 μM and higher. This is determined both by the sequence of the binding site of the regulatory proteins lux-box2 and lux-box1 within the respective promoters and by the properties of the two proteins LuxR1 and LuxR2, which differ in their ability to bind the AI.

**Keywords:** quorum sensing, luminescence, autoinducer, regulation, *Aliivibrio logei*, luxR.

## ИЗМЕНЧИВОСТЬ ГЕНОМА ЭНДОСИМБИОТИЧЕСКИХ БАКТЕРИЙ *Wolbachia* В ИНВАЗИВНЫХ ПОПУЛЯЦИЯХ АЗИАТСКОГО ТИГРОВОГО КОМАРА *Aedes albopictus* ЮГА РОССИИ

© 2022 г. Е. А. Коноров<sup>1, 2, \*</sup>, А. Н. Лукашев<sup>3</sup>, Н. Ю. Оюн<sup>3, 4, 5</sup>

<sup>1</sup>Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

<sup>2</sup>Федеральный научный центр пищевых систем им. В.М. Горбатова Российской академии наук, Москва, 109316 Россия

<sup>3</sup>Институт медицинской паразитологии, тропических и трансмиссивных заболеваний им. Е.И. Марциновского Первого Московского государственного медицинского университета им. И.М. Сеченова, Москва, 119435 Россия

<sup>4</sup>Московский государственный университет им. М.В. Ломоносова, Москва, 119991 Россия

<sup>5</sup>Национальный исследовательский центр эпидемиологии и микробиологии имени почетного акад. Н.Ф. Гамалеи, Москва, 123098 Россия

\*e-mail: casqy@yandex.ru

Поступила в редакцию 10.05.2021 г.

После доработки 22.06.2021 г.

Принята к публикации 29.06.2021 г.

Для изучения роли *Wolbachia* в продвижении азиатского тигрового комара *Aedes albopictus* в местообитания с более холодным климатом исследована изменчивость генома бактерий в популяциях комаров Краснодарского края. Было обнаружено, что относительное количество *Wolbachia* у комаров из г. Краснодара выше, чем у комаров из г. Сочи. По результатам филогенетического анализа все представители *Wolbachia* принадлежат к линии wAlbB, для которой ранее отмечалось увеличение приспособленности хозяина в условиях низких температур. Геномы *Wolbachia* из выборок Сочи отличались от геномов из выборок Краснодара по частотам аллелей SNP, расположенных в транспозазах семейства IS982 и обратных транскриптазах интронов группы II.

**Ключевые слова:** эндосимбионты, *Wolbachia*, *Aedes albopictus*, симбиотическая адаптация.

**DOI:** 10.31857/S0016675822020060

Азиатский тигровый комар *Aedes albopictus* является основным переносчиком лихорадки Денге, а также потенциальным переносчиком желтой лихорадки и Чикугуньи. На территории Российской Федерации он был впервые обнаружен в 2011 г. [1] в г. Сочи, с тех пор распространился по побережью [2] и вглубь Краснодарского края [3]. В Краснодарском крае отмечены завозные случаи лихорадки Денге, а в ходе мониторинга в отдельных случаях выявлена РНК возбудителя в комарах [4], однако сообщений о существовании очагов с местной передачей вируса на сегодняшний день нет.

*Wolbachia* – внутриклеточные бактерии, найденные в клетках членистоногих и нематод, чьи взаимоотношения с хозяином варьируют от паразитических до мутуалистических. Их обнаруживают в более чем 65% видов насекомых [5]. *Wolbachia* развили различные способы так называемого репродуктивного паразитизма, такие как цитоплазматическая несовместимость, индукция партеногенеза и феминизация самцов [6] для того, чтобы улучшить передачу бактерии зараженными

самками. Заражение некоторыми линиями паразитических *Wolbachia* было предложено и проверено в полевых условиях в качестве средства борьбы с комарами рода *Aedes*, понижающего продолжительность жизни и жизнеспособность яиц [7, 8]. Заражение *Wolbachia* не критически уменьшает приспособленность комаров, но уменьшает продолжительность их жизни, таким образом сокращая возможности передачи лихорадки Денге, Чикугунья, а также малярии [9, 10].

Как уже было отмечено, *Wolbachia* могут вступать в мутуалистические отношения с насекомым. Для *A. albopictus* было отмечено, что помимо цитоплазматической несовместимости *Wolbachia* может повышать приспособленность самок, повышая их плодовитость [11]. Некоторые линии *Wolbachia* повышают устойчивость своего хозяина к низким температурам [12]. В том числе это показано на комарах из рода *Aedes*, для которых особи, зараженные линией *Wolbachia* wAlbB, демонстрировали большую приспособленность в условиях низких температур по сравнению с незараженными [13].

**Таблица 1.** Результаты картирования чтений на сборку генома *Wolbachia*

	Среднее покрытие	Сайты генома с покрытием, %	Чтения <i>Wolbachia</i> по отношению к общему числу чтений на одну особь <i>Ae. albopictus</i> , %
Sochi_1	154×	97.9	0.012
Sochi_2	131×	95.3	0.018
Sochi_3	51×	96.8	0.005
Krasnodar_4	314×	98.7	0.019
Krasnodar_5	676×	99.8	0.038
Krasnodar_6	638×	99.7	0.039

В комарах *Ae. albopictus*, собранных с побережья Краснодарского края, были найдены ранее две линии *Wolbachia*: wAlbA и wAlbB [14]. На основе данных полногеномного секвенирования комаров из Краснодара и Сочи [15] проведен анализ изменчивости геномов *Wolbachia* с целью обнаружить различия между линиями из локаций с разными климатическими условиями и предположить роль *Wolbachia* в продвижении *Ae. albopictus* в местообитания с более холодным климатом.

## МАТЕРИАЛЫ И МЕТОДЫ

Данные секвенирования шести пулированных образцов комаров *Ae. albopictus* были получены из базы данных NCBI Sequence Read Archive, идентификаторы образцов SRR13570421–SRR13570426 [15]. Образцы представляли собой выделенную геномную ДНК из группы самок имаго (17–42 особей), собранных в трех локациях (удаленных друг от друга более, чем на 10 км) на территории Сочи и в трех локациях на территории Краснодара. Координаты мест сбора выборок: Sochi\_1 – 43.43 N 39.96 E; Sochi\_2 – 43.59 N 39.75 E; Sochi\_3 – 43.68 N 39.77 E; Krasnodar\_4 – 45.02 N 39.03 E; Krasnodar\_5 – 45.00 N 38.93 E; Krasnodar\_6 – 45.08 N 38.93 E.

Адаптеры и нуклеотиды с качеством прочтения меньшим 20, были отфильтрованы с помощью Trimmomatic v.0.38 с параметрами командной строки: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:80 [16]. Отфильтрованные чтения были картированы с помощью BBMap [17] на сборку генома (RefSeq ID: GCF\_004795415.1) *Wolbachia* эндосимбионта *Ae. albopictus* [18] отдельно для образцов из Сочи и Краснодара.

Для филогенетического анализа были отображены картированные чтения на ген поверхностного белка *wsp* *Wolbachia*, далее с помощью программы RegressHaplo [19] были выявлены гаплотипы данного гена и их частоты у *Wolbachia* из популяций Сочи и Краснодара. Обнаруженные гаплотипы и последовательности *wsp*, выделенных из других двукрылых, были выровнены с помощью MAFFT

[20] с параметрами max iterate 1000, localpair. Отдельно для разных позиций в кодоне была подобрана модель замен с помощью PartitionFinder [21], и на основании подобранной модели была реконструирована филогения байесовским методом с помощью Mr.Bayes [22], ngen = 2000000.

Для выявления полиморфных сайтов в геноме *Wolbachia* мы использовали пакет poolstat [23]. Для анализа главных компонент мы использовали ковариационную матрицу, вычисленную с помощью байесовского иерархического моделирования в программе BAYPASS [24].

## РЕЗУЛЬТАТЫ

По результатам картирования среднее покрытие геномов *Wolbachia* каждого из образцов составило 51–676× (табл. 1). Для всех выборок доля позиций в геноме, на которые было картировано хотя бы одно чтение, превышала 95%. В образцах, собранных в Краснодаре, доля чтений *Wolbachia* относительно общего числа чтений (табл. 1) была выше, чем в образцах из Сочи ( $U_{эмп} = 0$ ,  $p < 0.01$ , критерий Манна–Уитни).

Было обнаружено 95 SNP, большинство из них располагалось в транспозазах и в обратных транскриптазах интронов типа II (табл. 2), которые находились на границах скаффолдов сборки GCF\_004795415. Большая часть изменчивости наблюдалась внутри выборок Krasnodar\_5 и Krasnodar\_6, в то время как выборки из Сочи были мономорфны внутри генов транспозаз и обратных транскриптаз.

Изменчивости в генах *cidB* и *cidA* не было обнаружено.

Был обнаружен один общий гаплотип S/K1 гена *wsp*, распространенный в 87.7% геномах *Wolbachia*, эндосимбионтов *Ae. albopictus* из Сочи и в 86.7% Краснодара. Остальные гаплотипы отличались от него делецией одного из кодонов. Частота гаплотипа S2 в выборках Сочи составила 6.66%, гаплотипа S3 – 5.58%. В выборках из Краснодара встречались гаплотипы K2 (8.28%) и K3 (4.96%).

**Таблица 2.** Обнаруженные SNP в генах транспозаз и обратных транскриптаз

Ген	Скаффолд	Описание	Число SNP в данном гене	В каких выборках наблюдается полиморфизм
EJE47_06510	RWIK01000012	Group II intron reverse transcriptase/maturase	2	Krasnodar_5 и Krasnodar_6
EJE47_RS03060	RWIK01000004	IS982 family transposase	2	Все, кроме Sochi_2
EJE47_RS05180	RWIK01000007	IS982 family transposase	10	Krasnodar_5 и Krasnodar_6
EJE47_RS05190	RWIK01000008	Transposase	7	Krasnodar_5 и Krasnodar_6
EJE47_RS06905	RWIK01000015	IS982 family transposase	3	Krasnodar_5 и Krasnodar_6
EJE47_04765	RWIK01000006	IS982 family transposase	9	Krasnodar_5 и Krasnodar_6
EJE47_RS00005	RWIK01000001	Transposase	2	Krasnodar_5 и Krasnodar_6
EJE47_RS03760	RWIK01000005	IS982 family transposase	9	Krasnodar_5 и Krasnodar_6
EJE47_RS03765	RWIK01000005	Group II intron reverse transcriptase/maturase	19	Krasnodar_5 и Krasnodar_6
EJE47_05615	RWIK01000009	IS982 family transposase	4	Все выборки из Краснодара
EJE47_RS05830	RWIK01000010	IS982 family transposase	7	Krasnodar_5 и Krasnodar_6
EJE47_RS06075	RWIK01000010	Group II intron reverse transcriptase/maturase	3	Krasnodar_5 и Krasnodar_6
EJE47_RS06315	RWIK01000011	Group II intron reverse transcriptase/maturase	1	Krasnodar_5 и Krasnodar_6
EJE47_RS07085	RWIK01000017	Helix-turn-helix domain-containing protein	4	Krasnodar_5 и Krasnodar_6

Все перечисленные гаплотипы относились к супергруппе В, то есть к штамму wAlbB (рис. 1).

Анализ главных компонент на основе байесовского иерархического моделирования выявил различия между геномами *Wolbachia* из популяций Сочи и Краснодара (см. рис. 2). Основная часть различий между популяциями, как было отмечено выше, обусловлена различиями в частотах SNP в мобильных элементах.

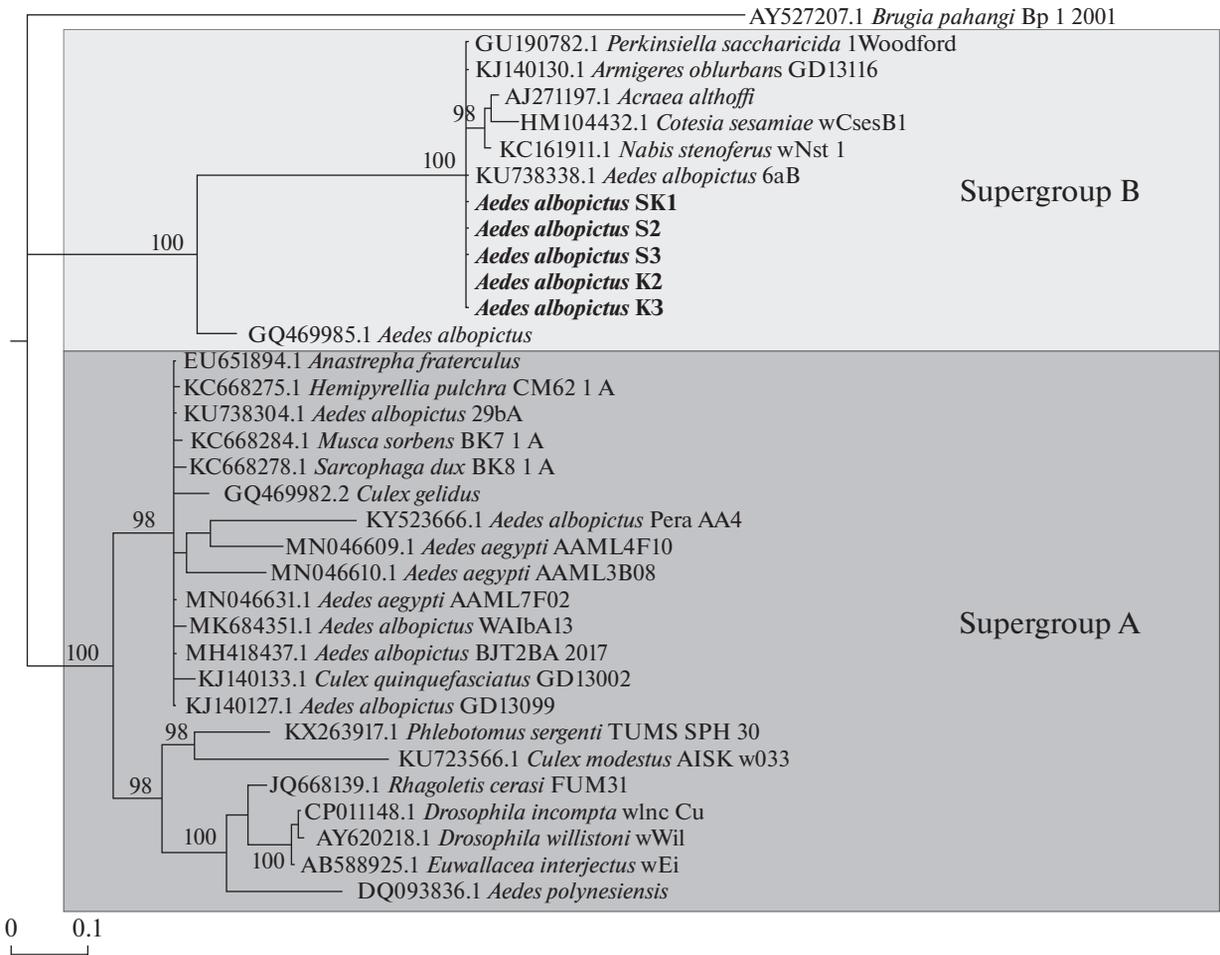
## ОБСУЖДЕНИЕ

В предыдущих исследованиях частоты зараженности *Wolbachia* комаров *Ae. albopictus* с побережья Черного моря России у 56% насекомых был найден хотя бы один из штаммов [25]. В то время как ранее встречались две линии *Wolbachia* (wAlbA и wAlbB), в том числе в 20% случаев обнаруживалось суперзаражение обеими линиями [14], в исследованных нами пулах комаров были обнаружены лишь бактерии wAlbB. Из-за особенностей подхода пулированного секвенирования невозможно оценить точное количество зараженных особей в каждой из выборок, но можно отметить, что относительное количество *Wolbachia* у комаров из Краснодара выше, чем у комаров из Сочи. Подобная ситуация наблюдалась на большем количестве точек сбора и климатических зон в Ки-

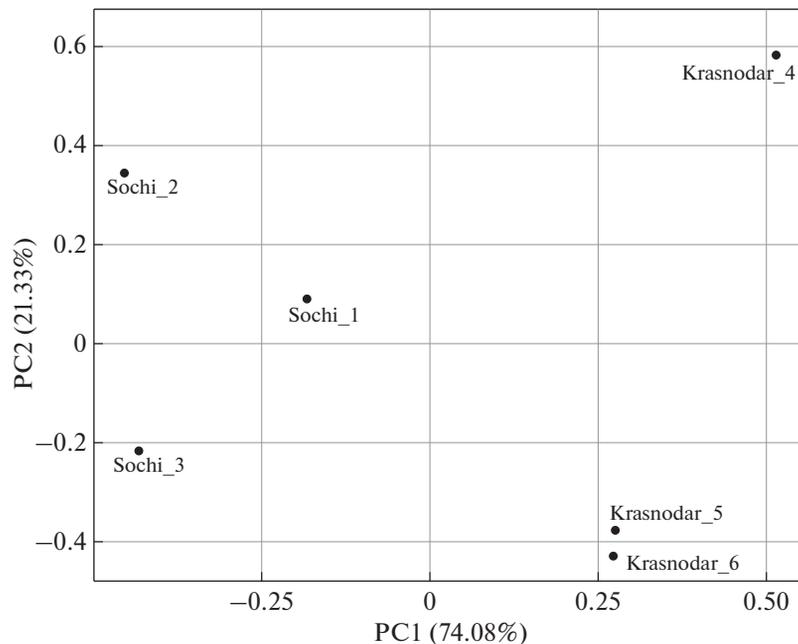
тае, где при продвижении от тропической зоны до умеренной увеличивалась зараженность комаров *Wolbachia* [26]. С одной стороны, данный факт вызывает беспокойство, поскольку представители этого рода, переносчики арбовирусов *Ae. aegypti*, зараженные *Wolbachia* линии wAlbB, обладают большей плодовитостью в условиях более низких температур [13]. Таким образом, увеличение зараженности бактериями может указывать на их возможный вклад в климатическую адаптацию и продвижение *Ae. albopictus* глубже в умеренные широты. С другой стороны, *Wolbachia* ограничивает передачу вируса лихорадки Денге, уменьшая его концентрацию в слюнных железах [27], что показано в том числе и в природных популяциях *Ae. aegypti* [28].

Мы изучили последовательности генов *cidB* и *cidA* *Wolbachia* из *Ae. albopictus*, поскольку было показано, что изменчивость в них коррелирует с фенотипическим разнообразием по цитоплазматической несовместимости у комаров *Culex pipiens* [29]. Ни одного полиморфного сайта в генах *cidB* и *cidA* не обнаружено, поэтому можно предположить, что нет изменчивости по цитоплазматической несовместимости у зараженных *Wolbachia* комаров *Ae. albopictus* из изученных выборок.

Самый распространенный гаплотип гена *wsp* в наших образцах был идентичен гаплотипу из



**Рис. 1.** Байесовское филогенетическое дерево *Wolbachia* двукрылых, построенное по выравниванию последовательностей гена *wsp*. Образцы из данного исследования выделены жирным.



**Рис. 2.** Результаты анализа главных компонент на основании ковариационной матрицы, вычисленной с помощью байесовского иерархического моделирования.

сборки генома *Wolbachia* с GenBank ID CP031221.1, которая была выделена из клеточной линии Aa23 [30]. Остальные гаплотипы не обладали 100% идентичностью ни с одной из последовательностей из GenBank, но также относились к супергруппе В. *Wolbachia* из выборок Сочи существенно отличались от выборок Краснодара по первой главной компоненте (рис. 2). Большая часть этих различий представляет собой разницу в частотах аллелей по SNP, расположенных в транспозазах семейства IS982 и обратных транскриптазах интронов группы II. Известные геномы wAlbV содержат большее количество IS-элементов семейства IS982, чем все остальные геномы *Wolbachia* [18, 30]. Предполагается, что экспансия IS-элементов у *Wolbachia* может быть недавним, возможно продолжающимся до сих пор событием, при этом сопровождающимся горизонтальным переносом от разных линий и псевдогенизацией [31]. Можно предположить, что IS-элементы накапливали замены в ходе перемещения по геному с последующей псевдогенизацией, вызывая изменения в окружающих точку инсерции генах.

В данной работе мы показали, что комары *Ae. albopictus* из Сочи и Краснодара отличаются по степени зараженности *Wolbachia*, а также что геномы *Wolbachia* из этих образцов отличаются по частотам аллелей SNP, по большей части расположенных в IS-элементах. Предполагаем, что дополнительные исследования особей из промежуточных точек сбора между Краснодаром и Сочи позволят подтвердить зависимость частоты заражения *Wolbachia* от среднегодовой температуры.

Работа выполнена при поддержке РФФИ № 19-75-00091 от 09.08.2019.

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Ганушкина Л.А., Таныгина Е.Ю., Безжонова О.В., Сергиев В.П. Об обнаружении комаров *Aedes* (*Stegomyia*) *albopictus* Skus. на территории Российской Федерации // Мед. паразит. и паразитарные болезни. 2012. № 1. С. 3–4.
2. Ganushkina L.A., Patraman I.V., Rezza G. et al. Detection of *Aedes aegypti*, *Aedes albopictus*, and *Aedes koreicus* in the Area of Sochi, Russia // Vector-Borne and Zoonotic Diseases. 2016. V. 16. № 1. P. 58–60. <https://doi.org/10.1089/vbz.2014.1761>
3. Федорова М.В., Швеиц О.Г., Юничева Ю.В. и др. Современные границы распространения инвазивных комаров *Aedes* (*Stegomyia*) *aegypti* (L., 1762) и *Aedes* (*Stegomyia*) *albopictus* (Skuse, 1895) на юге Краснодарского края России // Проблемы особо опасных инфекций. 2018. № 2. С. 101–105. <https://doi.org/10.21055/0370-1069-2018-2-101-105>
4. Гальцева Г.В., Малай В.И., Швеиц О.Г. и др. Завоз геморрагической лихорадки Денге на территорию Краснодарского края // Инфекция и иммунитет. 2017. № 5. С. 415.
5. Hilgenboecker K., Hammerstein P., Schlattmann P. et al. How many species are infected with *Wolbachia*? – a statistical analysis of current data // FEMS Microbiology Letters. 2008. V. 281. № 2. P. 215–220. <https://doi.org/10.1111/j.1574-6968.2008.01110.x>
6. Горячева И.И. Бактерии рода *Wolbachia* – репродуктивные паразиты членистоногих // Усп. совр. биологии. 2004. Т. 124. № 3. С. 246–259.
7. McMeniman C.J., Lane R.V., Cass B.N. et al. Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti* // Science. 2009. V. 323. № 5910. P. 141–144. <https://doi.org/10.1126/science.1165326>
8. Ritchie S.A., Townsend M., Paton C.J. et al. Application of wMel Pop *Wolbachia* strain to crash local populations of *Aedes aegypti* // PLoS Negl. Trop. Dis. 2015. V. 9. № 7. P. e0003930. <https://doi.org/10.1371/journal.pntd.0003930>
9. Moreira L.A., Iturbe-Ormaetxe I., Jeffery J.A. et al. A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and *Plasmodium* // Cell. 2009. V. 139. № 7. P. 1268–1278. <https://doi.org/10.1038/nature10355>
10. Hoffmann A.A., Montgomery B.L., Popovici J. et al. Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission // Nature. 2011. V. 476. № 7361. P. 454–457. <https://doi.org/10.1038/nature10356>
11. Dobson S.L., Rattanadechakul W., Marsland E.J. Fitness advantage and cytoplasmic incompatibility in *Wolbachia* single- and superinfected *Aedes albopictus* // Heredity. 2004. V. 93. № 2. P. 135–142. <https://doi.org/10.1038/sj.hdy.6800458>
12. Mazzucco R., Nolte V., Vijayan T., Schlotterer C. Long-term dynamics among *Wolbachia* strains during thermal adaptation of their *Drosophila melanogaster* hosts // Frontiers in Genetics. 2020. V. 11. P. 482. <https://doi.org/10.3389/fgene.2020.00482>
13. Lau M.J., Ross P.A., Endersby-Harshman N.M., Hoffmann A.A. Impacts of low temperatures on *Wolbachia* (Rickettsiales: Rickettsiaceae)-infected *Aedes aegypti* (Diptera: Culicidae) // J. Med. Entomol. 2020. V. 57. № 5. P. 1567–1574. <https://doi.org/10.1093/jme/tjaa074>
14. Shaikovich E.V., Patraman I.V., Bogacheva A.S. et al. Invasive mosquito species *Aedes albopictus* and *Aedes aegypti* on the Black Sea coast of the Caucasus: genetics (COI, ITS2), *Wolbachia* and *Dirofilaria* infections // Vavilov Journal of Genetics and Breeding. 2018. V. 22. № 5. P. 574–585. <https://doi.org/10.18699/VJ18.397>
15. Konorov E.A., Patraman I.V., Yurchenko V. et al. Role of genetic drift and genomic selection on differentiation and local adaptation of *Aedes albopictus* introduced populations in southern Russia // PeerJ. 2021. V. 9. P. e11776. <https://doi.org/10.7717/peerj.11776>
16. Bolger A.M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data // Bioinformatics. 2014. V. 30. № 15. P. 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

17. *Bushnell B.* BBMap: A Fast, Accurate, Splice-Aware Aligner. Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States), 2014. № LBNL-7065E.
18. *Lefoulon E., Vaisman N., Frydman H.M. et al.* Large enriched fragment targeted sequencing (LEFT-SEQ) applied to capture of *Wolbachia* genomes // *Scientific Reports*. 2019. V. 9. № 1. P. 1–10. <https://doi.org/10.1038/s41598-019-42454-w>
19. *Leviyang S., Griva I., Ita S., Johnson W.E.* A penalized regression approach to haplotype reconstruction of viral populations arising in early HIV/SIV infection // *Bioinformatics*. 2017. V. 33. № 16. P. 2455–2463. <https://doi.org/10.1093/bioinformatics/btx187>
20. *Katoh K., Standley D.M.* MAFFT multiple sequence alignment software version 7: improvements in performance and usability // *Mol. Biol. Evol.* 2013. V. 30. № 4. P. 772–780. <https://doi.org/10.1093/nar/gki198>
21. *Lanfear R., Frandsen P.B., Wright A.M. et al.* Partition-Finder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses // *Mol. Biol. Evol.* 2017. V. 34. № 3. P. 772–773. <https://doi.org/10.1093/molbev/msw260>
22. *Huelsenbeck J.P., Ronquist F.* MRBAYES: Bayesian inference of phylogenetic trees // *Bioinformatics*. 2001. V. 17. № 8. P. 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
23. *Hivert V., Leblois R., Petit E.J. et al.* Measuring genetic differentiation from Pool-seq data // *Genetics*. 2018. V. 210. № 1. P. 315–330. <https://doi.org/10.1534/genetics.118.300900>
24. *Gautier M.* Genome-wide scan for adaptive divergence and association with population-specific covariates // *Genetics*. 2015. V. 201. № 4. P. 1555–1579. <https://doi.org/10.1534/genetics.115.181453>
25. *Shaikevich E., Bogacheva A., Ganushkina L.* *Dirofilaria* and *Wolbachia* in mosquitoes (Diptera: Culicidae) in central European Russia and on the Black Sea coast // *Parasite*. 2019. V. 26. P. 2. <https://doi.org/10.1051/parasite/2019002>
26. *Hu Y., Liu X., Wang J. et al.* Identification and molecular characterization of *Wolbachia* strains in natural populations of *Aedes albopictus* in China // *Parasites & Vectors*. 2020. V. 13. № 1. P. 1–14. <https://doi.org/10.1186/s13071-020-3899-4>
27. *Mousson L., Zouache K., Arias-Goeta C. et al.* The native *Wolbachia* symbionts limit transmission of dengue virus in *Aedes albopictus* // *PLoS Negl. Trop. Dis.* 2012. V. 6. № 12. P. e1989. <https://doi.org/10.1371/journal.pntd.0001989>
28. *Ahmad N.A., Mancini M.V., Ant T.H. et al.* *Wolbachia* strain wAlbB maintains high density and dengue inhibition following introduction into a field population of *Aedes aegypti* // *Philosophical Transactions of the Royal Society B*. 2021. V. 376. № 1818. P. 20190809. <https://doi.org/10.1098/rstb.2019.0809>
29. *Bonneau M., Caputo B., Ligier A. et al.* Variation in *Wolbachia cidB* gene, but not *cidA*, is associated with cytoplasmic incompatibility mod phenotype diversity in *Culex pipiens* // *Mol. Ecol.* 2019. V. 28. № 21. P. 4725–4736. <https://doi.org/10.1111/mec.15252>
30. *Sinha A. Li Z., Sun L., Carlow C.K.* Complete genome sequence of the *Wolbachia* wAlbB endosymbiont of *Aedes albopictus* // *Genome Biology and Evolution*. 2019. V. 11. № 3. P. 706–720. <https://doi.org/10.1093/gbe/evz025>
31. *Cerveau N., Leclercq S., Leroy E. et al.* Short- and long-term evolutionary dynamics of bacterial insertion sequences: insights from *Wolbachia* endosymbionts // *Genome Biology and Evolution*. 2011. V. 3. P. 1175–1186. <https://doi.org/10.1093/gbe/evr096>

## Genome Variation of Endosymbiotic *Wolbachia* in Introduced Populations of Asian Tiger Mosquito *Aedes albopictus* from Krasnodar Krai

E. A. Konorov<sup>a, b, \*</sup>, A. N. Lukashev<sup>c</sup>, and N. Yu. Oyunc<sup>c, d, e</sup>

<sup>a</sup>Vavilov Institute of General Genetics of Russian Academy of Science, Moscow, 119991 Russia

<sup>b</sup>Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences, Moscow, 109316 Russia

<sup>c</sup>Martsinovskiy Institute of Medical Parasitology, Tropical and Vector-Borne Diseases, Sechenov University, Moscow, 119435 Russia

<sup>d</sup>Lomonosov Moscow State University, Moscow, 119991 Russia

<sup>e</sup>Gamaleya National Research Centre for Epidemiology, Moscow, 123098 Russia

\*e-mail: casqy@yandex.ru

For determining the impact of *Wolbachia* endosymbiont on *Aedes albopictus* spread to colder environments we have studied variation in *Wolbachia* genome from introduced Krasnodar krai mosquito populations. We found that mosquitoes from Krasnodar have higher *Wolbachia* infection rate than Sochi mosquitoes. All *Wolbachia* we found belong to wAlbB strain. It was shown in previous studies that this strain increases mosquitoes' egg hatching rate and fertility in low temperature conditions. *Wolbachia* genomes from Sochi and Krasnodar differ on allele frequency spectrum on SNPs in IS982 family transposases and group II intron reverse transcriptases/maturases.

**Keywords:** endosymbionts, *Wolbachia*, *Aedes albopictus*, symbiotic adaptation.

## РАСПРОСТРАНЕНИЕ АССОЦИИРОВАННЫХ С БОЛЕЗНЬЮ ХРОНИЧЕСКОГО ИЗНУРЕНИЯ АЛЛЕЛЕЙ ГЕНА *PRNP* У ДИКИХ И ДОМАШНИХ СЕВЕРНЫХ ОЛЕНЕЙ *Rangifer tarandus* НА ТЕРРИТОРИИ РОССИИ

© 2022 г. К. А. Курбаков<sup>1, 2, \*</sup>, Е. А. Коноров<sup>1, 2</sup>, М. Т. Семина<sup>1</sup>, Ю. А. Столповский<sup>1</sup>

<sup>1</sup>Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

<sup>2</sup>Федеральный научный центр пищевых систем им. В.М. Горбатова Российской академии наук, Москва, 109316 Россия

\*e-mail: homo\_ludens@vniimp.ru

Поступила в редакцию 09.04.2021 г.

После доработки 30.09.2021 г.

Принята к публикации 01.10.2021 г.

Болезнь хронического изнурения – это прионное заболевание, встречающееся у представителей семейства оленьих. Аминокислотные замены Ser138, Ser225 и 2M-129S-169M гена прионного белка *PRNP* увеличивают риск заболевания. В работе мы приводим данные о вариабельности участка гена *PRNP* с 108-ого по 240-ой кодон у *Rangifer tarandus* на территории азиатской части России и Ненецкого АО. Обнаружено семь полиморфных сайтов с несинонимичными заменами, пять из которых наблюдались ранее у оленей из Европы. Более резистентный к болезни хронического изнурения вариант Asp138 встречается примерно в четверти гаплотипов домашних оленей, другой устойчивый вариант Trp225 встретился лишь у небольшой доли оленей с Чукотки. Дикие особи отличались от домашних по частотам аллелей с заменами Val169 и Gly129, аллель предрасположенности 2M-129S-169M встречался, в основном, у одомашненных особей. Из-за большой изменчивости гена *PRNP* у северного оленя России в случае возникновения у него болезни хронического изнурения возможно повторение ситуации, наблюдающейся в Норвегии – появление разных штаммов с отличающимися клиническими симптомами и проявлением болезни.

**Ключевые слова:** *Rangifer tarandus*, болезнь хронического изнурения, прионы.

**DOI:** 10.31857/S0016675822020102

Ген *PRNP* кодирует прионный белок PrP, который имеет две конформационные изоформы. Инфекционная изоформа способна менять конформацию нормальной клеточной, накапливаясь в виде резистентного для протеаз агрегата в нервной ткани [1]. У разных видов млекопитающих данный процесс приводит к развитию трансмиссивных губчатых энцефалопатий, таких как болезнь Крейтцфельда–Якоба у человека, коровье бешенство или губчатая энцефалопатия КРС, скрепи у овец и коз, болезнь хронического изнурения у оленьих. Передача инфекционной формы может происходить даже между разными видами животных, например, возможна передача коровьего бешенства от КРС человеку при употреблении в пищу зараженных животных. Случаи передачи болезни хронического изнурения и скрепи от животных к человеку не были отмечены, скорее всего, из-за особенностей последовательности и структуры PrP человека, затрудняющие конформационный переход [2, 3]. Таким образом, риск передачи болезни хронического

изнурения от оленьих к человеку довольно низок, хотя нельзя исключать возможность адаптации прионов к новому хозяину или обнаружения гаплотипов *PRNP*, имеющих большую способность к трансмиссии [4].

Распространение и изменчивость *PRNP* изучена для разных представителей семейства Cervidae: для европейского лося, косули, лани [5], белохвостого [6, 7], благородного, пятнистого [7] и северного оленей [5, 8]. Для северного оленя прослежена также ассоциация отдельных гаплотипов с частотой болезни хронического изнурения [9], например, распространенный аллель Ser225 и делеция Trp84\_Gly91del увеличивает риск развития болезни в пять раз в гетерозиготе и в более чем в 40 раз, находясь в гомозиготном состоянии. Гомозиготы Trp225, наоборот, были обнаружены только у здоровых животных. Поиск ассоциаций был проведен также для белохвостого оленя *Odocoileus virginianus*, где эффект от замены Ser225 не был статистически значимым, однако обнаружена повышенная устойчивость к болезни при на-

**Таблица 1.** Объемы выборок, места сбора и показатели генетического разнообразия

Выборка	Объем выборки	Место сбора материала	$\pi \times 10^3$	<i>D</i> Таджимы
Чукотские, дикая популяция	17	О-в Айон, Певек	4.26	-0.66
Эвенкийские (тоджинские)	18	Тува, Тоджинский р-н	3.68	0.023
Чукотка, домашняя порода	16	П-ов Камчатка	4.54	0.66
Эвенкийские	20	Амурская обл., Зейский р-он	3.65	-0.29
Ненецкие	18	СПК Коопхоз "ЕРВ" (племярепродуктор) НАО, западная часть Большеземельной тундры Ненецкого АО	3.14	-0.33

личии несинонимической замены Ser95 [10]. Замена на аспарагин в 138-ом кодоне также предположительно связана с резистентностью [11].

По вышеперечисленным позициям в природных и одомашненных популяциях наблюдается изменчивость, а в некоторых случаях дифференциация. В Северной Америке отдельные локальные популяции северных оленей демонстрируют избыток гетерозигот гомозигот с Asn138, что может быть результатом действия балансирующего отбора или генетического дрейфа [12, 13]. На севере европейской части России было описано распространение вариантов гена *PRNP* у диких и одомашненных популяций северного оленя, при этом отмечалось более высокая доля аллелей с заменой Asn138 у домашних оленей по сравнению с дикими [8].

В настоящей работе мы приводим данные о вариативности участка гена *PRNP* у *Rangifer tarandus* с 108-ого по 240-ой кодон на территории азиатской части России и Ненецкого АО. На данном участке расположены основные нуклеотидные замены, ассоциированные с риском болезни хронического изнурения. Также мы провели поиск отличий между дикими и домашними оленями.

## МАТЕРИАЛЫ И МЕТОДЫ

Всего исследовано 89 оленей, из них 72 особи: ненецкой, эвенкийской, чукотской пород и 17 диких оленей (п-ов Чукотка) (табл. 1). Образцы дикой популяции были представлены мышечной тканью, домашних животных – ушным выщипом. ДНК выделялось с использованием набора QIAmp DNA mini Kit (QIAGEN, Хильден, Германия). Амплификация гена *PRNP* происходила с помощью qPCRmix-HS (Евроген, Москва, Россия) и праймеров RangPrP-F AGTCAGTGGAA-CAAGCCCAG и RangPrP-R TGAGGAGGATCA-CAGGAGGG, подобранных с помощью Primer-BLAST [14] и OligoAnalyzer v. 3 (<http://www.idtd->

[na.com/analyzer/applications/oligoanalyzer/](http://www.idtdna.com/analyzer/applications/oligoanalyzer/)). Длина амплифицируемого фрагмента составила 442 пар оснований (участок с 298 по 739 нуклеотид гена *PRNP*, GenBank – ID MK097270). Режим реакции ПЦР: предварительная денатурация при 95°C, 5 мин и 45 циклов амплификации (95°C, 15 с, 60°C, 20 с и 72°C 15 с). Секвенирование по Сенгеру полученных апликнов проводили в ЗАО "Евроген" (Евроген, Москва, Россия). Секвенирование проводили с использованием как форвард-, так и реверс-прайма. Анализ полученных хроматограмм проводили с использованием программы UGENE [15]. Обнаруженные аллельные варианты были депонированы в базе данных GenBank под идентификационными номерами MZ442704–MZ442716.

Выравнивание последовательностей гена *PRNP* с помощью Muscle [16] и тест Таджимы на нейтральность были проведены в пакете MEGAX [15]. TCS-сеть аллелей гена *PRNP* была построена в программе Popart [17]. Анализ главных компонент производился с помощью предустановленной в язык R функции `prcomp`.

## РЕЗУЛЬТАТЫ

В процессе исследования было получено 89 последовательностей гена *PRNP* северного оленя длиной 399 пн, начиная с 108-ого кодона и заканчивая 240-ым. Согласно показателям генетического разнообразия и *D* Таджимы, выборка домашних чукотских оленей единственная из изученных нами демонстрирует тенденцию на увеличение гетерозиготности (табл. 1, рис. 1).

Обнаружено семь полиморфных сайтов, все они несинонимичные. В кодоне 135 найдена новая для северных оленей замена Arg135, встретившаяся у двух диких особей чукотского оленя. Ранее не встречавшаяся замена Ser184 наблюдается в гетерозиготном состоянии в эвенкийской и ненецкой популяциях, по одной особи в каждой. Для остальных полиморфных сайтов наблюда-

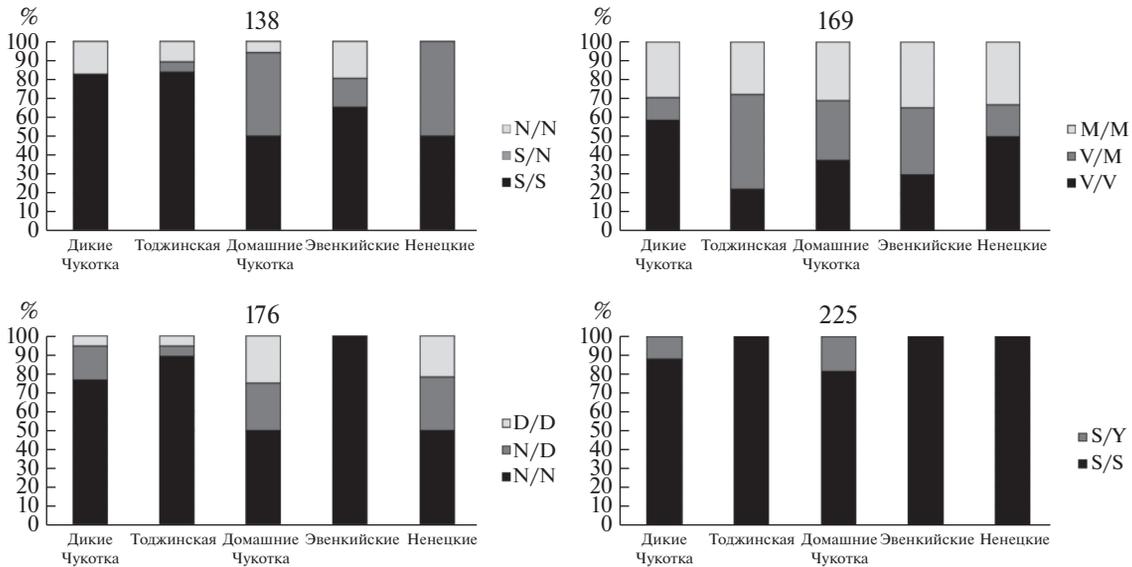


Рис. 1. Частоты генотипов для наиболее варибельных кодонов гена PRNP у особей R. tarandus, изученных в данной работе.

лось большее генетическое разнообразие (рис. 1, табл. 2). Связанный с резистентностью к болезни хронического изнурения вариант Asn138 встречается примерно в четверти гаплотипов чукотских, эвенкийских и ненецких пород оленей, в два раза меньше его у оленей тоджинской породы. У диких оленей из Чукотки Asn138 обнаружен только в гомозиготном состоянии, как и у популяций диких оленей с Европейского Севера России [8]. Доля резистентного варианта ниже по сравнению с выборками домашних оленей, исключая тоджинских.

У диких северных оленей с Чукотки и домашних оленей чукотской породы присутствуют аллели с Tyr225 (аллель В и аллель Asp176\_Tyr225), у остальных изученных в настоящей работе выборок подобный вариант отсутствует даже в гетеро-

зиготном состоянии. У оленей с Европейского Севера данные аллели наблюдались почти исключительно у диких, но не домашних особей [8].

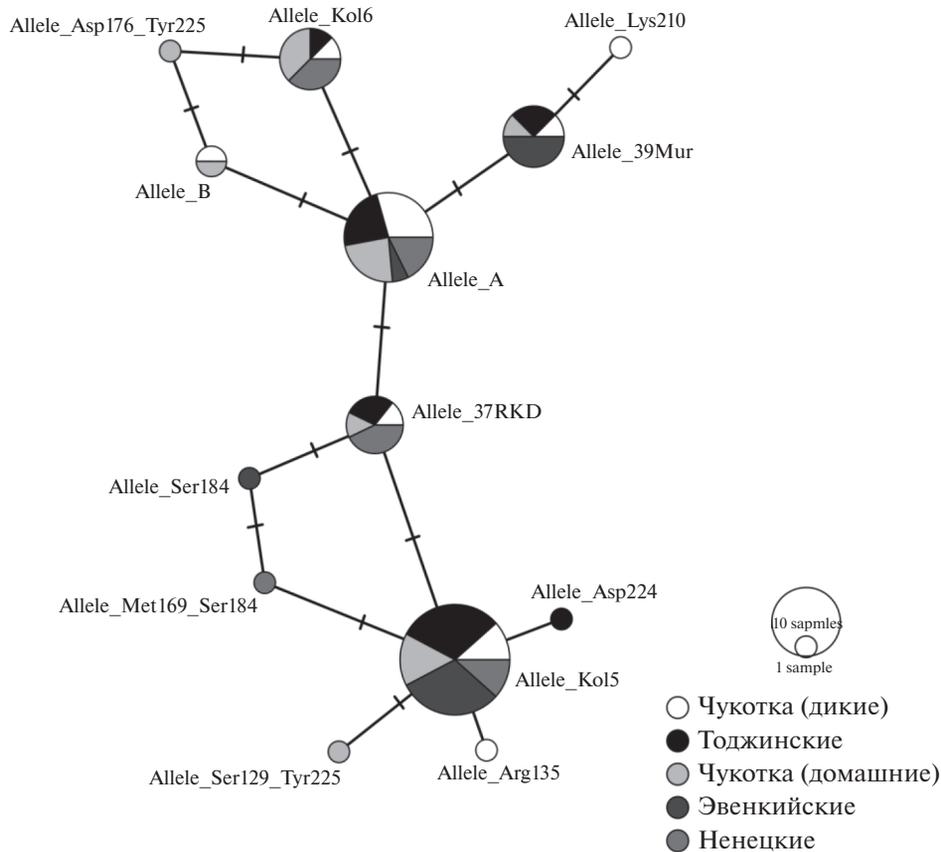
На рис. 2 приведены взаимоотношения между найденными в ходе исследования аллелями в виде TCS-сети. Наиболее распространенными аллелями в изученной выборке являются аллель, идентичный на рассмотренном участке аллелю А (Ser225, SH-SY5Y rePrP<sup>wt</sup>, [9]) и аллель Kol5 (MSSM, DQ154295, 2M-129S-169M). Реже встречались (у восьми–деяти особей) аллели 39Mur (VGNV, DQ154292), RKD37 (VSSV, DQ154294) и Kol6 (176D, JQ290075).

Анализ главных компонент на основании частот аллелей гена PRNP в разных выборках оленей с территории России выявил дифференциацию диких особей с европейской части России от

Таблица 2. Частоты аминокислотных замен в наиболее варибельных кодонах гена PRNP у R. tarandus на территории России

Выборка	129		138		169		176		225		
	G	S	S	N	V	M	N	D	S	Y	
Тоджинский Чукотка, домашние	0.35	0.65	0.86	0.14	0.47	0.53	0.92	0.08	1.00	0.00	
Эвенкийские	0.63	0.38	0.72	0.28	0.53	0.47	0.63	0.38	0.91	0.09	
Ненецкие	0.33	0.67	0.73	0.28	0.48	0.53	1.00	0.00	1.00	0.00	
Республика Коми	0.44	0.56	0.75	0.25	0.58	0.42	0.64	0.36	1.00	0.00	*
О-в Колгуев	0.53	0.47	0.72	0.28	0.78	0.22	0.94	0.06	1.00	0.00	*
Новая Земля	0.53	0.47	0.63	0.37	0.45	0.55	0.76	0.24	1.00	0.00	*
Кольский п-ов	0.85	0.15	0.91	0.09	0.87	0.13	0.89	0.11	0.98	0.02	*
Дикие: Архангельск, Республика Коми	0.89	0.11	0.89	0.11	0.93	0.07	1.00	0.00	0.86	0.14	*
Чукотка, дикие	1.00	0.00	0.82	0.18	1.00	0.00	1.00	0.00	0.86	0.14	*
	0.63	0.38	0.82	0.18	0.65	0.35	0.85	0.15	0.94	0.06	

Примечание. Верхняя строка в шапке – позиция кодона, \* – данные для выборок, взяты из исследования [8] для сравнения.



**Рис. 2.** TCS-сеть обнаруженных у исследуемых особей аллелей *PRNP*. Обозначения известных аллелей взяты из работ Kholodova et al. [8] и Güere et al. [9].

всех домашних оленей (рис. 3). Основным источником данной дифференциации являются аллели Val169 и Gly129, которые у диких оленей представлены в большей концентрации. Исключением является выборка диких оленей с п-ова Чукотка, которая близка по частотам Gly129 к оленям чукотской породы и не является статистическим выбросом из выборок домашних оленей по частоте Val169 (согласно тесту Диксона).

Выборки домашних оленей при этом также разбиваются на две группы: 1) с полным преобладанием Asn176 – тоджинская, эвенкийская породы и олени из выборки Коми; 2) со значительной долей гомо- и гетерозигот (до 50%) по замене Asp176 – олени ненецкой и чукотской пород и популяции с о-ва Колгуев.

## ОБСУЖДЕНИЕ

Случаи болезни хронического истощения северного оленя наблюдаются в соседних с Россией странах – Норвегии и Финляндии, с 2016 г. Как и в случае с другими прионными болезнями, в большинстве случаев предотвратить распространение инфекции возможно лишь с помощью уничтожения целых локальных популяций зара-

женных животных [19]. Болезнь трудно диагностировать у животных из-за неспецифичных симптомов, выражающихся в изменении поведения и потере веса, развивающихся за период от нескольких недель до месяцев [20]. На территории США болезнь распространялась со скоростью большей, чем почти все остальные известные прионные заболевания. На территории Норвегии, Швеции и Финляндии распространение было медленнее, но полностью остановить его не получилось несмотря на жесткие меры [21]. Появление болезни на территории России потребует срочных мер, внедрить которые будет сложно из-за большой территории и традиционного образа оленеводства различных народов Европейского Севера, Сибири и Дальнего Востока.

У исследованных на данный момент популяций одомашненного и дикого оленей, обитающих на территории России распространены аллели, ассоциированные ранее с предрасположенностью к болезни хронического истощения. У всех особей, генотипированных в ходе данного исследования, наблюдается замена Ser225, лишь у домашних и диких оленей из Чукотки обнаружен более резистентный вариант Tyr225 в гетерозиготном состоянии. Аллель Kol5 (2M-129S-169M),

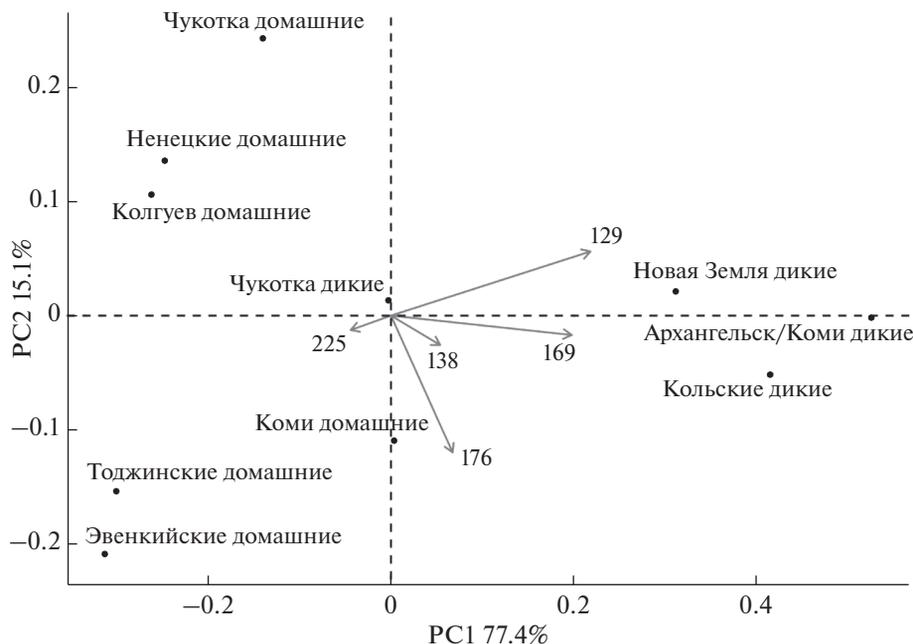


Рис. 3. Результаты анализа главных компонент (оси абсцисс и ординат) на основании частот аллелей гена *PRNP* в разных выборках северного оленя на территории России.

для которого в ходе анализа ассоциации также была обнаружена предрасположенность к болезни [9], встречается в основном у одомашненных оленей, как из европейской, так и азиатской части России, практически отсутствует у диких оленей Европейского Севера России.

В более старом очаге болезни хронического истощения оленьих – Северной Америке – предрасположенность связывали в основном с заменой Ser138 [12, 13], полиморфизма в 129-ом и 225-ом кодонах не наблюдалось. В то же время там, где болезнь впервые появилась в Европе, в Норвегии, нет полиморфизма в 138-ом кодоне, а ассоциация наблюдается с заменами Ser225 и Ser129 [9]. У северного оленя на территории России наблюдаются выявленные как европейскими, так и североамериканскими авторами факторы предрасположенности к болезни, такие как Ser138, так и Ser225 с Ser129. В ходе сравнения течения болезней у европейских и североамериканских оленей недавние исследования показали, что из-за разной конформации инфекционной изоформы белка PrP, кодируемого геном *PRNP*, штаммы прионов из Норвегии и Северной Америки обладают разными трансмиссионными и нейропатологическими паттернами; причем внутри норвежских штаммов также существует вариативность в патологической картине [22]. Из-за большой изменчивости гена *PRNP* у северного оленя России в случае возникновения у него болезни хронического истощения возможна ситуация, повторяющая такую же в Норвегии: большое разнообразие клинических картин болезни или появление нового, специфичного для российских популяций штам-

ма, что затруднит использование накопленного опыта борьбы с болезнью хронического истощения.

Работа выполнена при поддержке гранта РФФИ № 17-29-08003.

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Houston F., Andréoletti O. Animal prion diseases: the risks to human health // *Brain Pathology*. 2019. V. 29. № 2. P. 248–262. <https://doi.org/10.1111/bpa.12696>
2. Kurt T.D., Jiang L., Fernández-Borges N. et al. Human prion protein sequence elements impede cross-species chronic wasting disease transmission // *J. Clin. Investigation*. 2015. V. 125. № 4. P. 1485–1496. <https://doi.org/10.1172/JCI79408>
3. Concepcion G.P., David M.P. C., Padlan E.A. Why don't humans get scrapie from eating sheep? A possible explanation based on secondary structure predictions // *Med. Hypotheses*. 2005. V. 64. № 5. P. 919–924. <https://doi.org/10.1016/j.mehy.2004.11.021>
4. Hannaoui S., Schatzl H.M., Gilch S. Chronic wasting disease: Emerging prions and their potential risk // *PLoS Pathogens*. 2017. V. 13. № 11. P. e1006619. <https://doi.org/10.1371/journal.ppat.1006619>
5. Wik L., Mikko S., Klingeborn M. et al. Polymorphisms and variants in the prion protein sequence of European moose (*Alces alces*), reindeer (*Rangifer tarandus*), roe deer (*Capreolus capreolus*) and fallow deer (*Dama dama*)

- in Scandinavia // *Prion*. 2012. V. 6. № 3. P. 256–260. <https://doi.org/10.4161/pri.19641>
6. *Chafin T.K., Douglas M.R., Martin B.T. et al.* Age structuring and spatial heterogeneity in prion protein gene (*PRNP*) polymorphism in white-tailed deer // *Prion*. 2020. V. 14. № 1. P. 238–248. <https://doi.org/10.1080/19336896.2020.1832947>
  7. *Robinson S.J., Samuel M.D., O'Rourke K.I., Johnson C.J.* The role of genetics in chronic wasting disease of North American cervids // *Prion*. 2012. V. 6. № 2. P. 153–162. <https://doi.org/10.4161/pri.19640>
  8. *Kholodova M.V., Baranova A.I., Mizin I.A. et al.* A genetic predisposition to chronic wasting disease in the reindeer *Rangifer tarandus* in the Northern European Part of Russia // *Biol. Bulletin*. 2019. V. 46. № 6. P. 555–561. <https://doi.org/10.1134/S1062359019060074>
  9. *Güere M.E., Våge J., Tharaldsen H. et al.* Chronic wasting disease associated with prion protein gene (*PRNP*) variation in Norwegian wild reindeer (*Rangifer tarandus*) // *Prion*. 2020. V. 14. № 1. P. 1–10. <https://doi.org/10.1080/19336896.2019.1702446>
  10. *Rhyan J.C., Miller M.W., Spraker T.R. et al.* Failure of fallow deer (*Dama dama*) to develop chronic wasting disease when exposed to a contaminated environment and infected mule deer (*Odocoileus hemionus*) // *J. Wildlife Diseases*. 2011. V. 47. № 3. P. 739–744. <https://doi.org/10.7589/0090-3558-47.3.739>
  11. *Mitchell G.B., Sigurdson C.J., O'Rourke K.I. et al.* Experimental oral transmission of chronic wasting disease to reindeer (*Rangifer tarandus tarandus*) // *PLoS One*. 2012. V. 7. № 6. P. e39055. <https://doi.org/10.1371/journal.pone.0039055>
  12. *Cheng Y.C., Musiani M., Cavedon M., Gilch S.* High prevalence of prion protein genotype associated with resistance to chronic wasting disease in one Alberta woodland caribou population // *Prion*. 2017. V. 11. № 2. P. 136–142. <https://doi.org/10.1080/19336896.2017.1300741>
  13. *Happ G.M., Huson H.J., Beckmen K.B., Kennedy L.J.* Prion protein genes in caribou from Alaska // *J. Wildlife Diseases*. 2007. V. 43. № 2. P. 224–228. <https://doi.org/10.7589/0090-3558-43.2.224>
  14. *Ye J., Coulouris G., Zaretskaya I. et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction // *BMC Bioinformatics*. 2012. V. 13. № 1. P. 1–11. <https://doi.org/10.1186/1471-2105-13-134>
  15. *Okonechnikov K., Golosova O., Fursov M.* Ugene Team. Unipro UGENE: A unified bioinformatics toolkit // *Bioinformatics*. 2012. V. 28. № 8. P. 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>
  16. *Edgar R.C.* MUSCLE: Multiple sequence alignment with high accuracy and high throughput // *Nucl. Ac. Res.* 2004. V. 32. № 5. P. 1792–1797. <https://doi.org/10.1093/nar/gkh340>
  17. *Kumar S., Stecher G., Li M., Knyaz C., Tamura K.* MEGA X: Molecular evolutionary genetics analysis across computing platforms // *Mol. Biol. and Evol.* 2018. V. 35. № 6. P. 1547–1549. <https://doi.org/10.1093/molbev/msy096>
  18. *Leigh J.W., Bryant D.* Popart: full-feature software for haplotype network construction // *Met. in Ecol. and Evol.* 2015. V. 6. № 9. P. 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
  19. *Mysterud A., Rolandsen C.M.* A reindeer cull to prevent chronic wasting disease in Europe // *Nat. Ecol. & Evol.* 2018. V. 2. № 9. P. 1343–1345. <https://doi.org/10.1038/s41559-018-0616-1>
  20. *Benestad S.L., Telling G.C.* Chronic wasting disease: An evolving prion disease of cervids // *Handbook of Clin. Neurology*. 2018. V. 153. P. 135–151. <https://doi.org/10.1016/B978-0-444-63945-5.00008-8>
  21. *Maraud S., Roturier S.* Chronic wasting disease (CWD) in sami reindeer herding: The socio-political dimension of an epizootic in an indigenous context // *Animals*. 2021. V. 11. № 2. P. 297. <https://doi.org/10.3390/ani11020297>
  22. *Nonno R., Di Bari M.A., Pirisinu L. et al.* Studies in bank voles reveal strain differences between chronic wasting disease prions from Norway and North America // *PNAS*. 2020. V. 117. № 49. P. 31417–31426. <https://doi.org/10.1073/pnas.2013237117>

## Associated with Chronic Wasting Disease Alleles of *PRNP* Gene Distribution in Wild and Domesticated Reindeer *Rangifer tarandus* in Russia

K. A. Kurbakov<sup>a, b, \*</sup>, E. A. Konorov<sup>a, b</sup>, M. T. Semina<sup>a</sup>, and Yu. A. Stolpovsky<sup>a</sup>

<sup>a</sup>Vavilov Institute of General Genetics Russian Academy of Science, Moscow, 119991 Russia

<sup>b</sup>Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences, Moscow, 109316 Russia

\*e-mail: homo\_ludens@vniimp.ru

Chronic wasting disease (CWD) is a prion disease that affects cervids. Amino acid substitutions Ser138, Ser225 and 2M-129S-169M in prion protein coded by *PRNP* gene increase risk of the disease. In this study we present *PRNP* variability data in reindeer from Asian part of Russia. We discovered seven polymorphic codons with non-synonymous substitutions, polymorphism for five of them were reported before in European reindeer populations. Resistant to CWD variant Asn138 was observed in one quarter of *PRNP* haplotypes in domesticated reindeers, another resistance-associated variant, Tyr225, was found only in a small group of reindeers from Chukotka. Wild deers differentiate from domesticated ones in Val169 and Gly129 alleles frequencies. 2M-129S-169M allele associated with CWD predisposition was predominantly frequent in domesticated deers. Due to more *PRNP* gene variability in Russian reindeer, it is possible that CWD incidence in such population will lead to a situation similar to what is observed in Norway (appearing of different CWD strains with different neuropathological and transmission patterns).

**Keywords:** *Rangifer tarandus*, chronic wasting disease, prion.

## ГЕНЕТИЧЕСКИЕ МАРКЕРЫ ОВЕЦ КАРАЧАЕВСКОЙ ПОРОДЫ, ВЫЯВЛЕННЫЕ МЕТОДОМ ПОЛНОГЕНОМНОГО ПОИСКА АССОЦИАЦИЙ

© 2022 г. А. Ю. Криворучко<sup>1</sup>, О. А. Яцык<sup>1</sup>, А. В. Скокова<sup>1</sup> \*, А. А. Каниболоцкая<sup>1</sup>

<sup>1</sup>Северо-Кавказский федеральный научный аграрный центр, Ставропольский край, Михайловск, 356241 Россия

\*e-mail: antoninaskokova@mail.ru

Поступила в редакцию 14.05.2021 г.

После доработки 15.07.2021 г.

Принята к публикации 20.07.2021 г.

Проведен полногеномный поиск ассоциаций (GWAS) у овец карачаевской породы с использованием ДНК-биочипов Ovine Infinium HD BeadChip 600K для выявления вариантов однонуклеотидного полиморфизма (SNP), связанных с породной принадлежностью. На основании проведенного анализа было выявлено более 100 SNP, имеющих высокодостоверные различия в частоте встречаемости у животных карачаевской породы; для последующей оценки было отобрано по пять единичных SNP с наивысшими показателями на хромосомах 12, 13 и 14. Три замены находятся в экзонах генов, одна расположена в интроне, остальные – на разном удалении от генов. Самую высокую достоверность в ассоциациях с породной принадлежностью показали замены rs408511664, rs429132161 и rs409651063 ( $p = 4.637 \times 10^{-84}$ ), находящиеся на 14-й хромосоме и встречающиеся только у овец карачаевской породы. Большинство обнаруженных SNP локализованы рядом с генами, играющими немаловажную роль в процессе развития фенотипических признаков (*TUBB3*, *MC1R*, *BMP2*, *LYPLAL1-201*, *PAPPA2*). Проведенные исследования позволяют использовать совокупность выявленных SNP в качестве информативных молекулярных маркеров при идентификации породной принадлежности карачаевских овец.

**Ключевые слова:** овцы, молекулярные маркеры, ДНК-биочип, полногеномный поиск ассоциаций, однонуклеотидный полиморфизм.

**DOI:** 10.31857/S0016675822020096

Определение породной принадлежности овец является основой для селекционной работы. В настоящее время оно выполняется с использованием показателей фенотипа и данных родословной с подтверждением достоверности происхождения [1, 2]. Однако этой информации на современном уровне развития овцеводства уже недостаточно. Новые методы оценки породной принадлежности необходимы для выявления различий у близких по происхождению пород, для регистрации новой породы, определения происхождения гибридов, при выполнении экспертизы мясной продукции и целом ряде других мероприятий [3, 4]. Особенно это актуально в странах с большим количеством используемых пород. К таким относится и Россия, где официально зарегистрировано более 40 пород овец.

Наиболее перспективным для оценки породной принадлежности является применение современных достижений генетики. Методы кластеризации овец на отдельные породы по молекулярным ДНК-маркерам используются достаточно давно. Для

этого чаще всего применяют генотипирование по микросателлитным локусам [5]. Их полиморфизма достаточно для разделения животных на группы, соответствующие породе, по совокупности встречающихся с разной частотой аллелей [6].

Между тем обратный процесс, т.е. определение породной принадлежности конкретного животного по полиморфизму микросателлитных локусов, не может быть выполнен с высокой достоверностью. Это связано с тем, что уникальных аллелей микросателлитных локусов, встречающихся только у одной породы, не обнаружено. Поэтому оценка носит вероятностный характер и не достаточна для точной идентификации породы. Тем не менее при проведении генетической паспортизации или экспертизы продукции достоверность определения породы по результатам исследования ДНК должна приближаться к 100 процентам [7, 8].

Более точную оценку породной принадлежности можно сделать при использовании в качестве маркеров вариантов однонуклеотидного поли-

морфизма (Single Nucleotide Polymorphism, SNP). Их количество в геноме значительно превышает число микросателлитных локусов, а число аллелей, как правило, не более двух. Используя уникальную комбинацию из нескольких SNP, встречающихся только у исследуемой породы, вполне можно достичь требуемой точности определения породной принадлежности [9]. Для обнаружения этих SNP можно использовать проведение полногеномного поиска ассоциаций с помощью ДНК-биочипов высокой плотности [10]. Выполнение исследования целесообразно по принципу “случай–контроль”. При этом “случай” – животные исследуемой породы, а “контроль” – особи других пород. Это позволит выявить однонуклеотидные варианты полиморфизма, достоверно отличающие исследуемую породу на уровне генома. Обнаруженные SNP могут быть использованы как молекулярные маркеры породы [11]. Расположенные рядом с ними гены можно рассматривать как гены-кандидаты, их полиморфизм обеспечивает проявление фенотипических признаков, на основании которых животные были выделены в отдельную породу [12].

Интересным представляется поиск молекулярных маркеров породной принадлежности среди аборигенных пород овец, длительно существующих в условиях относительной изоляции. Одной из таких пород является карачаевская, разводимая в условиях Северного Кавказа на протяжении двухсот лет. По продуктивности она относится к грубошерстным, однако используется как универсальная. От овец получают шерсть (настриг у баранов до 3 кг, у ярок до 2 кг), мясо (вес баранов доходит до 70 кг) и молоко, выход которого составляет 50 кг в год. Несмотря на то что животные эти некрупные, порода является скороспелой, а мясо обладает высокими вкусовыми качествами. Овцы отличаются высокой устойчивостью к воздействию климатических факторов, так как разводятся в горной местности, а также резистентны ко многим инфекционным заболеваниям [13, 14].

Целью нашей работы стал поиск новых молекулярных маркеров карачаевской породы методом полногеномного поиска ассоциаций.

## МАТЕРИАЛЫ И МЕТОДЫ

Исследования проводились на базе лабораторий Всероссийского научно-исследовательского института овцеводства и козоводства (ВНИИОК) – филиала ФГБНУ “Северо-Кавказский федеральный научный аграрный центр”, Сколковского института науки и технологий “Сколтех”, Научно-диагностического и лечебного ветеринарного центра ФГБОУ ВО “Ставропольский государственный аграрный университет”, племенного завода “Восток” Ставропольского края.

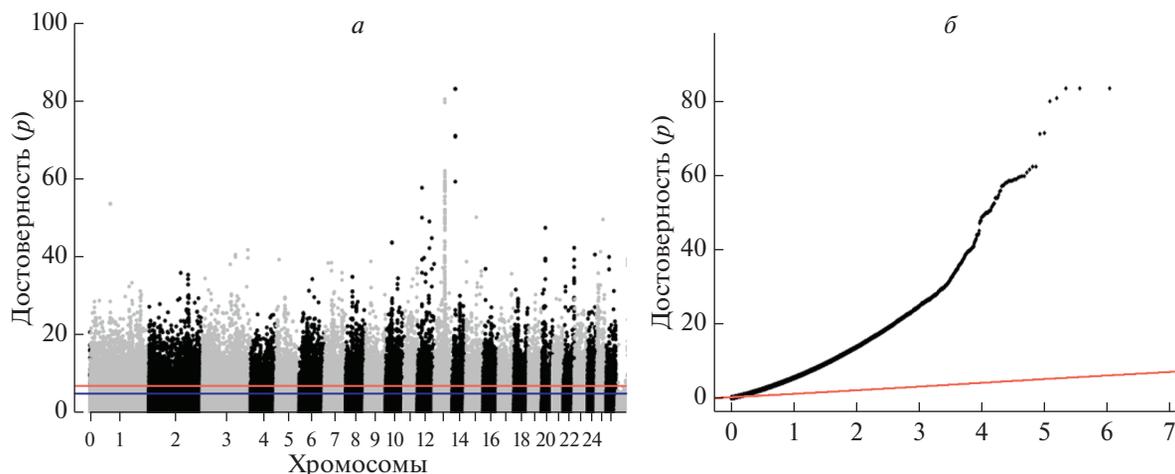
Исследование было организовано по принципу “случай–контроль”, где “случай” соответствовал баранчикам карачаевской породы, а “контроль” – баранчикам других пород. В группу “случай” по итогам комплексной оценки были отобраны 55 баранчиков карачаевской породы (из 417 баранчиков, полученных от маток селекционного ядра) в возрасте 12 мес. Критерии комплексной оценки (бонитировки) баранчиков – густота, длина, тонина шерсти, ее настриг в чистом волокне, качество жиропота, экстерьер, живая масса. По результатам проведенной бонитировки все 55 баранчиков были отнесены к классу “элита” и имели наиболее ярко выраженные признаки карачаевской породы. Группа “контроль” состояла из 220 животных (по 55 особей каждой породы) северокавказской мясо-шерстной породы, романовской породы, породы джалгинский меринос и российский мясной меринос. Они также отбирались из представителей селекционного ядра по тем же критериям комплексной оценки, как имеющие наиболее ярко выраженные признаки соответствующей породы. Все баранчики были клинически здоровыми.

### *Генотипирование*

Геномную ДНК выделяли из образцов цельной крови, взятой в асептических условиях из яремной вены, с использованием набора Pure Link Genomic DNA MiniKit (Invitrogen Life Technologies, США) в соответствии с протоколом производителя. Генотипирование животных выполнялось с использованием Ovine Infinium HD BeadChip 600K (Illumina, США) согласно протоколу производителя. Первичную обработку результатов генотипирования выполняли с использованием программного обеспечения Genome Studio 2.0 (Illumina).

### *Контроль качества генотипирования*

Контроль качества генотипирования проводили с использованием программного обеспечения PLINK V.1.07 [15]. В обработку данных были включены образцы с показателем количества выявленных SNP (Call Rate) больше 0.95. Из анализа были исключены SNP с частотой минорных аллелей (MAF – Minor Allele Frequency) меньше 0.01, частотой потерянных генотипов (missing genotype) больше 0.1. В качестве порогового значения по критерию Харди–Вайнберга (Hardy–Weinberg equilibrium) методом Фишера использовалось значение  $p = 0.00001$ . С полизиготностью результатом контроль качества генотипирования прошли все 275 образцов. Из 606006 SNP для дальнейшего анализа были использованы 550991 полиморфизм.



**Рис. 1.** Результаты полногеномного поиска ассоциаций у овец карачаевской породы. *а* – Манхэттенский график результатов GWAS с набором значений  $-\log_{10}(p)$  для исследуемых SNP. Нижняя линия обозначает порог различий с ожидаемой достоверностью различий при значении  $-\log_{10}(p) = 5$ , верхней линией указан порог высокой достоверности различий при значении  $-\log_{10}(p) = 7$ ; *б* – Q-Q-график для вероятностей распределения достоверности оценок SNP по всему геному. Точками обозначены значения  $-\log_{10}(p)$  для отдельных SNP. Линия обозначает ожидаемые значения при подтверждении нулевой гипотезы об отсутствии ассоциаций.

#### Генетический и статистический анализы

Полногеномный поиск ассоциаций выполняли с использованием программного обеспечения PLINK V.1.07 [15] на основе оценки значимости влияния SNP на породную принадлежность особей. Для подтверждения достоверности различий при множественных сравнениях использовали оценку  $p$  с поправкой Бонферрони. Визуализацию и построение графиков проводили с применением пакета “Qqman” на языке программирования R. Поиск генов-кандидатов выполнялся в области 500000 пн вокруг SNP (половина сантиморганиды), показавших достоверные различия по встречаемости среди животных исследуемых групп. Для картирования использовалась сборка генома Oar\_rambouillet\_v1.0. Аннотирование генов выполнялось с использованием геномных браузеров UCSC ([www.genome.ucsc.edu](http://www.genome.ucsc.edu)) и Ensembl ([www.ensembl.org](http://www.ensembl.org)).

### РЕЗУЛЬТАТЫ

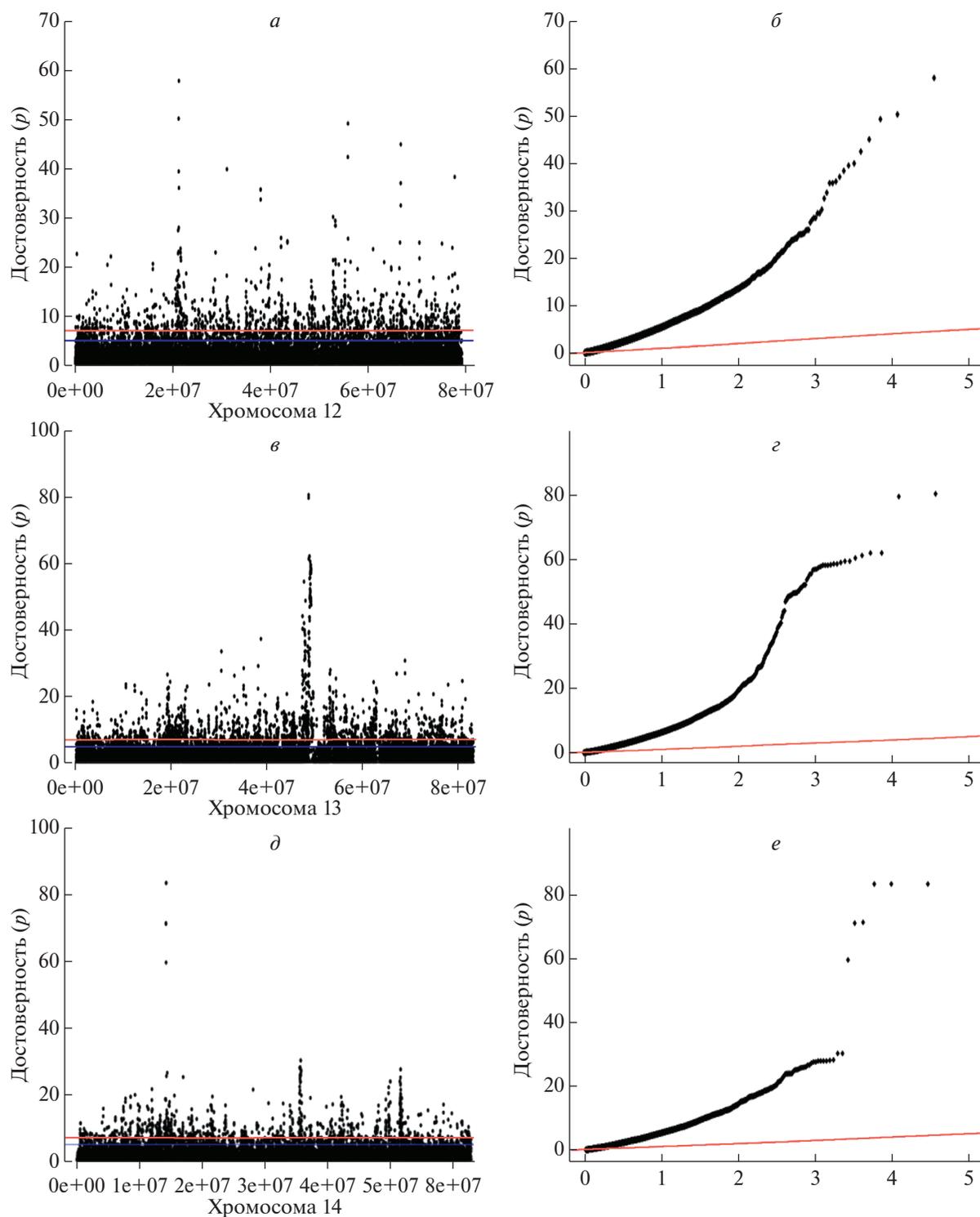
Полногеномный поиск ассоциаций частоты встречаемости вариантов однонуклеотидного полиморфизма с фенотипическими признаками карачаевской породы овец показал наличие очень большого количества локусов, преодолевших порог достоверности, установленный в соответствии с поправкой Бонферрони на уровне  $-\log_{10}(p) = 7$  (рис. 1, *а*). Квантиль-квантиль график (рис. 1, *б*) показывает, что абсолютное большинство исследуемых SNP имеет показатели, отличающиеся от таковых при подтверждении нулевой гипотезы (отсутствие достоверных

различий). Отклонение от ожидаемых значений наблюдается начиная от величины  $-\log_{10}(p) > 0.5$ .

Наиболее высокие показатели достоверности выявлены для SNP, расположенных на хромосомах 12, 13 и 14 (рис. 2). На 12-й хромосоме полиморфизмы, связанные с породной принадлежностью, находятся в локусах по всей длине хроматид. Максимальные отличия отмечены для rs424156280, находящейся в позиции 73451671. Хромосома 13 имеет целый пул SNP с высокими показателями достоверности, локализованный вблизи ее центральной области. На 14-й хромосоме нами были выявлены SNP с наибольшими показателями достоверности ассоциаций с породной принадлежностью, однако их количество существенно уступает хромосомам 12 и 13.

В связи с большим количеством SNP, достоверность связи которых с породной принадлежностью животных превышает установленный порог, для дальнейшего анализа их расположения в геноме относительно аннотированных генов были отобраны по пять SNP с максимальными показателями на хромосомах 12, 13 и 14 (табл. 1).

Из 15 рассматриваемых замен три находятся в экзонах генов, одна в интроне, а остальные расположены на разном удалении от кодирующих генов. Ближе всего к границам генов находятся замены rs408511664 и rs429132161 в 14-й хромосоме. Замена rs415954733 в хромосоме 12 располагается дальше всех рассматриваемых от ближайшего гена, однако это расстояние не превышает половины сантиморганиды. На основе сборки генома Oar\_rambouillet\_v1.0. нам не удалось картировать замену rs409651063, поэтому ее парамет-



**Рис. 2.** Результаты поиска ассоциаций по хромосомам у овец карачаевской породы. *а, в, д* – Манхэттенский график результатов GWAS с набором значений  $-\log_{10}(p)$  исследуемых SNP в отдельных локусах хромосом 12, 13 и 14. Нижняя линия обозначает порог различий с ожидаемой достоверностью различий при значении  $-\log_{10}(p) = 5$ , верхней линией указан порог высокой достоверности различий при значении  $-\log_{10}(p) = 7$ ; *б, г, е* – Q-Q-график для вероятностей распределения достоверности оценок SNP, локализованных на хромосомах 12, 13 и 14. Точками обозначены значения  $-\log_{10}(p)$  для отдельных SNP. Линия обозначает ожидаемые значения при подтверждении нулевой гипотезы об отсутствии ассоциаций.

**Таблица 1.** Характеристика SNP с наибольшими показателями достоверности ассоциации с породной принадлежностью животных при проведении GWAS

№	SNP	Хромосома/ позиция	A1	A2	F_A	F_U	<i>p</i>	Ген/расстояние до гена
1	rs408511664	14/15483844	A	G	0.75	0	4.637e-84	<i>TUBB3</i> /перед геном 6999 пн
2	rs429132161	14/15489751	A	G	0.75	0	4.637e-84	<i>TUBB3</i> /перед геном 1092 пн
3	rs409651063 (Oar_v3.1)	14/14231948	A	G	0.75	0	4.637e-84	<i>MC1R</i> /экзон 3, 105 аа, D/N (миссенс-мутация)
4	rs413629721	14/15451699	C	A	0.75	0.02	4.861e-72	LncRNA/экзон 10 (некодирующий транскрипт)
5	rs412344125	14/15448879	A	C	0.74	0.02	8.764e-72	LncRNA/экзон 4 (некодирующий транскрипт)
6	rs427301675	13/51330485	A	G	0.80	0.02	1.966e-81	<i>BMP2</i> /129839 пн
7	rs411672645	13/51327453	A	G	0.80	0.02	1.402e-80	<i>BMP2</i> /126807 пн
8	rs412370368	13/51502951	A	G	0.79	0.06	5.580e-63	<i>BMP2</i> /302305 пн
9	rs401025140	13/51506415	G	A	0.79	0.06	5.580e-63	<i>BMP2</i> /305769 пн
10	rs425709044	13/51358671	G	A	0.81	0.07	3.400e-62	<i>BMP2</i> /158025 пн
11	rs426400216	12/25137722	A	G	0.68	0.03	1.133e-58	<i>LYPLAL1</i> -201/377126 пн
12	rs415954733	12/25076577	A	G	0.72	0.07	5.472e-51	<i>LYPLAL1</i> -201/315981 пн
13	rs419909433	12/61840136	A	G	0.75	0.09	5.502e-50	<i>PAPPA2</i> /перед геном 1724 пн
14	rs424156280	12/73451671	G	A	0.52	0.02	1.016e-45	<i>KCNK2</i> /интрон 3
15	rs406878732	12/61816311	A	G	0.75	0.11	3.803e-43	<i>PAPPA2</i> /25549 пн

Примечание. A1 – минорный аллель; A2 – главный аллель; F\_A – частота минорного аллеля в группе карачаевских овец; F\_U – частота минорного аллеля у животных группы сравнения.

ры определены на основе сборки Oar\_v3.1, согласно которой она попадает в экзон гена. Пять замен, локализованных на 13-й хромосоме, расположены в одном межгенном пространстве, на различном расстоянии от гена *BMP2* (Bone Morphogenetic Protein-2, *BMP2*). Две пары замен на 12-й хромосоме также находятся достаточно близко друг от друга. Гены, расположенные в области выявленных SNP, можно считать генами-кандидатами с возможным влиянием на фенотипические признаки породной принадлежности овец карачаевской породы.

### ОБСУЖДЕНИЕ

Проведенное исследование подтвердило наличие генетических основ внутривидовой (породной) изменчивости у овец. Выявлено достаточно большое количество SNP, достоверно различающихся по частоте встречаемости у овец карачаевской породы и других пород овец. При этом три замены на 14-й хромосоме были обнаружены только у карачаевских овец и не встречались ни у одной особи остальных пород. Две из них, rs408511664 и rs429132161, расположены рядом с геном тубулина бета 3 (Tubulin Beta 3 Class III, *TUBB3*). Кодированный им белок вместе с тубулином альфа образуют структурную основу мик-

ротрубочек цитоскелета [16]. Кроме чисто механической поддержки формы клетки и расположения органелл, микротрубочки играют важную роль в процессах транспорта внутриклеточных компонентов и отдельных молекул. Также они обеспечивают подвижность клеток, организацию работы нейронов, секрецию везикул и многие другие функции клеток [17]. Учитывая участие белка тубулина в таком количестве жизненно важных процессов, мы считаем, что его ген может обеспечивать фенотипические особенности карачаевских овец.

Замена rs409651063 находится в экзоне гена, кодирующего рецептор меланокортина 1 (Melanocortin 1 Receptor, *MC1R*). Так же как и две описанные выше замены она встречается только у карачаевских овец и не обнаруживается в группе контроля. Рецептор меланокортина участвует в регуляции синтеза меланина и влияет на цвет шерсти и волос у млекопитающих [18]. Наличие обнаруженной нами замены уже рассматривается как маркер темного цвета шерсти у некоторых видов овец [19, 20]. Стоит отметить, что карачаевская порода характеризуется шерстью черного цвета, а животные в группе контроля имеют белый цвет шерсти. Таким образом, наши исследования подтверждают результаты, полученные другими учеными, и позволяют считать ген *MC1R*

связанным с породной характеристикой карачаевских овец.

Еще два SNP на 14-й хромосоме, rs413629721 и rs412344125, располагаются в экзонах гена длинной некодирующей РНК. Функции некодирующих РНК до конца не изучены. Однако уже известно, что подобные РНК участвуют в регуляции многих внутриклеточных процессов [21]. Это не только влияние на трансляцию и транскрипцию ДНК, но и прямое воздействие на работу молекулярных комплексов в цитоплазме [22]. Эти замены обнаружены нами всего у 2% животных группы контроля, а у карачаевских овец носителями являются 75% выборки. Исходя из приведенных данных, ген длинной некодирующей РНК может быть связан с породной принадлежностью изучаемых овец.

На 13-й хромосоме все пять отобранных нами замен располагались на разном удалении от одного и того же гена, кодирующего костный морфогенетический белок типа 2 (Bone Morphogenetic Protein-2, *BMP2*). Вместе с другими белками этой группы он регулирует развитие костной и хрящевой тканей, стимулируя остеогенную дифференцировку клеток [23]. Было обнаружено, что полиморфизм гена *BMP2* связан с размерами тела и развитием мышц у крупного рогатого скота [24]. У овец выявлена связь этого гена с фенотипом жирнохвостых овец [25], к которым относится и карачаевская порода. Поэтому мы считаем возможным связывать наличие обнаруженных нами замен с породными особенностями карачаевских овец.

Обнаруженные нами замены rs426400216 и rs415954733 локализируются в области гена лизофосфолипазы I (Lysophospholipase-like protein 1, *LYPLAL1*). Несмотря на то что дистанция до гена достаточно велика, она укладывается в половину сантиморганы. Это значит, что исследуемый локус с большой вероятностью наследуется вместе с указанным геном. Лизофосфолипаза принимает участие в метаболизме фосфолипидов клеточной стенки и органелл. Однако, как и у других фосфолипаз, у нее имеются и регуляторные функции за счет включения в процессы синтеза эйкозаноидов и других медиаторов межклеточной коммуникации [26]. Исходя из этого, мы считаем ген *LYPLAL1* возможным кандидатом, влияющим на фенотип породы у овец.

Еще две замены на 12-й хромосоме, rs419909433 и rs406878732, находятся очень близко к гену паппализина 2 (Pappalysin 2, *PAPPA2*). Эта металлопротеиназа специфически разрезает протеин, связывающийся с инсулиноподобным фактором роста 5, важным стимулятором роста костей. Для гена паппализина 2 была установлена связь с размерами тела у мышей [27], что позволяет считать этот ген каким-то образом связанным с реализа-

цией фенотипических признаков карачаевской породы.

В интроне гена, кодирующего один из белков калиевого канала 2 типа (Potassium two pore domain channel subfamily K member 2, *KCNK2*), нами выявлена замена rs424156280. Этот тип калиевых каналов широко распространен в клетках организма, особенно в нейронах, и относится к классу липидно-управляемых. Обеспечивая поддерживающие трансмембранного потенциала покоя, каналы этого типа участвуют во всех клеточных процессах, связанных с изменениями электрических параметров мембраны [28]. Важность этих функций указывает на то, что ген *KCNK2* вполне может оказывать влияние на параметры фенотипа как через связь с процессами трансмембранного транспорта, так и посредством регуляции вегетативной нервной системой.

Выявленные нами замены, достоверно различающиеся по частоте встречаемости у овец карачаевской породы, по сравнению с другими породами, располагаются рядом с генами, выполняющими важные функции в развитии фенотипических признаков. Для некоторых из них уже имеются данные о связи с внешними параметрами животных. Мы считаем возможным использовать совокупность выявленных замен в качестве молекулярно-генетических маркеров для подтверждения породной принадлежности карачаевских овец. Предложенные гены-кандидаты, которые могут быть напрямую связаны с фенотипом животных, требуют дальнейшего изучения для поиска полиморфизмов, обеспечивающих реализацию признаков.

Работа выполнена при поддержке Минобрнауки РФ в рамках Государственного задания Федерального государственного бюджетного научного учреждения “Северо-Кавказский федеральный научный аграрный центр” АААА-А19-119072690006-3 (тема № 0513-2019-0002).

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Yurchenko A., Yudin N., Aitnazarov R. et al. Genome-wide genotyping uncovers genetic profiles and history of the Russian cattle breeds // *Heredity*. 2018. V. 120(2). P. 125–137. <https://doi.org/10.1038/s41437-017-0024-3>
2. Eriksson C., Pettit A. Designing cattle: The social practice of constructing breeds // *Anthrozoös*. 2020. V. 33(2). P. 175–190. <https://doi.org/10.1080/08927936.2020.1719758>
3. Li D., Sun G., Zhang M. et al. Breeding history and candidate genes responsible for black skin of Xichuan

- black-bone chicken // BMC Genomics. 2020. V. 21(1). P. 1–15.  
<https://doi.org/10.1186/s12864-020-06900-8>
4. *Barcaccia G., Lucchin M., Cassandro M.* DNA barcoding as a molecular tool to track down mislabeling and food piracy // Diversity. 2016. V. 8(1). P. 2.  
<https://doi.org/10.3390/d8010002>
  5. *Широкова Н.В., Гетманцева Л.В., Колосов Ю.А. и др.* Полиморфизм микросателлитных локусов OarCP549, CSRD247, FCB20 и MAF65 у овец // Аграрная наука Евро-Северо-Востока. 2017. № 5. С. 57–62.
  6. *Арнаутовский И.Д., Шарвадзе Р.Л., Гоголов В.А. и др.* Племенному животноводству – инновационные, молекулярно-генетические, биотехнические технологии и современные кадры // Дальневосточный аграрный вестник. 2017. Т. 3(43). С. 75–83.
  7. *McClure M.C., McCarthy J., Flynn P. et al.* SNP data quality control in a national beef and dairy cattle system and highly accurate SNP based parentage verification and identification // Frontiers Genet. 2018. № 9. P. 84.  
<https://doi.org/10.3389/fgene.2018.00084>
  8. *Muñoz M., García-Casco J.M., Alves E. et al.* Development of a 64 SNV panel for breed authentication in Iberian pigs and their derived meat products // Meat Sci. 2020. V. 167. P. 108152.  
<https://doi.org/10.1016/j.meatsci.2020.108152>
  9. *Pasupa K., Rathasamuth W., Tongsim S.* Discovery of significant porcine SNPs for swine breed identification by a hybrid of information gain, genetic algorithm, and frequency feature selection technique // BMC Bioinformatics. 2020. V. 21. P. 1–28.  
<https://doi.org/10.1186/s12859-020-3471-4>
  10. *Денискова Т.Е., Петров С.Н., Доцев А.В. и др.* Поиск QTL и генов-кандидатов, ассоциированных с интенсивностью роста овец, на основе GWAS // Биотехнология в растениеводстве, животноводстве и с.-х. микробиологии. 2020. С. 95–97.
  11. *Gebrehiwot N.Z., Strucken E.M., Marshall K. et al.* SNP panels for the estimation of dairy breed proportion and parentage assignment in African crossbred dairy cattle // Genet. Selection Evol. 2021. № 53(1). P. 1–18.  
<https://doi.org/10.1186/s12711-021-00615-4>
  12. *Tinarelli S., Ribani A., Utzeri V. et al.* Redefinition of the mora romagnola pig breed herd book standard based on DNA markers useful to authenticate its “Mono-Breed” products: An example of sustainable conservation of a livestock genetic Resource // Animals. 2021. V. 11(2). P. 526.  
<https://doi.org/10.3390/ani11020526>
  13. *Шевхужева Л.А., Арова О.З., Шевхужева С.А.* О мерах по дальнейшему расширению производства мяса в Карачаево-Черкесской Республике // Изв. ОГАУ. 2012. № 1-1. С. 166–169.
  14. *Болатчиев А.Т., Гоголев Э.Х., Шаманова О.А. и др.* Молочная продуктивность и качество молока овцематок карачаевской породы при разных условиях содержания // Овцы, козы, шерстяное дело. 2020. № 1. С. 19–20.
  15. *Purcell S., Neale B., Todd-Brown K. et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses // Am. J. Hum. Genet. 2007. V. 81. № 3. P. 559–575.  
<https://doi.org/10.1086/519795>
  16. *Lee Y.H., Park N.H.* A Complex cortical malformation caused by a mutation in the tubulin-encoding TUBB3 gene // J. Korean Soc. Radiol. 2020. V. 81(5). P. 1246–1249.  
<https://doi.org/10.3348/jksr.2019.0175>
  17. *Mutch C.A., Poduri A., Sahin M. et al.* Disorders of microtubule function in neurons: imaging correlates // Am. J. Neuroradiology. 2016. V. 37(3). P. 528–535.  
<https://doi.org/10.3174/ajnr.A4552>
  18. *Goud T.S., Upadhyay R.C., Onteru S.K. et al.* Identification and sequence characterization of melanocortin 1 receptor gene (MC1R) in *Bos indicus* versus (*Bos taurus* × *Bos indicus*) // Animal Biotechnology. 2020. V. 31(4). P. 283–294.  
<https://doi.org/10.1080/10495398.2019.1585866>
  19. *Hepp D., Gonçalves G., Moreira G. et al.* Identification of the e allele at the Extension locus (MC1R) in Brazilian Creole sheep and its role in wool color variation // Genet. Mol. Res. 2012. V. 11(11). P. 2997–3006.  
<https://doi.org/10.4238/2012.May.22.5>
  20. *Koseniuk A., Ropka-Molik K., Rubiś D. et al.* Genetic background of coat colour in sheep // Archives Animal Breeding. 2018. V. 61(2). P. 173–178.  
<https://doi.org/10.5194/aab-61-173-2018>
  21. *Мустафин Р.Н., Хуснутдинова Э.К.* Некодирующие части генома как основа эпигенетической наследственности // Вавиловский журн. генетики и селекции. 2017. Т. 21(6). С. 742–749.
  22. *Рябых Г.К., Жарикова А.А., Миронов А.А.* Анализ данных РНК-ДНК взаимодействий: Сб. трудов 42-й междисциплинарной школы-конф. ИППИ РАН “Информационные технологии и системы 2018”. 2018. С. 659–661.
  23. *Lu Z., Liu J., Han J. et al.* Association between BMP2 functional polymorphisms and sheep tail type // Animals. 2020. V. 10(4). P. 739.  
<https://doi.org/10.3390/ani10040739>
  24. *Ma T., Liu G., Lin D. et al.* BMP2 gene is a direct target gene of miR-378 in cattle // Indian J. Animal Sci. 2017. V. 87(6). P. 768–769. <http://www.icar.org.in/>.
  25. *Baazaoui I., Bedhiaf-Romdhani S., Mastrangelo S. et al.* Genome-wide analyses reveal population structure and identify candidate genes associated with tail fatness in local sheep from a semi-arid area // Animal. 2021. 100193.  
<https://doi.org/10.1016/j.animal.2021.100193>
  26. *Wepy J.A., Galligan J.J., Kingsley P.J. et al.* Lysophospholipases cooperate to mediate lipid homeostasis and lysophospholipid signaling // J. Lipid Res. 2019. V. 60(2). P. 360–374.  
<https://doi.org/10.1194/jlr.M087890>
  27. *Christians J.K., Hoeflich A., Keightley P.D.* PAPP2, an enzyme that cleaves an insulin-like growth-factor-binding protein, is a candidate gene for a quantitative trait locus affecting body size in mice // Genetics. 2006. V. 173(3). P. 1547–1553.  
<https://doi.org/10.1534/genetics.106.057513>
  28. *Es-Salah-Lamoureux Z., Steele D.F., Fedida D.* Research into the therapeutic roles of two-pore-domain potassium channels // Trends in Pharmacological Sci. 2010. V. 31(12). P. 587–595.  
<https://doi.org/10.1016/j.tips.2010.09.001>

## Genetic Markers of Karachaevsky Sheep Identified by Genome-Wide Associations Study

A. Yu. Krivoruchko<sup>a</sup>, O. A. Yatsyk<sup>a</sup>, A. V. Skokova<sup>a, \*</sup>, and A. A. Kanibolotskaya<sup>a</sup>

<sup>a</sup>North Caucasus Federal Agrarian Research Centre, Stavropolskii kray, Mikhailovsk, 356241 Russia

\*e-mail: antoninaskokova@mail.ru

A genome-wide associations study (GWAS) was carried out in Karachaevsky sheep using Illumina Ovine Infinium HD BeadChip 600K DNA biochips to identify single nucleotide polymorphisms (SNPs) associated with the breed. Based on the analysis, more than 100 SNPs were identified that have highly significant differences in the frequency of occurrence in animals of the Karachaevsky breed, for subsequent assessment, 5 single polymorphisms were selected with the highest indicators on the 12th, 13th and 14th chromosomes. Three substitutions are located in the exons of genes, one is located in the intron, and the rest are located at different distances from the genes. The substitutions rs408511664, rs429132161, and rs409651063 ( $p = 4.637 \times 10^{-84}$ ), located on chromosome 14 and found only in sheep of the Karachaevsky breed, showed the highest reliability in associations with breed affiliation. Most of the detected SNPs are localized near genes that play an important role in the development of phenotypic traits (*TUBB3*, *MC1R*, *BMP2*, *LYPLAL1-201*, *PAPPA2*). The studies make it possible to use the set of identified SNPs as informative molecular markers in identifying the breed of Karachaevsky sheep.

**Keywords:** sheep, molecular markers, DNA biochip, genome-wide associations study, single nucleotide polymorphism.

**ФИЛОГЕОГРАФИЯ *Lacerta media* LANTZ et CYRÉN, 1920  
(Lacertidae: Sauria) ПО РЕЗУЛЬТАТАМ АНАЛИЗА  
МИТОХОНДРИАЛЬНОГО ГЕНА ЦИТОХРОМА *b***

© 2022 г. М. А. Доронина<sup>1</sup>, И. В. Доронин<sup>1</sup>, \*, \*\*, С. А. Луконина<sup>2</sup>,  
Л. Ф. Мазанова<sup>3</sup>, А. В. Барабанов<sup>1</sup>

<sup>1</sup>Зоологический институт Российской академии наук, Санкт-Петербург, 199034 Россия

<sup>2</sup>Пензенский государственный университет, Пенза, 440026 Россия

<sup>3</sup>Дагестанский государственный университет, Махачкала, 367025 Россия

\*e-mail: igor.doronin@zin.ru

\*\*e-mail: ivdoronin@mail.ru

Поступила в редакцию 24.05.2021 г.

После доработки 24.08.2021 г.

Принята к публикации 07.09.2021 г.

Проведен анализ *cytb* митохондриальной ДНК (1143 пн) у представителей *Lacerta media*. Проанализированы 34 образца из 19 локалитетов. Особое внимание уделено ранее неисследованным популяциям на северной границе ареала в пределах Северного Кавказа и Западного Закавказья. Реконструировано два типа дендрограмм филогенетических отношений, показавших идентичную топологию — байесовская (BA) и максимального правдоподобия (ML). Филогенетическая реконструкция выделяет четыре поддерживаемые клады. Аналогично на медианной сети гаплотипов обособляются четыре гаплогруппы. На более ранних этапах процесса филогенеза *L. m. media* выделяются популяции восточной части Армянского нагорья и Северного Загроста. В дальнейшем происходит обособление популяций центральной части Армянского нагорья, Большого Кавказа, а на завершающем этапе — Центрального Загроста и восточной части Малоазиатского нагорья. Полученные данные не поддерживают ранее предложенный Ф. Ахмадзаде с соавт. сценарий расселения номинативного подвида с запада на восток и северо-восток. Взамен его мы рассматриваем противоположное направление — с востока на север, юг и запад. Выявлено две географические области, где обитают представители нескольких клад/гаплогрупп — Армянское нагорье и горы Загроста.

**Ключевые слова:** *Lacerta media*, митохондриальная ДНК, филогения, филогеография, цитохром *b*.

**DOI:** 10.31857/S0016675822020035

Средняя ящерица, *Lacerta media* Lantz et Cyrén, 1920 — представитель надвидового комплекса *Lacerta (trilineata)*, находящегося в фокусе исследований филогении и филогеографии рода *Lacerta* Linnaeus, 1758 [1]. В ряде публикаций последних лет [2–6] популяции с территории Кавказа, где обитает номинативный подвид (по месту сбора лектотипа — долина р. Квабисхеви, окр. г. Боржом Грузии [7]), были представлены выборкой лишь из пяти локалитетов с Армянского нагорья. Однако не были исследованы ящерицы с Северного Кавказа и Западного Закавказья — Дагестана и Краснодарского края, т.е. с северной границы видового ареала [8, 9]. В ряде работ эта территория вовсе отсутствует на картах распространения *L. media* [2, 3]. Обозначенные дагестанские и черноморские периферические популяции изолированы от закавказских и представляют интерес с позиции исследования микроэволюционных процессов.

В настоящее время отсутствует ясность с определением времени появления и расселения представителей комплекса *Lacerta (trilineata)*. Обнаружение в северокавказском позднемиоценовом местонахождении “Солнечнодольск” ископаемых костных фрагментов *L. trilineata* [10] ставят под сомнение существующие сценарии филогенеза комплекса, рассматривающие в качестве областей дивергенции основных филогенетических клад Малоазиатское нагорье (Анатолию) и район Эгейского моря в плиоцене—раннем плейстоцене [3, 6] или в позднем миоцене—раннем плиоцене [5]. В отношении *L. media* также имеется противоречивая информация: на основе сведений по палеогеографии региона время вселения средней ящерицы в горный Дагестан относят к сармату (поздний миоцен) [11], а в Южную Осетию — к концу миоцена—середине плиоцена [12], тогда как по молекулярно-генетическим датировкам

вид появился примерно 3.5 млн лет назад в позднем плиоцене [4, 6] или примерно 7.12 млн лет назад в позднем миоцене [5].

Вышесказанное свидетельствует о необходимости более обстоятельного исследования филогении и филогеографии средней ящерицы; на Кавказе мы можем ожидать обнаружения рефугиума, свидетельствующего о древней колонизации региона, либо подтверждение относительно недавнего расселения вида с территории Анатолии – предполагаемого центра ее происхождения и радиации. Подчеркнем, что палеонтологические свидетельства, которые могут служить реперными точками при реконструкции филогении вида, ограничиваются упоминанием средней ящерицы в мустьерских слоях (верхний плейстоцен) пещеры Швалиети в Грузии [13].

Немаловажным является то обстоятельство, что *L. m. media* включена в региональные [14, 15] и национальные [16, 17] списки охраняемых таксонов животных (Красные книги); это подчеркивает актуальность данного исследования с позиции изучения и сохранения генетического разнообразия.

## МАТЕРИАЛЫ И МЕТОДЫ

Для выделения геномной ДНК использовали части регенерированных хвостов или фаланги пальцев передних конечностей ящериц, фиксированные в 96%-ном этаноле, от пяти экземпляров, собранных в трех локалитетах Большого Кавказа и одном локалитете Загроса (рис. 1). Ваучерные экземпляры хранятся в коллекции Зоологического института РАН (ZISP) и Зоологическом музее МГУ (ZMMU). В филогенетический анализ были включены 29 последовательностей *L. m. media* [2, 3], взятых из GenBank (<http://www.ncbi.nlm.nih.gov/gene/>) (табл. 1).

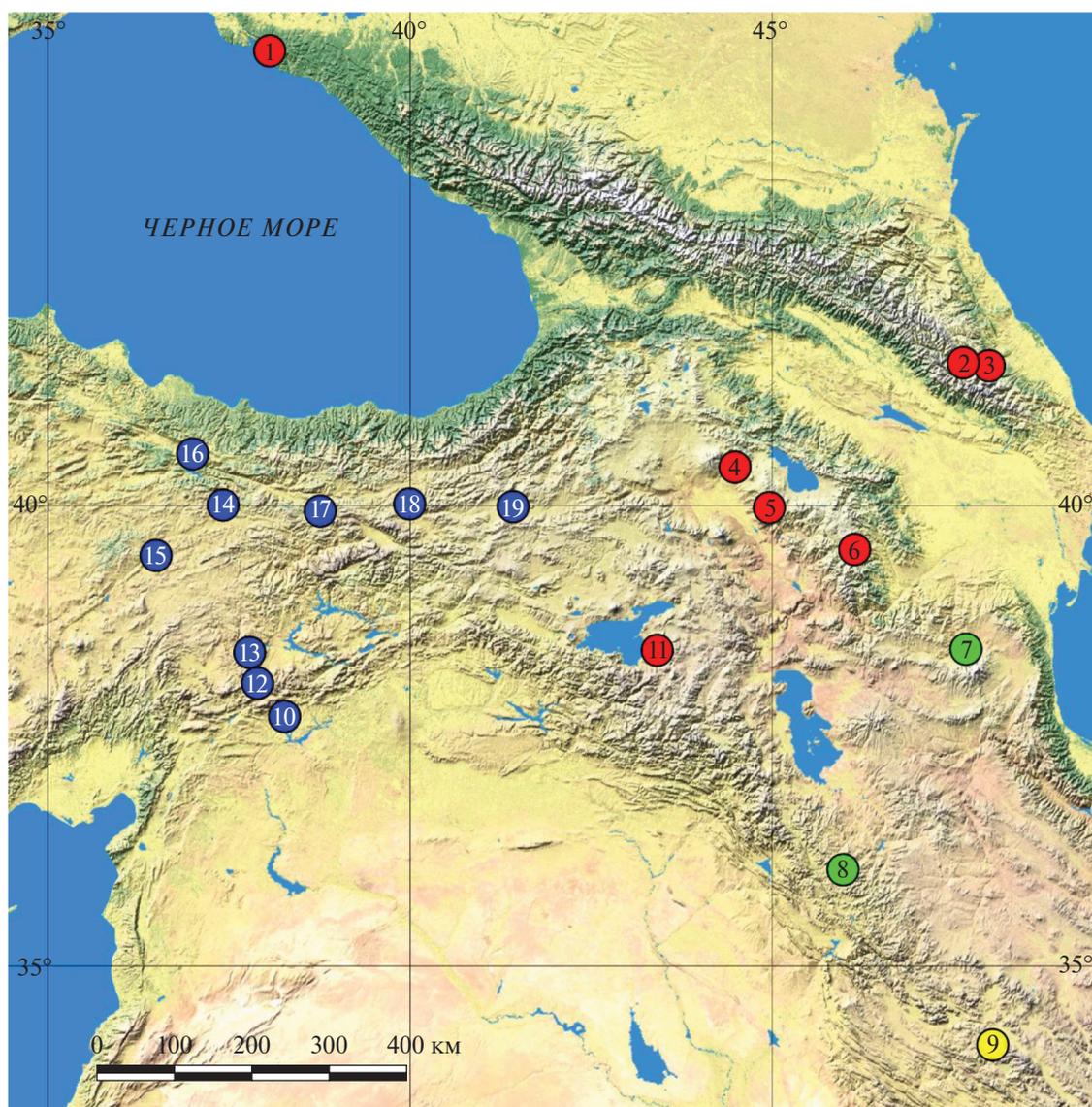
Выделение тотальной ДНК проводили стандартным солевым методом с лизированием протеиназой К [18]. Амплификацию участка гена *cytb* осуществляли с использованием ПЦР-амплификатора T100TM Thermal Cycler (Bio-Rad) и праймеров 5'-GCC CCA AAA TAA GGA GAT GG-3' и 5'-TAG TGA TGG GGG ATT AAA GC-3' [19, 20] по следующей схеме: начальная денатурация 95°C (3 мин), затем 32 цикла (95°C – 30 с, 56°C – 30 с, 72°C – 90 с) и конечная элонгация (72°C – 5 мин). Реакционная смесь для ПЦР (25 мкл) содержала 50–100 нг ДНК, 0.5 мкМ каждого праймера, 0.2 мМ dNTPs, 1.5 мМ MgCl<sub>2</sub>, 2.5 мкл 10× ПЦР буфера (10 мМ Tris-HCl, pH 8.3, 50 мМ KCl) и 2 ед. Таq-полимеразы (Thermo Scientific). Секвенирование проводили на генетическом анализаторе ABI PRISM 3500xL (Applied Biosystems). Окончательная длина амплифицированного фрагмента ДНК составила 1143 пн. Полученные последовательно-

сти (MZ065554–065558) депонировали в GenBank NCBI.

Выравнивание последовательностей проводили в программах Geneious Prime 2021.0.1 (<https://www.geneious.com>) и Ali View 1.6 [21]. Филогенетические отношения были реконструированы с использованием метода Байеса (Bayesian analysis, BA) в программе Mr. Bayes 3.1.2 [22–24] и методом максимального правдоподобия (Maximum Likelihood, ML) в программном пакете MEGA X [25]. Для BA использовали 5 млн поколений; статистическую надежность узлов ML-деревьев оценивали путем бутстреп-анализа (1500 псевдореплик). Выбор модели эволюции для нуклеотидных последовательностей проводили в программе MrModeltest 2.4 [26] с использованием информационного критерия Акаике (AIC). Была выбрана модель GTR (General Time Reversible) с параметрами  $I = 0.3909$  и  $G = 0.7440$ . Графические изображения деревьев получены с помощью программы FigTree 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree>). Построение сети гаплотипов осуществляли методом TCS в программе PopART [27].

Для изучения внутривидовой структуры *L. media* в анализ были включены последовательности из работы Ф. Ахмадзаде с соавт. [3] – *L. m. ciliciensis* Schmidtler, 1975 (KC896967–896974), *L. m. israelica* Peters, 1964 (KC896975) и *L. m. wolterstorffi* Mertens, 1922 (KC897005–897007). Для укоренения дендрограмм филогенетических отношений в качестве внешней группы были выбраны последовательности *L. strigata* Eichwald, 1831 (LN835024.1), *L. agilis* Linnaeus 1758 (AY616285, AY616398) и *Eremias stummeri* Wettstein, 1940 (NC029878.1) из GenBank. При анализе генетической изменчивости с использованием программы DnaSP v.5.10.01 [28] вычисляли следующие параметры: общее количество полиморфных позиций ( $S$ ), общее количество замен ( $\eta$ ), количество гаплотипов ( $H$ ), разнообразие гаплотипов ( $h$ ), разнообразие нуклеотидов на сайт ( $\pi$ ), среднее количество нуклеотидных замен ( $K$ ), значение теста Таджимы (Tajima's  $D$ ) и тест на нейтральность Фу (Fu's  $F_s$ ). Генетические дистанции ( $p$ -дистанции) вычисляли в программе MEGA X.

Для проведения анализа таксономических границ по молекулярно-генетическим признакам нами были использованы алгоритмы ASAP (Automated simultaneous analysis phylogenetics) [29] и ABGD (Automatic barcode gap discovery) [30]. Они позволяют выявить условные молекулярные операционные таксономические единицы (Molecular phylogenetic taxonomic units, MOTU), представляющие собой монофилетические клады неопределенного ранга. Анализ проводили с использованием следующих параметров:  $P_{\min}$  (минимальная априорная дистанция) = 0.01,  $P_{\max}$  (максимальная априорная дистанция) = 0.1,  $X$  (относительная



**Рис. 1.** Пункты сбора экземпляров *Lacerta media*, использованных для молекулярно-генетического анализа. Обозначение клад: А – желтые, В – синие, С – красные, D – зеленые. Нумерация и буквенное обозначение соответствуют таковым в табл. 1.

ширина интервала) = 0.5; была использована матрица попарных K80-дистанций.

### РЕЗУЛЬТАТЫ

34 последовательности гена *cytb* *L. m. media* образуют 24 гаплотипа. Выявлены 54 (4.7% от общей длины фрагмента) вариабельные позиции, из них – 44 (3.8%) парсимони информативные. Соотношение транзаций к трансверсиям составило 4.84. Нуклеотидный состав участка на L-цепи смещен в сторону уменьшения доли гуанина (А – 26.7, Т – 31.3, С – 30, G – 12%). Показатели генетического разнообразия в выборке:  $S = 54$ ,  $\eta = 57$ ,  $h = 0.964 (\pm 0.021)$ ,  $\pi = 0.0160 (\pm 0.0010)$ ,  $K = 14.01$ , Tajima's  $D = 0.02$  ( $p > 0.1$ ), Fu's  $F_s = -1.80$

( $p > 0.1$ ). Показатель  $F_s$ -тест незначим. Недостаточность значительных отрицательных величин  $D$  свидетельствует об отсутствии в прошлом внезапного роста популяции (экспансии) и/или положительного отбора, а положительных – об отсутствии недавнего сокращения численности популяции (т.н. бутылочного горлышка), фрагментации и/или балансирующего отбора.

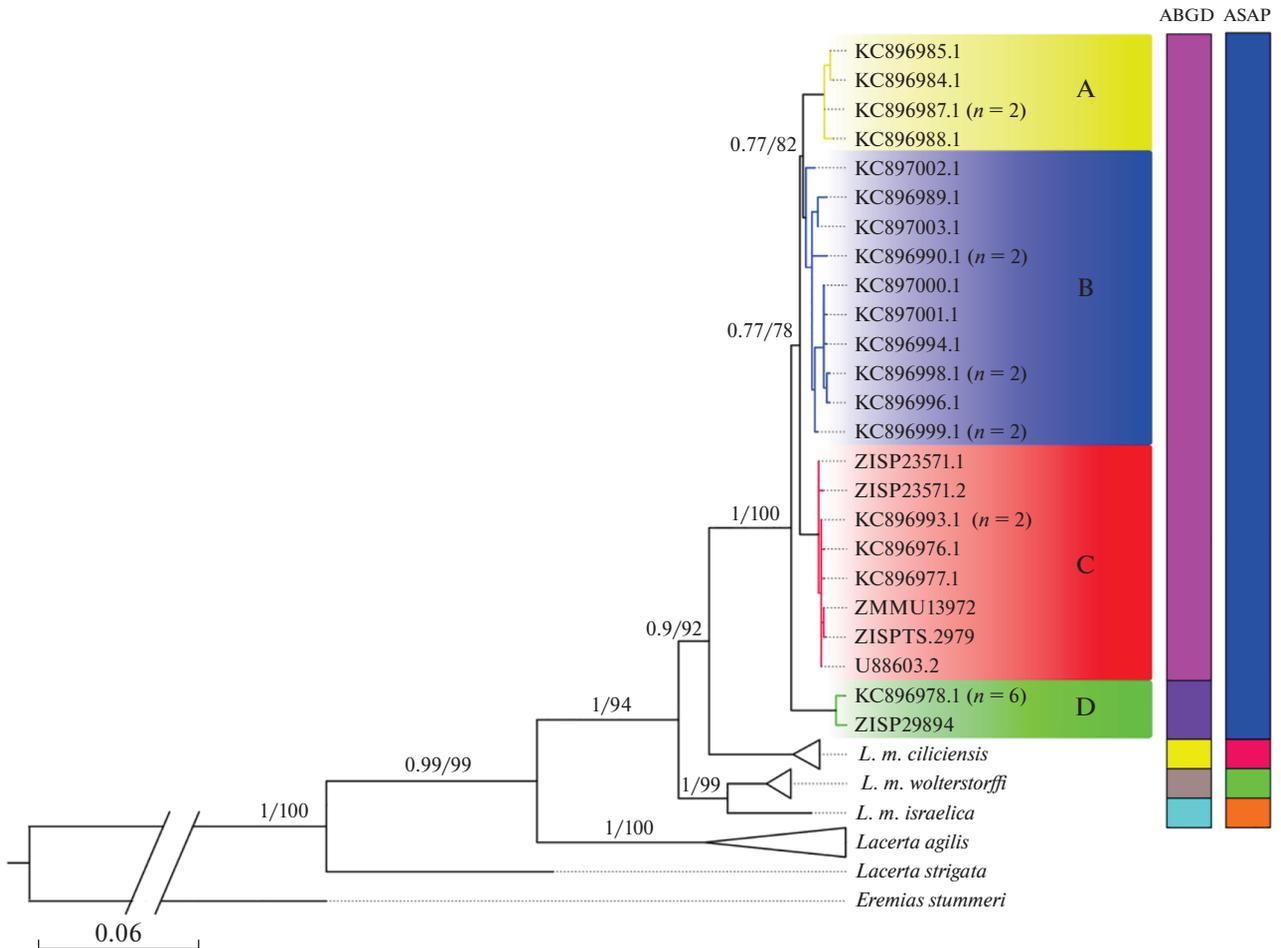
По результатам реконструкции филогенетических взаимоотношений представителей *L. m. media* были получены дендрограммы с идентичной топологией, на которых можно выделить четыре митохондриальные клады (рис. 2). Последовательности с территории Большого Кавказа вошли в кладу С, внутри которой нет значимой генетической гетерогенности ( $h = 0.889$ ,  $K = 1.67$ ). Для

Таблица 1. Коллекционные экземпляры *Lacerta media*, использованные при молекулярно-генетическом анализе

Номера локалитетов на рис. 1	Локалитеты	Коллекционные номера	Номера в Генбанке	Клада	Даты сбора	Географические координаты	Высота, м н. у. м.	Коллектор(-ы)	Источник информации
1	Россия, Краснодарский край, окр. г. Геленджик, Ю склон Маркотхского хребта	ZISP 23571.1-2	MZ065557 – MZ065558	C	28.04.2006	38.0624, 44.5992	160	С.В. Островских, К.Д. Мильго	Данная публикация
2	Россия, Дагестан, Ахтынский р-н, с. Кака	ZISP TS 2979	MZ065555	C	07.07.2020	47.6328, 41.4790	1130	З.С. Исмаилова	»
3	Россия, Дагестан, Ахтынский р-н, окр. с. Ахты, склон Ю экспозиции г. Ухиндаг	ZMMU 13972	MZ065556	C	03–04.05.2013	47.6916, 41.4478	1450	А.А. Аскендеров	»
4	Армения, Арагаотнская обл., гора Аралер (=Аралер, Аралер, Ара, Араилер)	–	U88603	C	–	44.4833, 40.4000	2150	–	[2]
5	Армения, Араратская обл., заповедник “Хосровский лес”	–	КС896976	C	–	44.9579, 39.9763	1500	–	[4]
6	Армения, Сюникская обл., трасса М-2 между с. Шаки и г. Горис	–	КС896977	C	–	46.1359, 39.5358	2010	–	»
7	Иран, Ардебиль, с. Араблу	–	КС896978– КС896983	D	–	47.6720, 38.4753	1110	–	»
8	Иран, Курдистан, 35 км ЮЗ г. Секкес, около трассы Секкес–Бане	ZISP 29894	MZ065554	D	08.08.2018	45.9746, 36.0727	2070	А.Н. Барабанов	Данная публикация
9	Иран, Лурестан, окр. г. Нурабад	–	КС896984– КС896988	A	–	48.0397, 34.1097	1800	–	[4]

Таблица 1. Окончание

Номера локалитетов на рис. 1	Локалитеты	Коллекционные номера	Номера в Генбанке	Клада	Даты сбора	Географические координаты	Высота, м н. у. м.	Коллектор(-ы)	Источник информации
10	Турция, Адьяман, г. Адьяман	–	КС897002– КС897003	В	–	38.2667, 37.7500	640	–	[4]
11	Турция, Ван, окр. г. Ван	–	КС896993, КС896995	С	–	43.4107, 38.4624	1810	–	»
12	Турция, Малатья, СВ г. Доганшехир	–	КС896989	В	–	37.8947, 38.1093	1180	–	»
13	Турция, Малатья, перевал Карахан	–	КС896990– КС896991	В	–	37.7855, 38.3393	1870	–	»
14	Турция, Сивас, дорога Доганшар–Хафик	–	КС897000	В	–	37.4197, 40.0029	1660	–	»
15	Турция, Сивас, окр. д. Казанджик	–	КС897001	В	–	36.4942, 39.4717	1630	–	»
16	Турция, Токат, между г. Никсар и г. Башчифт-лик	–	КС896994	В	–	37.0004, 40.5374	640	–	»
17	Турция, Эрзинджан, д. Кюрелик	–	КС896996– КС896998	В	–	38.7608, 39.9397	1730	–	»
18	Турция, Эрзинджан, окр. д. Кючюкоглукебели	–	КС896999	В	–	39.9939, 40.0147	2070	–	»
19	Турция, Эрзурум, 20 км СВ г. Эрзурум	–	КС896992	В	–	41.4220, 39.9864	2020	–	»



**Рис. 2.** Дендрограмма филогенетических отношений представителей *Lacerta media* по результатам анализа фрагмента гена *cytb* (1143 пн). В узлах указаны байесовы апостериорные вероятности и бутстреп-поддержки (BA/BS). Буквенное обозначение клад соответствует таковому в табл. 1 и 2. В двух столбцах разными цветами указаны выделенные молекулярные операционные таксономические единицы (MOTU).

сравнения наибольшую структурированность мы наблюдаем в кладе В ( $h = 0.961$ ,  $K = 6.08$ ). Аналогичное прослеживается при анализе медианной сети гаплотипов: выделяются четыре гаплогруппы (I–IV), наиболее удаленная из которых – IV (рис. 3). От II группы она отдалена на 25 нуклеотидных замен, от III – на 22, а от I – на 29.

**Таблица 2.** Средние генетические  $p$ -расстояния (%) (под диагональю) и их стандартные отклонения (над диагональю) между кладами *Lacerta media* по данным анализа последовательностей фрагмента гена *cytb*

Клады	A	B	C	D
A		0.35	0.38	0.51
B	1.58		0.33	0.48
C	1.63	1.39		0.49
D	2.79	2.59	2.45	

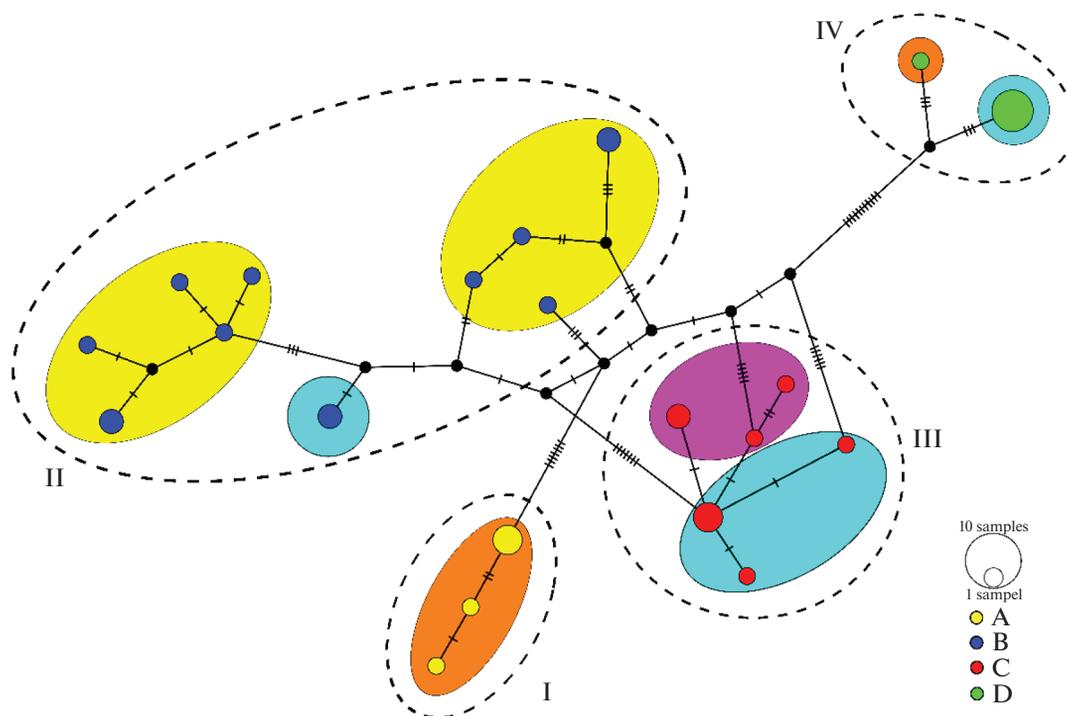
Примечание. Буквенное обозначение клад соответствует таковому в табл. 1 и на рис. 2.

При расчете  $p$ -расстояний максимальное значение было получено при сравнении клад A/D, B/D, а минимальное – B/C (табл. 2).

Анализ ASAP/ABGD показал наличие четкого разрыва между значениями популяционной и подвидовой изменчивости (threshold distance = 3.27%, barcoding gap = 2.20%). Применение данных алгоритмов позволило выявить в объеме *L. media* соответственно 4 и 5 монофилетических MOTU (рис. 2). Если в первом случае молекулярные операционные таксономические единицы строго соответствуют признаваемым подвидам средней ящерицы, то во втором дополнительно была делимитирована клада D.

### ОБСУЖДЕНИЕ

Согласно топологии полученной дендрограммы, первой на подвидовом уровне обособляется пара *L. m. israelica* и *L. m. wolterstorffi*, в дальнейшем – *L. m. ciliciensis* и *L. m. media*. Согласно пред-



**Рис. 3.** Медианная сеть гаплотипов *cytb* представителей *Lacerta m. media* (1143 пн). Каждый цветной кружок – уникальный гаплотип, чей размер пропорционален встречаемости в выборке; цвет кружка соответствует кладе на рис. 2: А – желтые, В – синие, С – красные, D – зеленые. Соединительные линии – вероятные эволюционные связи, засечки – нуклеотидные замены, черные кружки в узлах линий – предсказанные гаплотипы. Цветной заливкой обозначены географические регионы: оранжевой – Иранское нагорье, желтой – Малоазиатское нагорье, бирюзовой – Армянское нагорье, розовой – Большой Кавказ. Пунктирной линией ограничены группы гаплотипов.

шествующим публикациям [3, 4], Малоазиатское нагорье может рассматриваться как исходная территория, откуда происходило множественное независимое расселение средней ящерицы на Армянское нагорье (включая самые восточные турецкие популяции в районе оз. Ван), в горы Эльбурса на северо-западе Ирана и Загроса на западе Ирана; именно популяции Анатолии характеризуются наибольшим генетическим разнообразием внутри номинативного подвида. В процитированной работе [3] локалитет Араблун на северо-западе Ирана, расположенный на северном склоне потухшего вулкана Сабалан (=Савалан, Савелан или Кухе-Себелан), был отнесен к “Эльбурскому” рефугиуму. С этим мы не можем согласиться, поскольку хребет Савелан-Даг, в который входит вулкан, выступает восточной оконечностью Армянского вулканического нагорья [31]. Кроме того, *L. media* традиционно включают в восточносредиземноморскую зоогеографическую группу видов рептилий, в то время как Эльбурс – рефугиум гирканской герпетофауны [32]. Представители последней не обитают в районе Сабалана [33, 34].

Безусловно предковая форма *L. m. media* первоначально проникла на обозначенную территорию с запада, где происходила основная радиация

вида [3, 4]. Однако согласно полученной нами дендрограмме (рис. 2) на более ранних этапах процесса филогенеза *L. m. media* выделяются популяции восточной части Армянского нагорья и Северного Загроса (клада D). В дальнейшем происходит обособление популяций центральной части Армянского нагорья и Большого Кавказа (С), а на завершающем этапе – Центрального Загроса (А) и восточной части Малоазиатского нагорья (В). Взамен сценария расселения номинативного подвида с запада на восток и северо-восток мы предлагаем противоположное направление – с востока на север, юг и запад. Наши выводы о становлении ареала *L. m. media* соответствуют схеме, которую предложил в 1924 г. О. Сирен [35]. Уместно вспомнить, что И.С. Даревский [36] и Т.А. Мухелишвили [37] считали Армянское нагорье “центром формирования” средней ящерицы.

На Армянском нагорье обнаружены представители трех клад (В, С, D) и трех гаплогрупп (II, III, IV), что наглядно продемонстрировано на сети гаплотипов (рис. 3). Еще одна область, где выявлены несколько клад (А и D) и гаплогрупп (I и IV), – горы Загроса. Эти клад/гаплогруппы генетически максимально удалены друг от друга в выборке *L. m. media*. Обособление популяций, составляющих клад А, В и С, согласно датировкам Ф. Ахмад-

заде с соавт. [3], произошло около 650 тыс. лет назад, а D – около 1 млн лет назад. Вероятно, это было обусловлено климатическими флуктуациями плейстоцена, оказавшими влияние на ландшафты этой горной территории [38] и хронологию средней ящерицы. Кладе А генетически наиболее близки гаплотипы из Адыямана, входящие в кладу В. Это отражает тесную географическую связь юго-востока Малоазиатского (Восточный Тавр) и севера-запада Иранского (Загрос) нагорий [39]. Отметим, что популяции с этой территории (остан Керманшах) были обозначены Е.Ф. Шмидтлером [40] как самостоятельная форма “*L. m. media*–Zagrosform”. Полученные данные по определению границ между таксонами (по алгоритму ABGD) указывают на сравнительно высокий таксономический статус популяций, полученные последовательности от которых образовали в нашем исследовании кладу D.

Обращает на себя внимание отсутствие значимой генетической дистанции в выборке с оконечностей Большого Кавказа – точки 1–3 на рис. 1, расстояние между которыми по прямой составляет примерно 850 км. Это можно интерпретировать как свидетельство быстрого расселения вида с юга в двух направлениях в обход Главного Кавказского хребта. Однако нельзя исключать и существование единого циркумкавказского ареала. Он мог распасться во время последнего ледникового (гляциального) максимума (около 22 тыс. лет назад) либо сформироваться (вторично?) в ксеротермический период голоцена (около 4500–2500 лет назад) и разделиться в связи с последующей бореализацией климата. На такой сценарий может указывать наличие изолированных северокавказских популяций другого представителя средиземноморской зоогеографической группы рептилий – оливкового полоза, *Platyceps najadum* (Eichwald, 1831), имеющего в регионе чрезвычайно сходные с *L. media* очертания ареала [41]. Снижение в этот период области вероятного распространения (пригодности местообитаний) для обитания средней ящерицы на северо-западе Большого Кавказа было показано при ГИС-моделировании [42].

В дальнейшем необходимо включить в анализ ящериц с пограничных между кладами территорий, что позволит детализировать историю расселения вида. Так, новые находки *L. media* в Талышских горах [43] могут свидетельствовать либо о существовании здесь малочисленных реликтовых популяций, относящихся к кладе D, либо недавнем проникновении представителей клады С в горнолесной пояс Талыша; область распространения клады В может простираться до северо-западного склона Малого Кавказа в пределах Грузии. Наконец, глубокого анализа требует изучение филогеографии вида в Загросе – все еще малоизученной территории, выступающей с од-

ной стороны барьером между Месопотамией и нагорьем, с другой – коридором для распространения “северных” видов герпетофауны на юг региона [33].

Авторы признательны Н.Б. Ананьевой, О.А. Ермакову, В.Ф. Орловой, А.А. Кидову, Б.С. Туниеву и Н.А. Шаповалу за ценные консультации и помощь в работе.

Поддержание коллекции ZISP осуществлено в рамках проекта Министерства науки и высшего образования РФ (№ 075-15-2021-1069); исследования выполнены при финансовой поддержке гранта РНФ (№ 22-24-00079).

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Kornilios P., Thanou E., Lymberakis P. et al. A phylogenomic resolution for the taxonomy of Aegean green lizards // Zool. Scripta. 2019. V. 49. № 1. P. 14–27. <https://doi.org/10.1111/zsc.12385>
2. Godinho R., Crespo E.G., Ferrand N. Harris D.J. Phylogeny and evolution of the green lizards, *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences // Amphibia–Reptilia. 2005. V. 26. № 3. P. 271–285. <https://doi.org/10.1163/156853805774408667>
3. Ahmadzadeh F., Flecks M., Carretero M.A. et al. Rapid lizard radiation lacking niche conservatism: ecological diversification within a complex landscape // J. Biogeography. 2013. V. 40. № 9. P. 1807–1818. <https://doi.org/10.1111/jbi.12121>
4. Ahmadzadeh F., Flecks M., Rödder D. et al. Multiple dispersal out of Anatolia: biogeography and evolution of oriental green lizards // Biol. J. Linnean Society. 2013. V. 110. № 2. P. 398–408. <https://doi.org/10.1111/bij.12129>
5. Sagonas K., Poulakakis N., Lymberakis P. et al. Molecular systematics and historical biogeography of the green lizards (*Lacerta*) in Greece: Insights from mitochondrial and nuclear DNA // Mol. Phylogenet. Evol. 2014. V. 76. P. 144–154. <https://doi.org/10.1016/j.ympev.2014.03.013>
6. Kornilios P., Thanou E., Lymberakis P. et al. Genome-wide markers untangle the green-lizard radiation in the Aegean Sea and support a rare biogeographical pattern // J. Biogeography. 2019. V. 46. P. 552–567. <https://doi.org/10.1111/jbi.13524>
7. Доронин И.В., Доронина М.А. Обзор типовых экземпляров *Lacerta media* Lantz et Cyrén, 1920 (Sauria: Lacertidae) // Тр. Зоол. ин-та РАН. 2019. Т. 323. № 2. С. 85–92. <https://doi.org/10.31610/trudyzin/2019.323.2.85>
8. Nettmann H.-K., Rykena S. *Lacerta trilineata* Bedriaga 1886 – Riesensmaragdeidechse // Handbuch der Reptilien und Amphibien Europas. Bd. 2/I. Echsen (Sau-

- ria) II. (Lacertidae II: *Lacerta*) / Ed. Böhme W. Aula-Verlag (Wiesbaden), 1984. P. 100–128.
9. *Agasyan A., Tuniyev B., Cogalniceanu D. et al. Lacerta media* // The IUCN Red List of Threatened Species. 2009. (publ. online). <https://doi.org/10.2305/IUCN.UK.2009.RLTS.T164749A5922769>
  10. *Čerňanský A., Syromyatnikova E.V.* The first Miocene fossils of *Lacerta* cf. *trilineata* (Squamata, Lacertidae) with a comparative study of the main cranial osteological differences in green lizards and their relatives // PLoS One. 2019. V. 14. № 8. (publ. online). <https://doi.org/10.1371/journal.pone.0216191>
  11. *Мазанаева Л.Ф., Туниев Б.С.* Зоогеографический анализ герпетофауны Дагестана // Совр. герпетология. 2011. Т. 11. Вып. 1/2. С. 55–76.
  12. *Лотиев К.Ю., Туниев Б.С.* Возможные пути и этапы формирования герпетофауны Южной Осетии // Горные экосистемы и их компоненты: Материалы VI Всерос. конф. с международным участием, посвященной Году экологии в России и 100-летию заповедного дела в России. Нальчик: Алеф, 2017. С. 159–160.
  13. *Векуа А.К.* История животного мира — позвоночные // Грузия в антропогене. Развитие компонентов ландшафта и палеогеографические реконструкции. Тбилиси: Изд. Сакартвело, 1991. С. 308–381.
  14. *Туниев Б.С., Островских С.В.* Ящерица средняя *Lacerta media* (Lantz et Cyren, 1920) // Красная книга Краснодарского края. Животные. III изд. Краснодар: Адм. Краснодар. края, 2017. С. 492–493.
  15. *Мазанаева Л.Ф., Аскендеров А.Д.* Средняя ящерица *Lacerta media* (Lantz et Cyren, 1920) // Красная книга Республики Дагестан. Махачкала: Типография ИП Джамалудинов М.А., 2020. С. 532–533.
  16. *Туниев Б.С., Лотиев К.Ю.* Ящерица средняя *Lacerta media* Lantz et Cyren, 1920 // Красная книга Республики Южная Осетия. Нальчик: Издательство М. и В. Котляровых (ООО “Полиграфсервис и Т”), 2017. С. 226–227.
  17. Об утверждении Перечня объектов животного мира, занесенных в Красную книгу Российской Федерации. Приказ Минприроды РФ от 24.03.2020 № 162. Зарегистрировано в Минюсте РФ 02.04.2020 № 57940. <https://minjust.consultant.ru/documents/45937>
  18. *Aljanabi S.M., Martinez I.* Universal and rapid salt-extraction of high genomic DNA for PCR-based techniques // Nucl. Acids Res. 1997. V. 25. № 22. P. 4692–4693. <https://doi.org/10.1093/nar/25.22.4692>
  19. *Marzahn E., Mayer W., Joger U. et al.* Phylogeography of the *Lacerta viridis* complex: Mitochondrial and nuclear markers provide taxonomic insights // J. Zool. Syst. and Evol. Res. 2016. V. 54. № 2. P. 85–105. <https://doi.org/10.1111/jzs.12115>
  20. *Кукушкин О.В., Ермаков О.А., Иванов А.Ю. и др.* Филогеография прыткой ящерицы в Крыму по результатам анализа гена цитохрома *b*: древний рефугиум на полуострове, поздняя экспансия с севера и первые свидетельства гибридизации подвидов *Lacerta agilis tauridica* и *L. a. exigua* (Lacertidae: Sauria) // Тр. Зоол. ин-та РАН. 2020. Т. 324. № 1. С. 56–99. <https://doi.org/10.31610/trudyzin/2020.324.1.56>
  21. *Larsson A.* AliView: A fast and lightweight alignment viewer and editor for large data sets // Bioinformatics. 2014. V. 30. № 22. P. 3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>
  22. *Huelsenbeck J., Ronquist F.* MRBAYES: Bayesian inference of phylogenetic trees // Bioinformatics. 2001. V. 17. № 8. P. 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
  23. *Ronquist F., Huelsenbeck J.P.* MrBayes 3: Bayesian phylogenetic inference under mixed models // Bioinformatics. 2003. V. 19. № 12. P. 1572–1574. <https://doi.org/10.1093/bioinformatics/btg180>
  24. *Altekar G., Dwarkadas S., Huelsenbeck J.P., Ronquist F.* Parallel metropolis coupled markov chain Monte Carlo for Bayesian phylogenetic inference // Bioinformatics. 2004. V. 20. № 3. P. 407–415. <https://doi.org/10.1093/bioinformatics/btg427>
  25. *Kumar S., Stecher G., Li M. et al.* MEGA X: Molecular evolutionary genetics analysis across computing platforms // Mol. Biol. Evol. 2018. V. 35. № 6. P. 1547–1549. <https://doi.org/10.1093/molbev/msy096>
  26. *Nylander J.A.A.* Mr Modeltest version 2. Program distributed by the author. Evolutionary Biology Centre, Uppsala Univ., 2004.
  27. *Leigh J.W., Bryant D.* PopART: Full-feature software for haplotype network construction // Methods Ecol. Evol. 2015. V. 6. № 9. P. 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
  28. *Librado P., Rozas J.* DnaSP v5: A software for comprehensive analysis of DNA polymorphism data // Bioinformatics. 2009. V. 25. № 11. P. 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
  29. *Sarkar I.N., Egan M.G., Coruzzi G. et al.* Automated simultaneous analysis phylogenetics (ASAP): An enabling tool for phylogenomics // BMC Bioinformatics. 2008. V. 9. № 103. P. 1–5. <https://doi.org/10.1186/1471-2105-9-103>
  30. *Puillandre N., Lambert A., Brouillet S., Achaz G.* ABGD, Automatic Barcode Gap discovery for primary species delimitation // Mol. Ecol. 2012. V. 21. № 8. P. 1864–1877. <https://doi.org/10.1111/j.1365-294X.2011.05239.x>
  31. *Прохоров А.М.* (гл. ред.). Себелан // Большая Советская Энциклопедия (в 30 томах). Т. 23. Сафлор — Соан. Изд. 3-е. М.: БСЭ, 1976. С. 101.
  32. *Tuniyev B.S., Orlov N.L., Ananjeva N.B., Aghasyan A.L.* Snakes of the Caucasus: Taxonomic diversity, distribution, conservation. St. Petersburg; M.: KMK Scientific Press, 2019. 276 p.
  33. *Anderson S.C.* The Lizards of Iran // Society for the Study of Amphibians and Reptiles. Oxford: Ohio, 1999. V. 15. 442 p.
  34. *Kidov A.A., Litvinchuk S.N.* Distribution and conservation status of the Hircanian wood frog (*Rana pseudodalmatina*) in Azerbaijan // Rus. J. Herpetol. 2021. V. 28. № 2. P. 97–107. <https://doi.org/10.30906/1026-2296-2021-28-2-97-107>

35. *Cyrén O.* Klima und Eidechsenverbreitung. Eine Studie der geographischen Variation und Entwicklung einiger Lacerten, in sbesondereunter Berücksichtigung der klimatischen Faktoren // Meddel and enfran Göteborgs Musei Zoologiska Avdelning. 1924. V. 27. № 3. P. 1–82 + 13 p. plates.
36. *Даревский И.С.* Фауна пресмыкающихся Армении и ее зоогеографический анализ. Автореферат дис. ... канд. биол. наук. Ереван, 1957. 28 с.
37. *Мухелишвили Т.А.* О распространении полосатой ящерицы (*Lacerta strigata* Eichwald) и средней ящерицы (*Lacerta trilineata media* Lantz et Cyren) в Восточной Грузии // Сообщ. Акад. наук Груз. ССР. 1966. Т. XLIII. № 3. С. 753–756.
38. *Kehl M.* Quaternary climate change in Iran – the state of knowledge // *Erdkunde*. 2009. V. 63. № 1. P. 1–17. <https://doi.org/10.3112/ERDKUNDE.2009.01.01>
39. *Atalay I., Efe R., Soykan A.* Mediterranean Ecosystems of Turkey: Ecology of Taurus Mountains // *Environment and Culture in the Mediterranean Region* / Eds Efe, Cravins, Öztürk, Atalay. Newcastle, UK: Cambr. Scholars Publ., 2008. P. 3–37. <https://doi.org/10.5848/csp.1087.00001>
40. *Schmidler J.F.* Orientalische Smaragdei dechsen: 2. Über Systematik und Synökologie von *Lacerta trilineata*, *L. media* und *L. pamphylica* // *Salamandra*. 1986. V. 22. № 2/3. P. 126–146.
41. *Tuniyev B.S.* On the Mediterranean influence on the formation of herpetofauna of the Caucasian isthmus and its main xerophilous refugia // *Rus. J. Herpetology*. 1995. V. 2. № 2. P. 95–119. <https://doi.org/10.30906/1026-2296-1995-2-2-95-119>
42. *Доронин И.В., Мазанаева Л.Ф., Доронина М.А.* Использование ГИС-моделирования для анализа распространения средней ящерицы, *Lacerta media* Lantz et Cyren, 1920, на территории Дагестана (Россия) // *Тр. Зоол. ин-та РАН*. 2018. Т. 322. № 4. С. 463–480. <https://doi.org/10.31610/trudyzin/2018.322.4.463>
43. *Кидов А.А., Доронин И.В., Пыхов С.Г., Доронина М.А.* О новой находке *Lacerta media* Lantz et Cyren, 1920 (Reptilia, Lacertidae) в Тальше // *Совр. герпетология*. 2019. Т. 19. Вып. 1/2. С. 68–73. <https://doi.org/10.18500/1814-6090-2019-19-1-2-68-73>

## Cytochrome *b* Mitochondrial Gene Analysis-Based Phylogeography of a *Lacerta media* Lantz et Cyren, 1920 (Lacertidae: Sauria)

M. A. Doronina<sup>a</sup>, I. V. Doronin<sup>a, \*, \*\*</sup>, S. A. Lukonina<sup>b</sup>, L. F. Mazanaeva<sup>c</sup>, and A. V. Barabanov<sup>a</sup>

<sup>a</sup>Zoological Institute of the Russian Academy of Sciences, Saint Petersburg, 199034 Russia

<sup>b</sup>Penza State University, Penza, 440026 Russia

<sup>c</sup>Dagestan State University, Makhachkala, Dagestan, 367025 Russia

\*e-mail: igor.doronin@zin.ru

\*\*e-mail: ivdoronin@mail.ru

The analysis of *cytb* of mitochondrial DNA (1143 bp) in *Lacerta media* was carried out using 34 individuals from 19 localities. Special attention was paid to previously unexplored populations on the northern border of the range within the North Caucasus and Western Transcaucasia. Two types of dendrograms of phylogenetic relations (Bayesian and Maximum Likelihood) were reconstructed, showing identical topology. Phylogenetic reconstruction identifies four supported clades. Similarly, four haplogroups are distinguished on the median network of haplotypes. The populations of the eastern part of the Armenian Highlands and the Northern Zagros are distinguished at earlier stages of the process of phylogeny *L. m. media*. Subsequently, the populations of the central part of the Armenian Highlands, the Greater Caucasus, and, at the final stage, the Central Zagros and the eastern part of the Asia Minor Highlands become isolated. The data obtained do not confirm the scenario of dispersal of the nominative subspecies from west to east and northeast, proposed by Ahmadzadeh et al. Instead, we consider the opposite direction – from east to north, south, and west. Two geographical regions were identified where representatives of several clades/haplogroups live, is the Armenian Highlands and the Zagros Mountains.

**Keywords:** cytochrome *b*, *Lacerta media*, mitochondrial DNA, phylogeny, phylogeography.

## ИЗУЧЕНИЕ РОЛИ ПОЛИМОРФИЗМА ГЕНОВ РЕПАРАЦИИ ДНК В ФОРМИРОВАНИИ ПРЕДРАСПОЛОЖЕННОСТИ К РАЗВИТИЮ РАКА ЛЕГКОГО У ЖЕНЩИН

© 2022 г. Р. А. Титов<sup>1, \*</sup>, В. И. Минина<sup>1, 2, \*\*</sup>, А. В. Торгунакова<sup>1</sup>, В. Ю. Буслаев<sup>1</sup>,  
Е. Н. Воронина<sup>3</sup>, А. Ю. Просеков<sup>2</sup>, В. А. Титов<sup>4</sup>, А. Н. Глушков<sup>1</sup>

<sup>1</sup>Федеральный исследовательский центр угля и углехимии Сибирского отделения  
Российской академии наук, Кемерово, 650065 Россия

<sup>2</sup>Кемеровский государственный университет, Кемерово, 650043 Россия

<sup>3</sup>Институт химической биологии и фундаментальной медицины Сибирского отделения  
Российской академии наук, Новосибирск, 630090 Россия

<sup>4</sup>Кузбасский клинический онкологический диспансер, Кемерово, 650036 Россия

\*e-mail: ruslan-tito00@rambler.ru

\*\*e-mail: vminina@mail.ru

Поступила в редакцию 16.04.2021 г.

После доработки 30.07.2021 г.

Принята к публикации 24.08.2021 г.

Обследованы 522 женщины русской национальности, проживающие в Кемеровской области России, в том числе 273 больные раком легкого и 249 женщин близкого возраста, не имеющих признаков онкологических заболеваний. Проведен сравнительный анализ полиморфных вариантов генов репарации ДНК *APEX1 444T>G* (*rs1130409*), *XRCC1 1839G>A* (*rs25489*), *hOGG1 977C>G* (*rs1052133*), *XPB 2251T>G* (*rs13181*), *XPG 3310G>C* (*rs17655*), *XPC 2815A>C* (*rs2228001*) у больных раком легкого и индивидов, не имеющих онкологических заболеваний, проживающих в той же местности. Анализ однолокусных эффектов показал значимые связи между риском рака легкого и вариантами гена *XPC 2815A>C* (*rs2228001*) (OR = 0.56, CI: 0.39–0.81,  $p = 0.0018$ ) в общей группе, гена *APEX1 444T>G* (*rs1130409*) (OR = 0.15, CI: 0.03–0.67,  $p = 0.0027$ ) в группе курящих, генов *XPC 2815A>C* (*rs2228001*) (OR = 0.36, CI: 0.18–0.69,  $p = 0.0051$ ) и *hOGG1 977C>G* (*rs1052133*) (OR = 0.57, CI: 0.38–0.85,  $p = 0.0055$ ) в группе некурящих. MDR-анализ ген-генных взаимодействий показал, что гены *XPB 2251T>G* и *XPC 2815A>C*, *APEX1 444T>G* и *XPB 2251T>G* тесно взаимодействуют и взаимно усиливают риск развития рака легкого у женщин Западной Сибири.

**Ключевые слова:** рак легкого, полиморфизм генов *APEX1*, *XRCC1*, *hOGG1*, *XPB*, *XPG*, *XPC*.

**DOI:** 10.31857/S0016675822020138

Рак легкого (РЛ), или карцинома легкого, является наиболее распространенной формой онкологических патологий, с которой связан повышенный уровень заболеваемости и смертности среди населения во всем мире. Согласно данным статистики каждый год в мире регистрируют до 1.6 млн новых случаев заболеваний РЛ и около 1.2 млн случаев летальных исходов от этой патологии [1–3]. Известно, что женщины болеют РЛ в целом реже, чем мужчины, однако соотношение заболевших мужчин и женщин меняется с каждым годом, отмечается значительный рост именно женской заболеваемости РЛ [4, 5]. Многие авторы связывают это с действием факторов среды (все больше женщин курят, подвергаются воздействию промышленных и бытовых канцерогенов) и с большей чувствительностью организма жен-

щины к действию таких факторов. Воздействие женских половых гормонов нередко рассматривают как дополнительный фактор канцерогенного риска [6, 7].

Нарушения в системе репарации ДНК могут способствовать накоплению мутаций [8, 9]. Формирующийся в таких условиях высокий уровень нестабильности генома является одним из основных факторов риска возникновения злокачественных новообразований [10, 11]. Накопление критических мутаций в ДНК может последовательно вызывать изменения клеточного цикла и апоптоза с развитием злокачественных опухолей [12].

Генотоксичные соединения, присутствующие в большом количестве в окружающей среде промышленных регионов, индуцируют формирова-

**Таблица 1.** Характеристика изученных групп

Характеристика групп		Женщины больные РЛ	Здоровые женщины
Всего обследовано		273	249
Возраст, лет (среднее значение $\pm$ стандартное отклонение)		58.5 $\pm$ 7.6	54.0 $\pm$ 6.9
Статус курения, <i>n</i> (%)	курящие	72 (26.4)	46 (18.1)
	некурящие	201 (73.4)	202 (81.9)
Гистологические формы опухоли, <i>n</i> (%)	аденокарцинома	114 (42)	
	плоскоклеточный	65 (24)	
	мелкоклеточный	16 (6)	
	другие	77 (28)	
Наличие метастаз, <i>n</i> (%)	есть	160 (58.6)	
	нет	113 (41.4)	

ние соматических мутаций, ингибируют системы репарации ДНК и тем самым способствуют злокачественной трансформации клеток [13, 14]. В опубликованных нами ранее работах были показаны статистически значимые ассоциации вариантов генов ферментов репарации ДНК с риском развития РЛ у мужчин, проживающих в угледобывающем регионе [15, 16].

Вопрос о влиянии унаследованных вариантов генов ферментов репарации ДНК на риск возникновения РЛ у женщин малоизучен. В связи с этим целью настоящего исследования является изучение вклада полиморфных вариантов генов репарации ДНК *APEX1 444T>G (rs1130409)*, *XRCC1 1839G>A (rs25489)*, *hOGG1 977C>G (rs1052133)*, *XPB 2251T>G (rs13181)*, *XPG 3310G>C (rs17655)*, *XPC 2815A>C (rs2228001)* в формировании наследственной предрасположенности к РЛ у женщин, проживающих в промышленном угледобывающем регионе России.

## МАТЕРИАЛЫ И МЕТОДЫ

Были обследованы 522 женщины, принадлежащие к русской этнической группе, проживающие на территории Кемеровской области РФ. От каждого человека было получено информированное согласие на участие в исследовании. Исследование выполнено в соответствии с Хельсинкской декларацией Всемирной медицинской ассоциации: этические принципы проведения медицинских исследований с участием человека с поправками 2000 г. и “правилами клинической практики в Российской Федерации”, утвержденными Министерством здравоохранения Российской Федерации 19.06.2003 г. Проведение исследования было утверждено комитетом по этике Института экологии человека Федерального исследовательского центра угля и углехимии СО РАН.

Группы формировались по принципу “случай–контроль” с учетом возраста, пола, этнической принадлежности и влияния факторов среды. Критерии включения: русские, проживание в Кемеровской области с момента рождения, возраст старше 40 лет. Критерии исключения: аллергические, аутоиммунные, наследственные, инфекционные заболевания, родственники с онкозаболеваниями, ранее диагностированный рак в других органах. В первую группу вошли 273 впервые выявленные больные РЛ. Диагноз РЛ устанавливался по результатам клинического, эндоскопического и морфологического обследования на базе Кузбасского областного онкологического диспансера. У 125 человек рак легкого был диагностирован на I или II стадии, а у 148 – на III/IV стадии заболевания. Анализ гистологического материала после операций позволил установить точный патоморфологический диагноз каждого донора. Контрольную группу составили 249 неродственных женщин без онкозаболеваний и без патологии дыхательной системы в анамнезе, проживающих в той же местности. Все здоровые доноры, включенные в контрольную группу, не имели хронических заболеваний, не принимали препаратов с известным мутагенным действием и не проходили рентгенологических процедур в течение трех месяцев до участия в исследовании. Полное описание групп представлено в табл. 1.

ДНК выделяли из венозной крови по стандартной методике фенольно-хлороформной экстракции [17]. Клетки крови были выделены и лизированы, белковое содержимое клеток гидролизировано протеиназой К (SibEnzyme, Новосибирск, Российская Федерация). Для анализа нами были выбраны полиморфные локусы генов *APEX1 444T>G (rs1130409)*, *XRCC1 1839G>A (rs25489)*, *hOGG1 977C>G (rs1052133)*, *XPB 2251T>G (rs13181)*, *XPG 3310G>C (rs17655)*, *XPC 2815A>C*

**Таблица 2.** Последовательности праймеров для полиморфизма генов *APEX1*, *XRCC1*, *hOGG1*, *XPD*, *XPG*, *XPC*

Ген	Нуклеотидная замена, SNP	Локализация гена	Праймер (5' → 3')
<i>APEX1</i>	444T>G <i>rs1130409</i>	14q11.2-q12	attgaggtctccacacagcaca aattctgttctattctatagcgag
<i>XRCC1</i>	839G>A <i>rs25489</i>	19q13.2	tggggcctggattgctgggtctg cagcaccactaccacacctgaagg
<i>hOGG1</i>	977C>G <i>rs1052133</i>	3p26.2	ggaaggtgcttggggaat actgtcactagtctcaccag
<i>XPD (ERCC2)</i>	2251T>G <i>rs13181</i>	19q13.32	tcaaacatcctgtccctact ctgccgattaaaggctgtgga
<i>XPG (ERCC5)</i>	3310G>C <i>rs17655</i>	13q33	ttacgtctttgcgacaaattcatt cattaagatgaacttcagcat
<i>XPC</i>	2815A>C <i>rs2228001</i>	3p25	tcccattgagaagctgtgag ttcccattgagcagctgtgagc

(*rs2228001*). Частота редкого аллеля в популяциях европеоидов учитывалась по данным базы National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). Типирование полиморфных маркеров проводилось с помощью аллель-специфической ПЦР (ООО “Литех НПП”, Москва, Россия). Амплификацию проводили на термоциклере “Терцик” по программе, рекомендованной производителем наборов реагентов. Продукты ПЦР анализировали методом электрофореза в 3%-ном агарозном геле в присутствии бромиды этидия с последующей визуализацией фрагментов ДНК в УФ-свете.

Статистическую обработку результатов проводили с использованием пакетов прикладных программ: SNPstat (<http://bioinfo.iconcologia.net/SNPstats>), STATISTICA 10.0, MDR (<http://www.multifactor dimensionality reduction.org>). Проводили оценку частоты аллелей и генотипов; анализировали соответствие распределения частот генотипов равновесию Харди–Вайнберга; оценивали статистическую значимость различий между группами по частотам аллелей и генотипов для теста  $\chi^2$  на гомогенность выборок (статистически значимыми считали различия при  $p < 0.05$  с учетом поправки Бонферрони). Логистическую регрессию использовали для выявления ассоциации полиморфных локусов в различных моделях (аддитивной, доминантной, сверхдоминантной, рецессивной, лог-аддитивной) с коррекцией на возраст, статус курения. Для выбора лучшей модели использовали информационный критерий Акайке (AIC). Исследование роли межгенных взаимодействий проводили в программе MDR – Multifactor Dimensionality Reduction (<http://www.multifactor dimensionality reduction.org>).

## РЕЗУЛЬТАТЫ

Частоты генотипов и аллелей исследованных полиморфных локусов показали соответствие равновесию Харди–Вайнберга как в группе больных раком легкого, так и в группе здоровых (табл. 3).

При сравнении изученных групп жителей Кемеровской области больных РЛ и здоровых были выявлены отличия частот вариантов генов *APEX1* 444T>G (*rs1130409*), *hOGG1* 1977C>G (*rs1052133*), *XRCC1* 1839G>A (*rs25489*), *XPC* 2815A>C (*rs2228001*). С учетом поправки на множественные сравнения (поправка Бонферрони) статистически значимыми следует признать отличия распределений аллелей гена *APEX1* 444T>G и генотипов гена *XPC* 2815A>C.

Расчеты отношения шансов для различных моделей наследования (табл. 4) показали, что для гена *APEX1* 444T>G ассоциация с риском РЛ наиболее значимо проявлялась в общей группе (все обследованные) в лог-аддитивной модели наследования ( $OR_{adj} = 0.71$ , 95%CI: 0.54–0.92;  $p_{adj} = 0.009$ , AIC 678.2). Для группы некурящих женщин ассоциация значима в лог-аддитивной модели ( $OR_{adj} = 0.72$ , 95%CI: 0.54–0.96;  $p_{adj} = 0.026$ , AIC 557.7), а в группе курящих в рецессивной модели ( $OR_{adj} = 0.15$ , 95%CI: 0.03–0.67;  $p_{adj} = 0.002$ , AIC 152.8).

Для гена *XPC* 2815A>C наиболее значимые ассоциации проявлялись в доминантной модели наследования в общей группе ( $OR_{adj} = 0.56$ , 95%CI: 0.39–0.81;  $p_{adj} = 0.0018$ , AIC 675.2) и в доминантной модели в группе некурящих ( $OR_{adj} = 0.51$ , 95%CI: 0.18–0.69;  $p_{adj} = 0.005$ , AIC 548.8).

**Таблица 3.** Распределение частот генотипов и аллелей полиморфных локусов генов репарации ДНК в изученных группах женщин

Ген, нуклеотидная замена, SNP	Генотипы и аллели	Больные РЛ <i>n</i> (%)	Здоровые женщины <i>n</i> (%)	$\chi^2$	<i>p</i>	<i>p</i> <sub>cor</sub>
<i>APEX1</i> 444T>G ( <i>rs1130409</i> )	<i>TT</i>	74 (27.1)	93 (37.4)	8.7	0.0128	0.0768
	<i>TG</i>	140 (51.3)	123 (49.4)			
	<i>GG</i>	59 (21.6)	33 (13.2)			
	<i>T/G</i>	288 (52.7)/258 (47.3)	309 (62.1)/189 (37.9)	8.2	0.0029	0.0174
	<i>p</i> <sup>HWE</sup>	0.72	0.50			
<i>XRCC1</i> 839G>A ( <i>rs25489</i> )	<i>GG</i>	202 (74)	205 (82.3)	4.7	0.0961	0.5766
	<i>GA</i>	64 (23.4)	39 (15.7)			
	<i>AA</i>	7 (2.6)	5 (2.0)			
	<i>G/A</i>	468 (85.7)/78 (14.3)	449 (90.2)/49 (9.8)	4.4	0.0358	0.2148
	<i>p</i> <sup>HWE</sup>	0.46	0.07			
<i>hOGG1</i> 977C>G ( <i>rs1052133</i> )	<i>CC</i>	145 (53.1)	162 (65.1)	6.9	0.0311	0.1866
	<i>CG</i>	109 (39.9)	74 (29.7)			
	<i>GG</i>	9 (7)	13 (5.2)			
	<i>C/G</i>	399 (73.1)/147 (26.9)	398 (79.9)/100 (20.1)	6.3	0.0115	0.069
	<i>p</i> <sup>HWE</sup>	0.88	0.24			
<i>XPD</i> 2251T>G ( <i>rs13181</i> )	<i>TT</i>	128 (46.9)	125 (50.2)	0.6	0.7482	4.4892
	<i>TG</i>	121 (44.3)	101 (40.6)			
	<i>GG</i>	24 (8.8)	23 (9.2)			
	<i>T/G</i>	377 (69.0)/169 (31.0)	351 (70.5)/147 (29.5)	0.2	0.6629	3.9774
	<i>p</i> <sup>HWE</sup>	0.67	0.76			
<i>XPG</i> 3310G>C ( <i>rs17655</i> )	<i>GG</i>	143 (52.4)	145 (58.2)	2.0	0.3629	2.1774
	<i>GC</i>	111 (40.7)	85 (34.1)			
	<i>CC</i>	19 (7)	19 (7.6)			
	<i>G/C</i>	397 (72.7)/149 (27.3)	375 (75.3)/123 (24.7)	0.7	0.3775	2.265
	<i>p</i> <sup>HWE</sup>	0.76	0.23			
<i>XPC</i> 2815A>C ( <i>rs2228001</i> )	<i>AA</i>	91 (33.3)	117 (46.9)	12.8	0.0016	0.0096
	<i>AC</i>	130 (47.6)	107 (42.9)			
	<i>CC</i>	52 (19.1)	25 (10.0)			
	<i>A/C</i>	312 (57.1)/234 (42.9)	341 (68.5)/157 (31.5)	13.7	0.0002	0.0012
	<i>p</i> <sup>HWE</sup>	0.71	1.0			

Примечание. *p*<sup>HWE</sup> – значимость отличий распределения частот генотипов от равновесия Харди–Вайнберга, *p* – значимость отличий частоты встречаемости генотипа в группе больных и здоровых, критерий  $\chi^2$  с поправкой Йетса, *p*<sub>cor</sub> – значимость отличий с учетом поправки Бонферрони.

Для гена *hOGG1 977C>G* наиболее значимые ассоциации проявлялись в доминантной модели наследования в общей группе ( $OR_{adj} = 0.62$ , 95%CI: 0.43–0.89;  $p_{adj} = 0.01$ , AIC 678.4;  $p_{cor} = 0.01$ ) и в доминантной модели в группе некурящих ( $OR_{adj} = 0.57$ , 95%CI: 0.38–0.85;  $p_{adj} = 0.0055$ , AIC 555).

Для гена *XRCC1 1839G>A* наиболее значимые ассоциации проявлялись в доминантной модели наследования в общей группе ( $OR_{adj} = 0.55$ , 95%CI: 0.36–0.87;  $p_{adj} = 0.008$ , AIC 678.1), в сверхдоминантной модели в группе некурящих ( $OR_{adj} = 0.60$ , 95%CI: 0.36–0.99;  $p_{adj} = 0.042$ , AIC 558.6).

**Таблица 4.** Результаты анализа ассоциаций полиморфных локусов генов-кандидатов с риском РЛ у женщин

Ген, нуклеотидная замена, SNP	Группа	Модель	OR <sub>adj</sub>	<i>p</i>	<i>p</i> <sub>cor</sub>	AIC
<i>APEX1</i> 444T>G (rs1130409)	Общая	Лог-аддитивная	0.71 (0.54–0.92)	0.0095	0.057	678.2
	Курящие	Рецессивная	0.15 (0.03–0.67)	0.0027	0.0162	152.8
	Некурящие	Лог-аддитивная	0.72 (0.54–0.96)	0.026	0.156	557.7
<i>XRCC1</i> 839G>A (rs25489)	Общая	Доминантная	0.55 (0.36–0.87)	0.0082	0.0492	678.1
	Курящие	Лог-аддитивная	0.48 (0.20–1.16)	0.084	0.504	158.8
	Некурящие	Сверхдоминантная	0.60 (0.36–0.99)	0.042	0.252	558.6
<i>hOGG1</i> 977C>G (rs1052133)	Общая	Доминантная	0.62 (0.43–0.89)	0.01	0.06	678.4
	Курящие	Доминантная	0.71 (0.33–1.52)	0.37	2.22	161
	Некурящие	Доминантная	0.57 (0.38–0.85)	0.0055	0.033	555
<i>XPB</i> 2251T>G (rs13181)	Общая	Сверхдоминантная	0.82 (0.57–1.19)	0.3	1.8	683.9
	Курящие	Лог-аддитивная	1.25 (0.70–2.23)	0.45	2.7	161.2
	Некурящие	Доминантная	0.80 (0.54–1.18)	0.25	1.5	561.4
<i>XPG</i> 3310G>C (rs17655)	Общая	Сверхдоминантная	0.85 (0.58–1.23)	0.38	2.28	684.2
	Курящие	Лог-аддитивная	1.38 (0.80–2.40)	0.25	1.5	160.5
	Некурящие	Сверхдоминантная	0.66 (0.44–0.99)	0.046	0.276	558.7
<i>XPC</i> 2815A>C (rs2228001)	Общая	Доминантная	0.56 (0.39–0.81)	0.0018	0.0108	675.2
	Курящие	Сверхдоминантная	1.88 (0.88–4.00)	0.1	0.6	159.1
	Некурящие	Доминантная	0.36 (0.18–0.69)	0.0051	0.0306	548.8

Примечание. OR<sub>adj</sub> – показатель отношения шансов для редкого аллеля с поправкой на возраст, курение; *p* – значимость различий между группами; AIC – критерий Акайке.

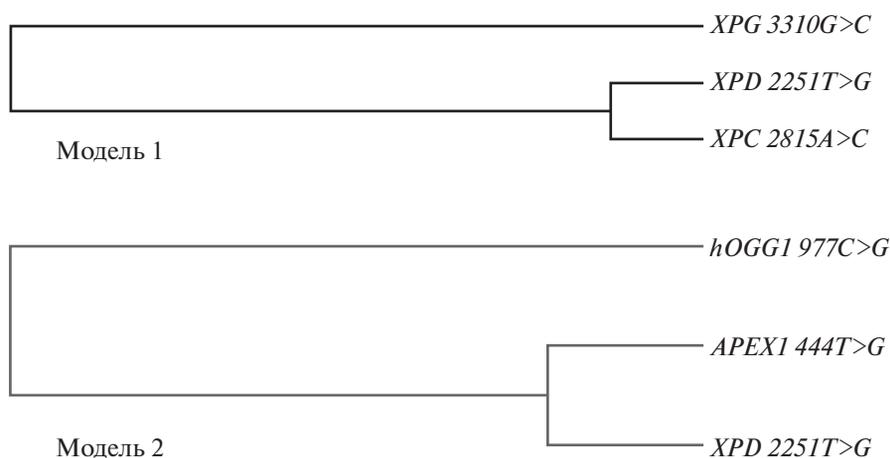
При изучении отдельных гистологических форм РЛ статистически значимая ассоциация была выявлена только в группе женщин больных немелкоклеточным раком легких. Ассоциация с развитием данного типа РЛ была выявлена для вариантов гена *XPC 2815A>C* в лог-аддитивной модели (OR = 1.56, 95%CI: 1.17–2.07; *p* = 0.002).

В результате анализа полиморфизма генов репарации у женщин в связи с риском развития РЛ с метастазами в отдаленные органы были выявлены ассоциации с вариантами гена *XRCC1 839G>A* в доминантной модели (OR = 2.30, 95%CI: 1.39–3.79; *p* = 0.0011), гена *XPC 2815A>C* в лог-аддитивной модели (OR = 1.74, 95%CI: 1.28–2.37; *p* = 0.0004).

Важным этапом ассоциативных исследований является выявление взаимодействий между генами

с целью определения комбинаций полиморфных локусов, которые имеют наибольшую патогенетическую значимость для развития заболевания. В результате анализа межгенных взаимодействий методом MDR были определены две трехлокусные модели, детерминирующие риск развития РЛ у женщин (табл. 5).

Первая модель включала в себя гены *XPB 2251T>G (rs13181)*, *XPG 3310G>C (rs17655)*, *XPC 2815A>C (rs2228001)*. Кластерный анализ позволил установить тесное взаимодействие и взаимное усиление эффектов (синергизм) между полиморфными локусами *XPB 2251T>G* и *XPC 2815A>C*, тогда как эффект гена *XPG 3310G>C* не зависел от других генов, входящих в модель. При этом наибольший вклад в развитие РЛ у женщин опреде-



**Рис. 1.** Дендрограмма межгенных взаимодействий у женщин больных РЛ. Короткие линии указывают на сильное взаимодействие генных локусов, длинные — на слабую связь; черным цветом указывается синергизм, т.е. взаимное усиление эффектов между локусами, серым — дублирование эффектов между локусами.

лен для гена *XPC 2815A>C* ( $H$  (энтропия) = 1.59%) (рис. 1).

Вторая модель ген-генного взаимодействия, ассоциированная с развитием РЛ у женщин, включала полиморфные варианты генов *APEX1 444T>G* (*rs1130409*), *hOGG1 977C>G* (*rs13181*), *XPD 2251T>G* (*rs13181*). Был выявлен кластер тесно взаимодействующих генов *APEX1 444T>G* и *XPD 2251T>G*, эффекты которых синонимичны. Данный кластер образовывал единый блок с геном *hOGG1 977C>G* с дублирующим эффектом данных генов (рис. 1). В данной модели наибольший вклад вносил ген-триггер *APEX1 444T>G* ( $H = 1.54\%$ ).

## ОБСУЖДЕНИЕ

Современные эпидемиологические данные указывают на то, что риск формирования РЛ среди женщин увеличивается и к 2030 г. заболевае-

мость РЛ превысит показатели рака молочной железы в большинстве развитых стран [18].

Известно, что курение является наиболее важным фактором риска развития РЛ, но за последние 30 лет заболеваемость и смертность от РЛ увеличивается и у некурящих женщин [19, 20]. Заболеваемость РЛ у некурящих женщин может быть связана как с генетическими факторами, так и с влиянием окружающей среды [21]. Сегну с коллегами обнаружили, что женщины с РЛ имели значительно меньшее воздействие табачного дыма, чем мужчины, и средний возраст женщин с диагностированным РЛ был ниже [22].

По мнению некоторых авторов, такие факторы как выбросы промышленных предприятий, кипящие растительные масла, пассивное курение могут выступать в качестве факторов риска РЛ для некурящих пациентов [23]. Некоторые канцерогены, похожие на табачный дым, были идентифицированы в нагретых маслах, включая поли-

**Таблица 5.** Анализ межгенных взаимодействий при формировании РЛ у женщин

Сочетания полиморфных локусов в модели	Тр. Val. Acc.	Test. Val. Acc.	Sign Test ( $p$ )	Se	Sp	CVC	Pre
<i>XPD 2251T&gt;G</i> <i>XPG 3310G&gt;C</i> <i>XPC 2815A&gt;C*</i>	0.6141	0.5724	0.0018	0.6667	0.4615	10/10	0.6977
<i>APEX1 444T&gt;G</i> <i>hOGG1 977C&gt;G</i> <i>XPD 2251T&gt;G*</i>	0.6038	0.5500	0.0001	0.4840	0.7192	10/10	0.6741

Примечание. Тр.Val.Acc. — тренировочная сбалансированная точность. Test.Val.Acc. — тестируемая сбалансированная точность; Sign Test ( $p$ ) — тест на значимость; Se — чувствительность; Sp — специфичность; CVC — повторяемость результата; Pre (Precision) — точность модели; \* алгоритм полного поиска (Exhaustive search algorithm).

циклические ароматические углеводороды, ароматические амины и нитрополициклические ароматические углеводороды [24]. Ряд авторов показали, что воздействие нагретых паров растительного масла может быть значимым фактором риска РЛ для некурящих женщин [25–27].

По данным литературы известно, что полиморфизм генов, кодирующих белки, участвующие в репарации ДНК, может быть значимо связан с риском развития немелкоклеточной формы РЛ у женщин [28, 29]. В результате проведенного настоящего исследования были выявлены ассоциации между унаследованными вариантами ряда генов ферментов репарации ДНК и риском развития РЛ у женщин, проживающих в Кузбассе. Ранее исследования, проводившиеся нами в смешанной группе индивидов (мужчин и женщин), проживающих в той же местности, показали другие результаты [16]. Была показана значимость вариантов гена *XPD 2251T>G (rs13181)* в подгруппе мужчин, *APEX1 444T>G (rs1130409)* – у курильщиков, *NBS1 553C>G (rs1805794)* – у работников угольной промышленности. В подгруппе женщин были выявлены ассоциации лишь с вариантами гена *ATM 557G>A (rs1801516)*, а по генам *APEX1 444T>G (rs1130409)*, *hOGG1 977C>G (rs1052133)*, *XPC 2815A>C (rs2228001)* статистически значимых ассоциаций выявлено не было. Однако объем выборки женщин в данном исследовании был очень мал: 41 человек больных РЛ и 72 здоровые женщины. Конечно, этот результат стоит рассматривать лишь как предварительный, требующий верификации. В настоящем исследовании выборка женщин была существенно увеличена (суммарно  $n = 522$ ), что позволило получить новые данные и выявить ассоциации вариантов генов *APEX1 444T>G (rs1130409)*, *hOGG1 977C>G (rs1052133)*, *XPC 2815A>C (rs2228001)* с риском формирования РЛ у женщин, проживающих в промышленном регионе.

Как известно, продукты генов *APEX1 444T>G* и *hOGG1 977C>G* играют важную роль в реализации механизмов эксцизионной репарации оснований (BER-base nucleotide excision repair) [30, 31]. Ген *APEX1* локализован в локусе 14q11.2-q12, кодирует специализированный фермент – апуриновую/апириимидиновую (AP)-эндонуклеазу, удаляющую из ДНК AP-сайты, которые способны нарушать процессы репликации ДНК и значительно увеличивать вероятность гибели клетки. Полиморфный вариант гена *APEX1 444T>G (rs1130409)*, несущий трансверсию тимина на гуанин ( $T \rightarrow G$ ) в позиции 444, приводит к замене аминокислоты аспарагин (Asp) на глутаминовую кислоту (Glu) в положении 148. Ранее было выявлено, что аллель *G* кодирует фермент со сниженной активностью

эндонуклеазы и сниженной эффективностью ДНК-связывания [32]. Генотип *GG* также ассоциирован с большей задержкой G2-фазы клеточного цикла [33]. Значимость полиморфизма данного гена достаточно много изучали у больных РЛ, но результаты оказались противоречивы [34, 35]. В немецкой популяции вариант *Glu148Glu* имел защитный эффект (OR = 0.77, 95%CI: 0.51–1.16) [36], в то время как у китайцев риск развития РЛ оказался в 2 раза выше при носительстве варианта *Glu148Glu*, чем *Asp148Asp* (OR = 2.13, 95%CI: 0.96–4.74) [37]. Риск развития РЛ у женщин в нашем исследовании был связан с наличием редкого аллеля *G*. Ассоциация с РЛ установлена в рецессивной модели в группе курящих женщин, а также в лог-аддитивной модели в группе некурящих и в общей выборке. Ранее ассоциацию вариантов гена *APEX1 444T>G (rs1130409)* и РЛ отмечали только у курильщиков [16], проживающих в угольном регионе, но доля женщин в обследованной выборке была невелика (16.7%).

Ген *hOGG1* (human 8-oxoguanine DNA glycosylase) кодирует ключевой фермент BER, удаляющий из ДНК остатки 8-оксогуанина, образующегося под действием активных форм кислорода. Обнаружены два типа оксогуанин-гликозилаз длиной 345 и 424 аминокислоты ( $\alpha$ -*hOGG1* и  $\beta$ -*hOGG1*). Мишенью этих белков являются ядро и митохондрии, причем ядерная  $\alpha$ -*hOGG1* вырезает охoG и 2'-6-диамино-4-гидрокси-5-N-метилформамид-пиримидин (Me-FapyGua). Полиморфный вариант гена *hOGG1 977C>G (rs1052133)*, несущий трансверсию цитозина на гуанин ( $C \rightarrow G$ ) в позиции 977, приводит к замене аминокислоты серин (Ser) на цистеин (Cys) в положении 326. Исследования показали, что вариант аллеля *G* гена *hOGG1 977C>G (rs1052133)* связан с более низкой активностью фермента [38], что способствует накоплению окислительных аддуктов (8-оксо-G) и может увеличить скорость мутаций [39].

В ряде исследований было выявлено влияние варианта *Ser326Cys* на риск развития РЛ. Kohno соавт. при исследовании полиморфного варианта *hOGG1 977C>G (rs1052133)* обнаружили, что замена *Ser326Cys* в составе белка способствует уменьшению возможности репарации 8-оксогуанина [40]. Мета-анализ полиморфизма *hOGG1 977C>G (rs1052133)* с использованием данных 18 исследований указал на возможность вклада полиморфных вариантов *Ser326Cys* в повышение риска развития немелкоклеточной формы РЛ в азиатской популяции [41]. Также положительную корреляцию между гомозиготным вариантом *Cys326Cys* гена *hOGG1 977C>G (rs1052133)* и развитием РЛ показали исследования, которые были проведены на курильщиках [42, 43]. Однако есть результаты, демонстриро-

вавшие противоположные результаты [44, 45]. Причины различий результатов не вполне ясны, но они могут быть связаны, в том числе, и со спецификой ген-генных и ген-средовых взаимодействий.

Ген *XPC* расположен в локусе 3p25 и содержит 16 экзонов и 15 интронов. Кодируемый им белок играет значимую роль в механизмах эксцизионной репарации нуклеотидов ДНК (NER – nucleotide excision repair). Биологическая роль *XPC* состоит в том, что этот белок объединяется с HR23Вс образованием *XPC*–HR23В комплекса, являющегося самым ранним детектором повреждения, который запускает репарацию генома в целом [46, 47]. Комплекс *XPC*–HR23В после связывания с повреждением индуцирует специфическое конформационное изменение (включая локальное раскрытие ДНК), которое затем инициирует другие факторы, такие как ХРА и РРА. Известно, что аллель *C* гена *XPC 2815A>C (rs2228001)* ассоциируется с уменьшением репаративной способности ДНК [48, 49]. По данным литературы известна роль полиморфных вариантов этого гена в развитии злокачественных новообразований. Ну с соавт. выявили ассоциацию гетерозиготного и гомозиготного генотипов по минорному аллелю локуса Lys939Gln с риском развития РЛ в Китае [50]. Была показана значимая ассоциация полиморфизма Gln939Gln гена *XPC 2815A>C (rs2228001)* и риска развития РЛ у женщин в работе Letkova с соавт. [51].

В последние годы стало известно, что многие гены системы репарации ДНК играют важную роль в метастазировании опухолей [52]. При исследовании генетического полиморфизма у женщин больных РЛ с метастазами нами были выявлены ассоциации метастатических форм РЛ с вариантами генов *XRCC1 1839G>A (rs25489)* и *XPC 2815A>C (rs2228001)*.

Ген *XRCC1* (X-ray cross-complementing group I, локус 19q13.2) кодирует белок, который является интегральным регулятором эксцизионной репарации оснований (BER) [53]. Белок является каркасным с важными участками в N-концевой области: домен 1 BRCT и домен 2 BRCT. Эти домены образуют комплекс как минимум с тремя различными ферментами: поли-АДФ-рибозной полимеразой (ПАРП), ДНК-лигазой III и ДНК-полимеразой  $\beta$  [54]. Данная система обеспечивает защиту клетки от негативного воздействия агентов, модифицирующих азотистые основания ДНК и разрушающих ее сахарофосфатный остов. К таковым относятся разнообразные экзогенные и эндогенные генотоксические факторы [55]. Полиморфные варианты в этих областях могут влиять на связывание этих доменов с соответствующими ферментами, таким образом нарушая путь BER. Было уста-

новлено, что полиморфный вариант Arg280His гена *XRCC1 1839G>A (rs25489)* влияет на эффективность репарации ДНК и ассоциирован с повышенным риском РЛ [56–58].

При разделении по стадиям заболевания больных РЛ женщин не было отмечено ассоциаций с генетическим полиморфизмом. Это может быть обусловлено пока недостаточным объемом выборки. В целом данный вопрос малоизучен. Несколько исследований показали, что у женщин с РЛ прогноз лучше, чем у мужчин, вне зависимости от стадии заболевания [59–61].

В рамках настоящего исследования проведено моделирование межгенных взаимодействий методом MDR. Для оценки межгенных взаимодействий с помощью MDR-анализа в отношении риска развития РЛ у женщин использовали алгоритм всестороннего поиска (Exhaustive search algorithm), который оценивал все возможные комбинации исследованных полиморфных локусов. В результате анализа были выявлены наиболее информативные модели межгенных взаимодействий, детерминирующих формирование РЛ у женщин: *XPB 2251T>G (rs13181)*, *XPG 3310G>C (rs17655)*, *XPC 2815A>C (rs2228001)* и *APEX1 444T>G (rs1130409)*, *hOGG1 977C>G (rs13181)*, *XPB 2251T>G (rs13181)*.

Первая модель межгенных взаимодействий включала в себя гены *XPB 2251T>G (rs13181)*, *XPG 3310G>C (rs17655)*, *XPC 2815A>C (rs2228001)*. Между генами *XPB* и *XPC* было выявлено взаимно усиливающее действие (синергизм), тогда как ген *XPG* представлял собой локус с независимым эффектом.

Ген *XPB* (excision repair cross-complementing rodent repair deficiency, complementation group 2 [Xeroderma pigmentosum group D]) кодирует АТФ-независимую хеликазу. Это ключевой белок BER, который узнает и исправляет различные мутации (сшивки оснований, тиминовые димеры, ДНК-аддукты, окислительные повреждения ДНК и др.), образующиеся, например, после УФ-облучения или оксидативного стресса. В составе транскрипционного комплекса хеликазы *XPB* раскручивает цепь ДНК, обеспечивая доступ эндонуклеаз к поврежденному участку [62]. Полиморфный вариант гена *XPB 2251T>G (rs13181)*, несущий трансверсию тимина на гуанин ( $T \rightarrow G$ ) в позиции 2251, приводит к замене аминокислоты лизин (Lys) на глутамин (Gln) в положении 751. Аллель *G* идентифицируется с более высоким уровнем аддуктов ДНК, низкой репарационной способностью, увеличением частоты хроматидных разрывов [63, 64].

Ген *XPG* кодирует белок, представляющий собой эндонуклеазу, которая вырезает поврежденный

участок ДНК. *XPG* взаимодействует с комплексом *TFIIH* и обеспечивает прикрепление хеликаз и правильное расплетание цепей ДНК, делая их доступными для действия эндонуклеаз. Нарушение взаимодействия *XPG–TFIIH* приводит к диссоциации хеликаз и нарушению белкового комплекса, прерывая процесс репарации [65].

В нашем исследовании, исходя из таблиц сопряженности, было выявлено четыре варианта наиболее рискованных комбинаций, включающих в себя *TT* (или *TG*) гена *XPB 2251T>G (rs13181)*, *GG* (или *GC*) гена *XPG 3310G>C (rs17655)* и минорный вариант *CC* гена *XPC 2815A>C (rs2228001)* (53 человека). Также сочетаниями риска развития РЛ у женщин являлись комбинации *TG* (или *GG*) гена *XPB 2251T>G (rs13181)*, *GG* (или *GC*) гена *XPG 3310G>C (rs17655)* и гетерозиготного генотипа *AC* гена *XPC 2815A>C (rs2228001)* (89 человек).

Вторую модель межгенных взаимодействий, связанную с развитием РЛ у женщин, составили полиморфные локусы *APEX1 444T>G (rs1130409)*, *hOGG1 977C>G (rs13181)*, *XPB 2251T>G (rs13181)*. Между генами *APEX1 444T>G* и *XPB 2251T>G* было выявлено сильное взаимодействие, но их влияние дублировалось, а ген *hOGG1 977C>G* образовывал единый кластер с дублирующим эффектом данных генов.

Наиболее рискованными в нашем исследовании являлись пять комбинаций генотипов, включающих в себя гетерозиготный (*CG*) или минорный (*GG*) генотипы гена *hOGG1 977C>G (rs13181)*. Из них три содержали гетерозиготный генотип *CG* гена *hOGG1 977C>G (rs13181)*, гетерозиготный генотип *TG* гена *XPB 2251T>G (rs13181)* и все варианты генотипов гена *APEX1 444T>G (rs1130409)* (64 человека). Одна из рискованных комбинаций состояла из минорного варианта *GG* гена *APEX1 444T>G (rs1130409)*, гетерозиготного генотипа *CG* гена *hOGG1 977C>G (rs13181)* и мажорного генотипа *TT* гена *XPB 2251T>G (rs13181)* (11 человек). Шесть женщин больных РЛ имели рискованную комбинацию, содержащую минорный генотип *GG* гена *hOGG1 977C>G (rs13181)*, гетерозиготный генотип *TG* гена *APEX1 444T>G (rs1130409)* и генотип *TT* гена *XPB 2251T>G (rs13181)*, тогда как в группе здоровых женщин данного варианта не было выявлено.

Таким образом, в нашем исследовании были найдены новые значимые маркеры риска возникновения РЛ, а также определены две трехлокусные модели взаимодействия генов репарации ДНК, детерминирующие риск развития РЛ у женщин: *XPB 2251T>G (rs13181)*, *XPG 3310G>C (rs17655)*, *XPC 2815A>C (rs2228001)* и *APEX1*

*444T>G (rs1130409)*, *hOGG1 977C>G (rs13181)*, *XPB 2251T>G (rs13181)*.

Исследование поддержано Государственным заданием на 2019–2021 гг. № 0352-2019-0011.

От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национально-го комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Spaggiari L., Leo F., Veronesi G. et al. Superior vena cava resection for lung and mediastinal malignancies: a single-center experience with 70 cases // *Ann. Thorac. Surg.* 2007. V. 83. № 1. P. 223–229.
2. Sihoe A.D. Reasons not to perform uniportal VATS lobectomy // *J. Thorac. Dis.* 2016. V. 8. P. 333–343.
3. Состояние онкологической помощи населению России в 2019 году / Под ред. Каприна А.Д., Старинского В.В., Шахзадовой А.О. М.: МНИОИ им. П.А. Герцена, 2020. 239 с.
4. Martín-Sánchez J.C., Lunet N., González-Marrón A. et al. Projections in breast and lung cancer mortality among women: A Bayesian analysis of 52 countries worldwide // *Cancer Res.* 2018. V. 78. № 15. P. 4436–4442. <https://doi.org/10.1158/0008-5472.CAN-18-0187>
5. Sher D.J., Fidler M.J., Liptay M.J. et al. Comparative effectiveness of neoadjuvant chemoradiotherapy versus chemotherapy alone followed by surgery for patients with stage IIIA non-small cell lung cancer // *Lung Cancer.* 2015. V. 88. № 3. P. 267–274.
6. Nga M.E., Mohd Omar M.F., Lim D.G. et al. Using immunohistochemistry to evaluate protein expression levels of female sex hormone receptors (ER, PR) and epidermal growth factor receptor family members (EGFR, HER2) in East Asian non-small cell lung cancers (NSCLC) // *J. Thorac. Oncology.* 2007. V. 2. № 8. P. 502–503. <https://doi.org/10.1097/01.jto.0000283496.66021.67>
7. Stapelfeld C., Neumann K.T., Maser E. Different inhibitory potential of sex hormones on NNK detoxification *in vitro*: A possible explanation for gender-specific lung cancer risk // *Cancer Letters.* 2017. V. 405. P. 120–126. <https://doi.org/10.1016/j.canlet.2017.07.016>
8. Hoeijmaker J.H. Genome maintenance mechanisms for preventing cancer // *Nature.* 2001. V. 411. P. 366–374.
9. Khanna K.K., Jackson S.P. DNA double-strand breaks: signaling, repair and the cancer connection // *Nat. Genet.* 2001. V. 27. P. 247–254.

10. Jackson S.P., Bartek J. The DNA-damage response in human biology and disease // *Nature*. 2009. V. 461. P. 1071–1078.
11. Hartlerode A.J., Scully R. Mechanisms of double-strand break repair in somatic mammalian cells // *Biochem. J.* 2009. V. 423. № 2. P. 157–168.
12. Paris C., Bioucas-Dias J., Bruzzone L. A hierarchical approach to superresolution of multispectral images with different spatial resolutions // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2017. P. 2589–2592.
13. Catana A., Popp R.A., Pop M. et al. Genetic polymorphism of DNA repair gene ERCC2/XPD (Arg156Arg) (A22541C) and lung cancer risk in Northern Romania // *Revista Română de Medicină de Laborator*. 2012. V. 20. № 2. P. 157–161.
14. Akhmadishina L.Z., Giliyazova I.R., Kutlyeva L.R. et al. DNA repair XPC1 and XPD genes polymorphism as associated with the development of bladder cancer and renal cell carcinoma // *Rus. J. Genet.* 2014. V. 50. № 4. P. 481–490.  
<https://doi.org/10.1134/S1022795414040024>
15. Bakanova M.L., Minina V.I., Savchenko Y.A. et al. Links between DNA-repair gene polymorphisms and chromosomal aberrations in lung-cancer patients // *Mol. Genet. Microbiol. Virol.* 2013. V. 28. P. 133–136.
16. Minina V.I., Bakanova M.L., Soboleva O.A. et al. Polymorphisms in DNA repair genes in lung cancer patients living in a coal-mining region // *Eur. J. Cancer Prev.* 2019. V. 28. № 6. P. 522–528.
17. Sambrook J., Fritsch E.F., Maniatis T. *Molecular Cloning. A Laboratory Manual*. N.Y.: Cold Spring Harbor Lab., 1989. 1659 p.
18. Martín-Sánchez J.C., Lunet N., González-Marrón A. et al. Projections in breast and lung cancer mortality among women: A Bayesian analysis of 52 countries worldwide // *Cancer Res.* 2018. V. 78. № 15. P. 4436–4442.
19. Thomas L., Doyle L.A., Edelman M.J. Lung cancer in women: Emerging differences in epidemiology, biology, and therapy // *Chest*. 2005. V. 128. P. 370–381.
20. Belani C.P., Marts S., Schiller J. et al. Women and lung cancer: Epidemiology, tumor biology, and emerging trends in clinical research // *Lung Cancer*. 2007. V. 55. P. 15–23.
21. Matullo G., Dunning A.M., Guarrera S. et al. DNA repair polymorphisms and cancer risk in non-smokers in a cohort study // *Carcinogenesis*. 2006. V. 27. P. 997–1007.
22. Cerny D., Cerny T., Ess S. et al. Lung cancer in the Canton of St. Gallen, Eastern Switzerland: sex-associated differences in smoking habits, disease presentation and survival // *Onkologie*. 2009. V. 32. P. 69–73.
23. Li T., Huang H., Liao D. et al. Genetic polymorphism in HLA-G 3'UTR 14-bp ins/del and risk of cancer: A metaanalysis of case-control study // *Mol. Genet. Genomics*. 2015. V. 290. P. 1235–1245.
24. Chiang T.A., Wu P.F., Ko Y.C. Identification of carcinogens in cooking oil fumes // *Environ Res.* 1999. V. 81. № 1. P. 18–22.
25. Li M., Yina Z., Guan P. et al. XRCC1 polymorphisms, cooking oil fume and lung cancer in Chinese women nonsmokers // *Lung Cancer*. 2008. V. 62. P. 145–151.
26. Ko Y., Cheng L., Lee C. Chinese food cooking and lung cancer in women nonsmokers // *Am. J. Epidemiol.* 2000. V. 151. P. 140–147.
27. Xue X., Yin Z., Lu Y. et al. The joint effect of hOGG1, APE1, and ADPRT polymorphisms and cooking oil fumes on the risk of lung adenocarcinoma in Chinese non-smoking females // *PLoS One*. 2013. V. 8. P. 71157.
28. Su Y., Zhang H., Xu F. et al. DNA repair gene polymorphisms in relation to non-small cell lung cancer survival // *Cell Physiol. Biochem*. 2015. V. 36. P. 1419–1429.
29. Xu W., Liu D., Yang Y. et al. Association of CHEK2 polymorphisms with the efficacy of platinum-based chemotherapy for advanced non-small-cell lung cancer in Chinese never-smoking women // *J. Thorac. Dis.* 2016. V. 8. № 9. P. 2519–2529.
30. Hung R.J., Hall J., Brennan P. et al. Genetic polymorphisms in the base excision repair pathway and cancer risk: a HuGE review // *Am. J. Epidemiol.* 2005. V. 162. P. 925–942.
31. Simonelli V., Mazzei F., D'Errico M. et al. Gene susceptibility to oxidative damage: From single nucleotide polymorphisms to function // *Mutat. Res.* 2012. V. 731. P. 1–13.
32. Hadi M.Z., Coleman M.A., Fidelis K. et al. Functional characterization of ApeI variants identified in the human population // *Nucl. Acids Res.* 2000. V. 28. № 20. P. 3871–3879.
33. Hu J.J., Smith T.R., Miller M.S. et al. Amino acid substitution variants of APE1 and XRCC1 genes associated with ionizing radiation sensitivity // *Carcinogenesis*. 2001. V. 22. P. 917–922.
34. De Ruyck K., Szaumkessel M., De Rudder I. et al. Polymorphisms in base-excision repair and nucleotide-excision repair genes in relation to lung cancer risk // *Mutat. Res.* 2007. V. 631. № 2. P. 101–110.  
<https://doi.org/10.1016/j.mrgentox.2007.03.010>
35. Kiyohara C., Takayama K., Nakanishi Y. Association of genetic polymorphisms in the base excision repair pathway with lung cancer risk: a meta-analysis // *Lung Cancer*. 2006. V. 54. P. 267–283.
36. Popanda O., Schattenberg T., Phong C.T. et al. Specific combinations of DNA repair gene variants and increased risk for non-small cell lung cancer // *Carcinogenesis*. 2004. V. 25. P. 2433–2441.
37. Shen M., Berndt S.I., Rothman N. et al. Polymorphisms in the DNA base excision repair genes APEX1 and XRCC1 and lung cancer risk in Xuan Wei, China // *Anticancer Res.* 2005. V. 25. P. 537–542.
38. Manini P., De Palma G., Andreoli R. et al. Biomarker of nucleic acid oxidation, polymorphism in, and expres-

- sion of, hOGG1 gene in styrene-exposed worke // *Toxicol. Lett.* 2009. V. 190. № 1. P. 41–47.
39. *Rihs H.P., Marczynski B., Lotz A. et al.* Modulation of oxidative DNA damage by repair enzymes XRCC1 and hOGG1 // *J. Toxicol. Environ. Health A.* 2012. V. 75. P. 588–596.
40. *Kohno T., Shinmura K., Tosaka M. et al.* Genetic polymorphisms and alternative splicing of the hOGG1 gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA // *Oncogene.* 1998. V. 16. P. 3219–3225.
41. *Guan P., Huang D., Yin Z. et al.* Association of the hOGG1 Ser326Cys polymorphism with increased lung cancer susceptibility in Asians: A meta-analysis of 18 studies including 7592 cases and 8129 controls // *Asian Pacific J. Cancer Pre.* 2012. V. 12. P. 1067–1072.
42. *Zienolddiny S., Campa D., Lind H. et al.* Polymorphisms of DNA repair genes and risk of non-small cell lung cancer // *Carcinogenesis.* 2006. V. 27. P. 560–567.
43. *Liu C.J., Hsia T.C., Tsai R.Y. et al.* The joint effect of hOGG1 single nucleotide polymorphism and smoking habit on lung cancer in Taiwan // *Anticancer Res.* 2010. V. 30. P. 4141–4146.
44. *Wikman H., Risch A., Klimek F.* hOGG1 polymorphism and loss of heterozygosity (LOH): Significance for lung cancer susceptibility in a Caucasian population // *Int J. Cancer.* 2000. V. 88. P. 932–937.
45. *Park J., Chen L., Tockman M.S. et al.* The human 8-oxoguanine DNA N-glycosylase 1 (hOGG1) DNA repair enzyme and its association with lung cancer risk // *Pharmacogenetics.* 2004. V. 14. P. 103–109.
46. *Sugasawa K.* XPC: its product and biological roles // *Adv. Exp. Med. Biol.* 2008. V. 637. P. 47–56.
47. *Volker M., Moné M.J., Karmakar P. et al.* Sequential assembly of the nucleotide excision repair factors *in vivo* // *Mol. Cell.* 2001. V. 8. P. 213–224.
48. *Qiao Y., Spitz M.R., Shen H. et al.* Modulation of repair of ultraviolet damage in the host-cell reactivation assay by polymorphic XPC and XPD/ERCC2 genotypes // *Carcinogenesis.* 2002. V. 23. P. 295–299.
49. *Vodicka P., Kumar R., Stetina R. et al.* Genetic polymorphisms in DNA repair genes and possible links with DNA repair rates, chromosomal aberrations and single-strand breaks in DNA // *Carcinogenesis.* 2004. V. 25. P. 757–763.
50. *Hu Z., Wang Y., Wang X. et al.* DNA repair gene XPC genotypes/haplotypes and risk of lung cancer in a Chinese population // *Int. J. Cancer.* 2005. V. 115. P. 478–483.
51. *Letkova L., Matakova T., Musak L. et al.* DNA repair genes polymorphism and lung cancer risk with the emphasis to sex differences // *Mol. Biol. Rep.* 2013. V. 40. № 9. P. 5261–5273.
52. *Broustas C., Lieberman H.* DNA damage response genes and the development of cancer metastasis // *Radiat. Res.* 2014. V. 181. № 2. P. 111–130.
53. *Jiang J., Liang X., Zhou X. et al.* DNA repair gene X-ray repair cross complementing group 1 Arg194Trp polymorphism on the risk of lung cancer: A meta-analysis on 22 studies // *J. Thorac. Oncol.* 2010. V. 5. P. 1741–1747.
54. *Hanssen-Bauer A., Solvang-Garten K., Akbari M. et al.* X-ray repair cross complementing protein 1 in base excision repair // *Int. J. Mol. Sci.* 2012. V. 13. P. 17210–17229.  
<https://doi.org/10.3390/ijms131217210>
55. *Wood R.D., Mitchel M., Sgouros J. et al.* Human DNA repair genes // *Science.* 2001. V. 291. P. 1284–1289.
56. *Ratnasinghe D., Yao S.X., Tangrea J.A. et al.* Polymorphisms of the DNA repair gene XRCC1 and lung cancer risk // *Cancer Epidemiol. Biomarkers Prev.* 2001. V. 10. P. 119–123.
57. *Misra R.R., Ratnasinghe D., Tangrea J.A. et al.* Polymorphisms in the DNA repair genes XPD, XRCC1, XRCC3, and APE/ref-1, and the risk of lung cancer among male smokers in Finland // *Cancer Lett.* 2003. V. 191. P. 171–178.
58. *Hung R.J., Brennan P., Canzian F. et al.* Large-scale investigation of base excision repair genetic polymorphisms and lung cancer risk in a multicenter study // *J. Natl Cancer Inst.* 2005. V. 97. P. 567–576.
59. *Asamura H., Goya T., Koshiishi Y. et al.* A Japanese lung cancer registry study: Prognosis of 13,010 resected lung cancers // *J. Thorac. Oncol.* 2008. V. 3. № 1. P. 46–52.
60. *Chang J., Asamura H., Kawachi R. et al.* Gender difference in survival of resected non-small cell lung cancer: histology-related phenomenon? // *J. Thorac. Cardiovasc. Surg.* 2009. V. 137. P. 807–812.
61. *Wang B., Huang J., Cheng C. et al.* Lung cancer and prognosis in Taiwan a population-based cancer registry // *J. Thorac. Oncol.* 2013. V. 8. № 9. P. 1128–1135.
62. *Broughton B.C., Steingrimsdottir H., Lehmann A.R.* Five polymorphisms in the coding sequence of the Xeroderma pigmentosum group D gene // *Mutat. Res.* 1996. V. 362. P. 209–211.
63. *Hou S.M., Failt S., Andelini S. et al.* The XPD variant alleles are associated with increased aromatic DNA adduct level and lung cancer risk // *Carcinogenesis.* 2002. V. 23. P. 599–603.
64. *Au W.W., Salama S.A., Sierra-Torres C.H.* Functional characterization of polymorphism in DNA repair genes using cytogenetic challenge assays // *Environ. Health Perspect.* 2003. V. 111. P. 1842–1850.
65. *Ito S., Kuraoka I., Chymkowitch P. et al.* XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: Implications for Cockayne syndrome in XP-G/CS patients // *Mol. Cell.* 2007. V. 26. P. 231–243.

## Studying the Role of DNA Repair Genes Polymorphism in Formation of Preposition to Lung Cancer Development in Women

R. A. Titov<sup>a, \*</sup>, V. I. Minina<sup>a, b, \*\*</sup>, A. V. Torgunakova<sup>a</sup>, V. Yu. Buslaev<sup>a</sup>,  
E. N. Voronina<sup>c</sup>, A. Yu. Prosekov<sup>b</sup>, V. A. Titov<sup>d</sup>, and A. N. Glushkov<sup>a</sup>

<sup>a</sup>Federal Research Center of Coal and Coal Chemistry, Siberian Branch,  
Russian Academy of Sciences, Kemerovo, 650065 Russia

<sup>b</sup>Kemerovo State University, Kemerovo, 650043 Russia

<sup>c</sup>Institute of Chemical Biology and Fundamental Medicine, Siberian Branch,  
Russian Academy of Sciences, Novosibirsk, 630090 Russia

<sup>d</sup>Kuzbass Clinical Oncologic Dispensary, Kemerovo, 650036 Russia

\*e-mail: ruslan-tito00@rambler.ru

\*\*e-mail: vminina@mail.ru

We examined 522 women of Russian nationality residing in Kemerovo oblast, Russia, including 273 patients with lung cancer and 249 women of similar age with no signs of cancer. The aim of this study was a comparative analysis of polymorphic variants of DNA repair genes: *APEX1 444T>G* (*rs1130409*), *XRCC1 1839G>A* (*rs25489*), *hOGG1 977C>G* (*rs1052133*), *XPD 2251T>G* (*rs13181*), *XPG 3310G>C* (*rs17655*), *XPC 2815A>C* (*rs2228001*) in lung cancer patients and individuals without cancer living in the same area. Single-locus effects analysis showed significant associations between lung cancer risk and variants of the *XPC 2815A>C* gene (*rs2228001*) (OR = 0.56, CI: 0.39–0.81,  $p = 0.0018$ ) in the overall group, *APEX1 444T>G* (*rs1130409*) gene (OR = 0.15, CI: 0.03–0.67,  $p = 0.0027$ ) in the smoking group, *XPC 2815A>C* (*rs2228001*) gene (OR = 0.36, CI: 0.18–0.69,  $p = 0.0051$ ) and *hOGG1 977C>G* (*rs1052133*) (OR = 0.57, CI: 0.38–0.85,  $p = 0.0055$ ) in the nonsmoker group. MDR analysis of gene-gene interactions showed that *XPD 2251T>G* and *XPC 2815A>C*; *APEX1 444T>G* and *XPD 2251T>G* genes closely interact and mutually increase lung cancer risk in West Siberian women.

**Keywords:** lung cancer, polymorphism of the *APEX1*, *XRCC1*, *hOGG1*, *XPD*, *XPG*, *XPC* genes.

## АНАЛИЗ АССОЦИАЦИЙ ПОЛИМОРФНЫХ ЛОКУСОВ ГЕНА *LOXL1* С РАЗВИТИЕМ ПЕРВИЧНОЙ ОТКРЫТОУГОЛЬНОЙ ГЛАУКОМЫ У ЖЕНЩИН ЦЕНТРАЛЬНО-ЧЕРНОЗЕМНОГО РЕГИОНА РОССИИ

© 2022 г. Н. В. Елисеева<sup>1</sup>, И. В. Пономаренко<sup>1</sup>, М. И. Чурнос<sup>1</sup>, \*

<sup>1</sup>Белгородский государственный национальный исследовательский университет, Белгород, 308015 Россия

\*e-mail: churnosov@bsu.edu.ru

Поступила в редакцию 08.06.2021 г.

После доработки 23.07.2021 г.

Принята к публикации 10.08.2021 г.

Проведено репликативное исследование ассоциаций полиморфизма гена *LOXL1* с первичной открытоугольной глаукомой (ПОУГ) у женского населения Центрально-Черноземного региона России. Работа выполнена в дизайне “больные–контроль”. Выборка для исследования составила 290 женщин больных ПОУГ и 220 женщин контрольной группы. Проведено генотипирование трех полиморфных локусов гена *LOXL1* (rs2165241, rs4886776, rs893818) методом ПЦР (технология Tag-Map зондов). Выявлено, что SNPs гена *LOXL1* (rs2165241, rs4886776, rs893818) ассоциированы с ПОУГ у женщин Центрального Черноземья России: аллель С rs2165241 (OR = 0.33–0.45 при  $p_{\text{perm}} \leq 0.0005$ ), аллели А rs4886776 (OR = 0.62–0.63 при  $p_{\text{perm}} \leq 0.031$ ) и rs893818 (OR = 0.53–0.62 при  $p_{\text{perm}} \leq 0.007$ ) и гаплотип САА rs2165241–rs4886776–rs893818 (OR = 0.56 при  $p_{\text{perm}} = 0.022$ ) имеют протективное значение для развития заболевания, а гаплотип TGG (OR = 2.19 при  $p_{\text{perm}} = 0.001$ ) ассоциирован с повышенным риском развития ПОУГ у женщин.

**Ключевые слова:** *LOXL1*, ассоциации, первичная открытоугольная глаукома, полиморфизм, женщины.

**DOI:** 10.31857/S0016675822020047

Глаукома – это группа гетерогенных нейро-дегенеративных заболеваний с общими патогенетическими путями, в основе которых лежит прогрессирующая потеря ганглиозных клеток сетчатки и аксонов зрительного нерва, приводящая к дефектам поля зрения [1]. Эпидемиологические исследования показывают, что во всем мире 64.3 млн человек в возрасте 40–80 лет страдают глаукомой и по прогнозам ученых распространенность заболевания увеличится до 111.8 млн в 2040 г. [1]. Первичная открытоугольная глаукома (ПОУГ) – это наиболее распространенная форма глаукомы, которая является одной из основных причин слепоты во всем мире [2, 3]. Следует отметить, что наиболее часто ПОУГ встречается среди женщин и несколько реже среди мужчин [2, 3].

Согласно литературным данным генетические факторы играют важную роль в развитии глаукомы [4, 5]. Результаты полногеномных ассоциативных исследований (GWAS) выявили ряд генов-кандидатов, ассоциированных с глаукомой, включая ген лизилоксидазоподобного фермента (*LOXL1*) [6–8]. Продукт гена *LOXL1* модулирует биогенез внеклеточного матрикса путем сшивания эластина и коллагена в соединительных тканях [9]. Эластин является основным компонен-

том эластических волокон внеклеточного матрикса решетчатой пластинки, и деформация решетчатой пластинки может приводить к повреждению аксонов ганглиозных клеток сетчатки [9]. Показана связь ряда полиморфных локусов гена *LOXL1* с уровнем его экспрессии [8, 10].

Важно отметить, что до недавнего времени считалось, что полиморфизм гена *LOXL1* ассоциирован только с псевдоэксфолиативным синдромом/псевдоэксфолиативной глаукомой (ПЭС/ПЭГ) и не связан с развитием других видов глаукомы (ПОУГ, первичная закрыто-угольная глаукома, пигментная глаукома) [11]. Однако в последние годы появились данные об ассоциации полиморфизма гена *LOXL1* с ПОУГ [12], в том числе в полногеномных исследованиях [7].

Цель настоящей работы – репликативное исследование ассоциаций полиморфизма гена *LOXL1* с ПОУГ у женского населения Центрально-Черноземного региона России.

### МАТЕРИАЛЫ И МЕТОДЫ

Выборка для исследования представлена 510 женщинами, из которых 290 – больные ПОУГ и 220 – контрольная группа. Специализированное

обследование проводилось на клинической базе отделения офтальмохирургии Белгородской областной клинической больницы. Все больные ПОУГ и индивидуумы контрольной группы дали письменное информированное согласие на участие в исследовании.

В группу больных включались индивидуумы с диагнозом ПОУГ, который был верифицирован в результате клинического и клинико-инструментального обследования пациентов. Диагноз ПОУГ устанавливался на основе соответствующих критериев – высокое внутриглазное давление (ВГД выше 21 при пневмотонометрии и выше 25 при тонометрии по Маклакову), глаукоматозная экскавация диска зрительного нерва, характерные изменения периферического поля зрения [13]. Среди пациенток с ПОУГ сердечно-сосудистые заболевания встречались у 72.76%, болезни эндокринной системы – 15.86%, нервной – 14.83%, пищеварительной – 12.07%, половой – 11.03%, мочевыделительной – 7.24%. В группу контроля включены индивидуумы, не имеющие ПОУГ (ВГД ниже 21 при пневмотонометрии и ниже 25 при тонометрии по Маклакову, отсутствие глаукоматозной экскавации диска зрительного нерва и характерных изменений периферического поля зрения), других заболеваний глаз и тяжелой сопутствующей соматической патологии, сопровождающейся поражением глаз [14]. В исследуемые выборки больных и контроля включались неродственные индивидуумы русской национальности, родившиеся и проживающие в Центрально-Черноземном регионе России [15, 16]. Возраст больных ПОУГ и контрольной группы достоверно не отличался ( $62.24 \pm 11.45$  и  $61.78 \pm 11.06$  лет соответственно при  $p > 0.05$ ).

В качестве объекта для экспериментального исследования использовалась геномная ДНК (выделена из образцов венозной крови фенольно/хлороформным методом экстракции) [17]. Для молекулярно-генетического исследования был проведен отбор полиморфных локусов гена *LOXL1* на основе следующих критериев [18, 19]: 1) ассоциации с глаукомой согласно результатам проведенных ранее полногеномных исследований; 2) наличие регуляторного потенциала и связи с экспрессией генов [20]; 3) частота минорного аллеля 5% и более.

При отборе SNPs использовались каталог полногеномных исследований (GWAS) (<http://www.genome.gov/gwastudies/>) и база данных HaploReg (<http://archive.broadinstitute.org/mammals/haploreg/>). Согласно вышеуказанным критериям для настоящего исследования были отобраны три полиморфных локуса гена *LOXL1* (rs2165241, rs4886776, rs893818). Все три полиморфизма ассоциированы с глаукомой (экссфолиативная глаукома/синдром) по данным GWAS [6, 8, 21, 22], имели значимый регуляторный потенциал и были связаны с экс-

прессией генов, частота минорных аллелей превышала 5%.

Для генотипирования SNPs применяли метод полимеразной цепной реакции (технология Tag-Map зондов) и локус-специфические наборы, разработанные и синтезированные ООО “Тест-Ген” (г. Ульяновск). Экспериментальные исследования выполняли на приборе-амплификаторе CFX96 (фирма-производитель Bio-Rad) согласно протоколу фирмы-разработчика.

Для анализа ассоциаций полиморфных локусов и их гаплотипов с ПОУГ использовали метод логистической регрессии [23, 24], имплементированный в программе plink 1.06 (представлена в свободном доступе на электронном ресурсе <http://pngu.mgh.harvard.edu/Épurcell/plink>). Расчеты выполняли в рамках доминантной, аддитивной и рецессивной генетических моделей [25]. Неравновесие по сцеплению SNPs гена *LOXL1* оценивалось на основе коэффициентов Левонтина ( $D'$ ) и Пирсона ( $r^2$ ). Для оценки характера ассоциации использовали показатель отношения шансов (OR) и его 95%-ный интервал (95%CI) [26]. При расчетах проводили коррекцию на множественные сравнения (выполнялась адаптивная пермутационная процедура) и ковариаты (возраст). За статистически значимый принимали показатель  $p_{\text{perm}} < 0.05$  [27].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

По всем изучаемым трем полиморфным локусам гена *LOXL1* (rs2165241, rs4886776, rs893818) как среди больных, так и в контроле в распределении генотипов (наблюдаемом и ожидаемом) было выполнено равновесие Харди–Вайнберга (с учетом поправки Бонферрони на количество анализируемых локусов,  $n = 3$ ,  $p_{\text{bonf}} > 0.017$ ) (табл. 1). Встречаемость минорных (редких) аллельных вариантов была выше 17%.

Установлены ассоциации изученных полиморфных локусов гена *LOXL1* (rs2165241, rs4886776 и rs893818) с развитием ПОУГ у женщин (табл. 2). Полиморфизм rs2165241 ассоциирован с ПОУГ в рамках всех трех анализируемых генетических моделей: аддитивной – OR = 0.45 ( $p = 7.55 \times 10^{-7}$ ,  $p_{\text{perm}} = 1.00 \times 10^{-6}$ ), доминантной – OR = 0.37 ( $p = 6.46 \times 10^{-6}$ ,  $p_{\text{perm}} = 7.00 \times 10^{-6}$ ) и рецессивной – OR = 0.33 ( $p = 0.0004$ ,  $p_{\text{perm}} = 0.0005$ ). Полиморфные локусы rs4886776 и rs893818 связаны с развитием заболевания у женщин согласно аддитивной (OR = 0.62,  $p = 0.012$ ,  $p_{\text{perm}} = 0.015$  и OR = 0.62,  $p = 0.008$ ,  $p_{\text{perm}} = 0.007$  соответственно) и доминантной (OR = 0.63,  $p = 0.032$ ,  $p_{\text{perm}} = 0.031$  и OR = 0.53,  $p = 0.004$ ,  $p_{\text{perm}} = 0.004$  соответственно) генетических моделей. При этом следует отметить, что все альтернативные варианты данных полиморфных

**Таблица 1.** Характеристика распределения полиморфных локусов гена *LOXLI* среди пациенток с ПОУГ и женщин контрольной группы

Полиморфизм	Редкий аллель	Частый аллель	Частота редкого аллеля	Кол-во изученных хромосом	Данные о распределении генотипов, число (%)*	Наблюдаемая гетерозиготность ( $H_o$ )	Ожидаемая гетерозиготность ( $H_e$ )	$P_{HWE}$
Больные ПОУГ ( $n = 290$ )								
rs2165241	C	T	0.276	576	24/111/153 (8.33/38.54/53.13)	0.385	0.400	0.556
rs4886776	A	G	0.189	572	4/100/182 (1.39/34.97/63.64)	0.350	0.306	0.019
rs893818	A	G	0.176	568	9/82/193 (3.17/28.87/67.96)	0.289	0.290	1.000
Контрольная группа ( $n = 220$ )								
rs2165241	C	T	0.461	438	45/112/62 (20.55/51.14/28.31)	0.511	0.497	0.786
rs4886776	A	G	0.271	432	15/87/114 (6.94/40.28/52.78)	0.403	0.395	0.864
rs893818	A	G	0.259	432	14/84/118 (6.48/38.89/54.63)	0.389	0.384	1.000

\* Приведены данные в формате гомозиготы по редкому аллелю/гетерозиготы/гомозиготы по частому аллелю.

локусов (аллель C rs2165241, аллели A rs4886776 и rs893818) имеют протективное значение для формирования заболевания ( $OR < 1$ ).

Выявлено, что три изученных нами полиморфизма гена *LOXLI* (rs2165241, rs4886776, rs893818) расположены рядом (физическое расстояние между ними равно 6 тыс. пар нуклеотидов), находятся в состоянии неравновесия по сцеплению ( $r^2 = 0.31-0.72$ ,  $D' = 0.78-0.87$ ) и образуют единый гапоблок. Определены ассоциации с формированием ПОУГ у женщин трех гаплотипов rs2165241–rs4886776–rs893818 гена *LOXLI* (табл. 3): TGG ( $OR = 2.19$ ,  $p = 2 \times 10^{-6}$ ,  $p_{perm} = 0.001$ ), CGG ( $OR = 0.44$ ,  $p = 0.0001$ ,  $p_{perm} = 0.002$ ), САА ( $OR = 0.56$ ,  $p = 0.006$ ,  $p_{perm} = 0.022$ ).

Полученные нами результаты согласуются с литературными данными по этому вопросу. В первом полногеномном исследовании глаукомы, выполненном в 2007 г. G. Thorleifsson et al. [8] в популяциях Исландии и Швеции, в исследование были включены 274 больных с ПЭГ, 290 пациентов с ПОУГ (таким образом, выборка больных с глаукомой в целом составила 564 человека) и 14672 человека контрольной группы. Авторами были установлены ассоциации rs2165241 гена *LOXLI* с развитием как ПЭГ ( $OR = 3.62$ ,  $p = 1.00 \times 10^{-27}$ ), так и глаукомы в целом (объединенная выборка ПЭГ и ПОУГ) ( $OR = 1.96$ ,  $p = 1.30 \times 10^{-16}$ ).

Значимая роль полиморфизма rs2165241 гена *LOXLI* в формировании псевдоэкзофалиативного синдрома без глаукомы была показана и в полногеномном исследовании K. Zagajewska et al. [21] в польской популяции (работа выполнена на выборке из 209 индивидуумов, в том числе больных – 103 и контроля – 106). Авторы установили, что минорный аллель C является протективным фактором развития ПЭС ( $OR = 0.24$ ), тогда как референсный для него аллель T существенно повышает риск развития заболевания ( $OR = 4.2$ ,  $p = 2.77 \times 10^{-10}$ ).

Следует отметить, что наши данные полностью соответствуют результатам ранее выполненного исследования V. Zanon-Moreno et al. [12] в испанской популяции. В данной работе (проведено генетическое исследование 232 больных ПОУГ и 241 контроля) установлено, что полиморфизм rs2165241 гена *LOXLI*, ранее показавший, как отмечают авторы, ассоциации с псевдоэкзофалиативной глаукомой, также ассоциирован и с развитием ПОУГ в популяции средиземноморья, причем как и для ПЭГ повышает риск развития ПОУГ аллель T rs2165241 гена *LOXLI* (согласно рецессивной генетической модели для TT в сравнении с CC  $OR = 2.19$  95%CI 1.33–3.62, при коррекции на возраст и вес  $OR = 2.07$  95%CI 1.20–3.57).

Значимая роль аллеля T rs2165241 как фактора риска ПЭС/ПЭГ показана в исследованиях латино-американского населения [28], немецкой и итальянской популяций [29], жителей Испании

**Таблица 2.** Ассоциации полиморфизма гена *LOXLI* с ПОУГ у женщин

SNP	Аллели, генотипы	Больные, <i>n</i> (%)	Контроль, <i>n</i> (%)	OR (95%CI)	<i>p</i>
rs2165241	Объем выборки	288	219		
	<i>C</i> vs <i>T</i> (аллельная модель)	159/417 (27.60/72.40)	202/236 (46.12/53.88)	0.45 (0.34–0.58)	1.06E-09
	<i>C/C</i> vs <i>T/C</i> vs <i>T/T</i> (аддитивная модель)	24/111/153 (8.33/38.54/53.13)	45/112/62 (20.55/51.14/28.31)	0.45 (0.33–0.62)	7.55E-7
	<i>C/C</i> + <i>T/C</i> vs <i>T/T</i> (доминантная модель)	135/153 (46.87/53.13)	157/62 (71.69/28.31)	0.37 (0.24–0.57)	6.46E-6
	<i>C/C</i> vs <i>T/C</i> + <i>T/T</i> (рецессивная модель)	24/264 (8.33/91.67)	45/174 (20.55/79.45)	0.33 (0.18–0.61)	0.004
rs4886776	Объем выборки	286	216		
	<i>A</i> vs <i>G</i> (аллельная модель)	108/464 (18.88/81.12)	117/315 (27.08/72.92)	0.63 (0.47–0.84)	0.02
	<i>A/A</i> vs <i>G/A</i> vs <i>G/G</i> (аддитивная модель)	4/100/182 (1.39/34.97/63.64)	15/87/114 (6.94/40.28/52.78)	0.62 (0.43–0.90)	0.012
	<i>A/A</i> + <i>G/A</i> vs <i>G/G</i> (доминантная модель)	104/182 (36.36/63.64)	102/114 (47.22/52.78)	0.63 (0.41–0.96)	0.032
	<i>A/A</i> vs <i>G/A</i> + <i>G/G</i> (рецессивная модель)	4/282 (1.39/98.61)	15/201 (6.94/93.06)	0.30 (0.09–1.03)	0.056
rs893818	Объем выборки	284	216		
	<i>A</i> vs <i>G</i> (аллельная модель)	100/468 (17.61/82.39)	112/320 (25.93/74.07)	0.61 (0.45–0.83)	0.001
	<i>A/A</i> vs <i>G/A</i> vs <i>G/G</i> (аддитивная модель)	9/82/193 (3.17/28.87/67.96)	14/84/118 (6.48/38.89/54.63)	0.62 (0.43–0.88)	0.008
	<i>A/A</i> + <i>G/A</i> vs <i>G/G</i> (доминантная модель)	91/193 (32.04/67.96)	98/118 (45.37/54.63)	0.53 (0.35–0.82)	0.004
	<i>A/A</i> vs <i>G/A</i> + <i>G/G</i> (рецессивная модель)	9/275 (3.17/96.83)	14/202 (6.48/93.52)	0.69 (0.26–1.87)	0.469

Примечание. Результаты получены методом логистической регрессии; OR – отношение шансов, 95%CI – 95%-ный доверительный интервал; *p* – уровень статистической значимости.

**Таблица 3.** Ассоциации гаплотипов полиморфных локусов *LOXLI* с ПОУГ у женщин

Гаплотип	Частота гаплотипа		OR	<i>p</i>
	больные ( <i>n</i> = 290)	контроль ( <i>n</i> = 220)		
CAA	0.143	0.206	0.56	0.006
TAA	0.012	0.028	0.48	0.24
CGA	0.011	0.013	0.83	0.809
TGA	0.010	0.009	2.74	0.197
CAG	0.015	0.017	1.51	0.533
TAG	0.019	0.020	1.44	0.536
CGG	0.110	0.227	0.44	0.0001
TGG	0.680	0.480	2.19	2.00E-06

Примечание. Результаты получены методом логистической регрессии; OR – отношение шансов; *p* – уровень статистической значимости.

[30] и других популяций [31–33]. Результаты метаанализа, представленные в работе J.Z. Tang et al. [34], свидетельствуют о протективном значении генотипа *CC* rs2165241 при формировании псевдоэксфолиативного синдрома/глаукомы.

Следует отметить, что полиморфные варианты локуса rs2165241 гена *LOXL1* имеют разнонаправленный характер ассоциаций с развитием заболевания в популяциях разного этнического состава. Так, если в европейских популяциях, как установлено в работах, приведенных нами выше [8, 21, 29], фактором риска развития ПЭС/ПЭГ является аллель *T* rs2165241, то в азиатских популяциях фактором риска формирования заболевания служит аллель *C* rs2165241 [29, 35]. Также следует отметить, что в европейских популяциях, как правило, частым является аллель *T* rs2165241, а в азиатских – аллель *C* rs2165241 [12, 21, 35].

Связь рассмотренного в нашей работе SNP rs893818 гена *LOXL1* с ПЭГ при полногеномном уровне статистической значимости впервые была обнаружена в 2014 г. научным коллективом M. Nakato et al. [6] при изучении населения Японии. Ассоциации с ПЭС rs4886776 гена *LOXL1* впервые была продемонстрирована в GWAS исследовании T. Aung et al. [22]. При этом следует отметить, что в данной работе были обнаружены разнонаправленные ассоциации rs4886776 *LOXL1* с ПЭС в популяциях с разным этническим составом: аллельный вариант *A* rs4886776 служит фактором риска для ПЭС в популяции Японии ( $OR = 9.87, p = 2.35 \times 10^{-217}$ ) и что явилось, как указывают авторы в своей работе, “сюрпризом” для них, служит протективным фактором для ПЭС в неазиатских популяциях (в том числе и европеоидных) ( $OR = 0.49, p = 2.35 \times 10^{-317}$ ). Эти данные полностью согласуются с полученными в нашей работе результатами – в изученной нами популяции России (европеоидная популяция) протективным фактором развития ПОУГ является аллель *A* rs4886776 гена *LOXL1* ( $OR = 0.63$ ). Разнонаправленный характер ассоциации другого полиморфизма гена *LOXL1* – rs4886778 с развитием псевдоэксфолиативного синдрома выявлен и в работе F. Rasutto et al. [29]: в итальянской и немецкой популяциях фактором риска развития заболевания является аллель *C* rs4886778 ( $OR = 3.95, p = 1.54 \times 10^{-32}$  и  $OR = 3.10, p = 8.30 \times 10^{-22}$  соответственно), тогда как в японской популяции рисковое значение имеет аллель *A* ( $OR = 6.85, p = 1.53 \times 10^{-136}$ ).

Следует отметить, что изучаемые нами однонуклеотидные полиморфные локусы расположены в первом интроне гена *LOXL1*. Данные полиморфные варианты, согласно ранее опубликованным данным, модифицируют активность промоторного участка гена *LOXL1-ASI* [36], влияют на связывание этого региона ДНК с фактором транскрипции RXRa и вовлечены в процессы альтернатив-

ного сплайсинга гена *LOXL1* [29]. Ген *LOXL1-ASI* (*LOXL1 antisense RNA 1*) – это длинная некодирующая РНК, обладающая существенными регуляторными эффектами [37]. Литературные источники указывают на вовлеченность *LOXL1-ASI* в патофизиологию псевдоэксфолиативного синдрома за счет участия в клеточном стрессовом ответе, при котором наблюдается значительная дисрегуляция экспрессии этой некодирующей РНК [36]. Показана значимая роль lncRNA *LOXL1-ASI* в модуляции экспрессии генов, отвечающих за реакцию на оксидативный стресс, деградацию экстрацеллюлярного матрикса, вовлеченных в формирование коллагеновых структур (*HMOX1, TIMP3, LOXL4, ACTA2, COL6A3* и др.), и других генов, имеющих “ключевое” значение для патофизиологии глаукомы [38].

Полученные данные свидетельствуют о вовлеченности в формирование ПОУГ у женщин Центрально-Черноземного региона России полиморфных локусов гена *LOXL1* (rs2165241, rs4886776, rs893818). Аллельные варианты *C* rs2165241, *A* rs4886776, *A* rs893818, а также гаплотип САА rs2165241–rs4886776–rs893818 проявляют “защитную” роль” при формировании ПОУГ у женского населения Центрально-Черноземного региона России, тогда как гаплотип ТGG, наоборот, связан с более высоким риском возникновения ПОУГ.

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики.

От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Tham Y.C., Li X., Wong T.Y. et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis // *Ophthalmology*. 2014. V. 121. P. 2081–2090.
2. Grzybowski A., Och M., Kanclerz P. et al. Primary open angle glaucoma and vascular risk factors: A review of population based studies from 1990 to 2019 // *J. Clin. Med*. 2020. V. 9(3). P. 761.
3. Kreft D., Doblhammer G., Guthoff R.F., Frech S. Prevalence, incidence and risk factors of primary open angle glaucoma – a cohort study based on longitudinal data from a German public health in surname // *BMC Public Health*. 2019. V. 19. P. 851.
4. Liu Y., Allingham R.R. Molecular genetics in glaucoma // *Exp. Eye Res*. 2011. V. 93. P. 331–339.

5. *Eliseeva N., Ponomarenko I., Reshetnikov E. et al.* The haplotype of the *CDKN2B-AS1* gene is associated with primary open-angle glaucoma and pseudoexfoliation glaucoma in the Caucasian population of Central Russia // *Ophthalmic Genet.* 2021. V. 13. P. 1–8.
6. *Nakano M., Ikeda Y., Tokuda Y. et al.* Novel common variants and susceptible haplotype for exfoliation glaucoma specific to an Asian population // *Sci. Rep.* 2014. V. 4. P. 5340.
7. *Shiga Y., Akiyama M., Nishiguchi K.M. et al.* Genome-wide association study identifies seven novel susceptibility loci for primary open-angle glaucoma // *Hum. Mol. Genet.* 2018. V. 27. P. 1486–1496.
8. *Thorleifsson G., Magnusson K.P., Sulem P. et al.* Common sequence variants in the *LOXL1* gene confer susceptibility to exfoliation glaucoma // *Science.* 2007. V. 317. P. 1397–1400.
9. *Liu X., Zhao Y., Gao J. et al.* Elastic fiber homeostasis requires lysyl oxidase-like 1 protein // *Nat. Genet.* 2004. V. 36. P. 178–182.
10. *Schlötzer-Schrehardt U., Zenkel M.* The role of lysyl oxidase-like 1 (*LOXL1*) in exfoliation syndrome and glaucoma // *Exp. Eye Res.* 2019. V. 189. P. 107818.
11. *Aboobakar I.F., Allingham R.R.* Genetics of exfoliation syndrome and glaucoma // *Int. Ophthalmol. Clin.* 2014. V. 54(4). P. 43–56.  
<https://doi.org/10.1097/HIO.0000000000000042>
12. *Zanon-Moreno V., Zanon-Moreno L., Ortega-Azorin C. et al.* Genetic polymorphism related to exfoliative glaucoma is also associated with primary open-angle glaucoma risk // *Clin. Exp. Ophthalmol.* 2015. V. 43(1). P. 26–30.  
<https://doi.org/10.1111/ceo.12367>
13. *Tikunova E., Ovtcharova V., Reshetnikov E. et al.* Genes of tumor necrosis factors and their receptors and the primary open angle glaucoma in the population of Central Russia // *Int. J. Ophthalmol.* 2017. V. 10. P. 1490–1494.  
<https://doi.org/10.18240/ijjo.2017.10.02>
14. *Starikova D., Ponomarenko I., Reshetnikov E. et al.* Novel data about association of the functionally significant polymorphisms of the *MMP-9* gene with exfoliation glaucoma in the Caucasian population of Central Russia // *Ophthalmic Res.* 2021. V. 64(3). P. 458–464.  
<https://doi.org/10.1159/000512507>
15. *Reshetnikov E., Zarudskaya O., Polonikov A. et al.* Genetic markers for inherited thrombophilia are associated with fetal growth retardation in the population of Central Russia // *J. Obstet. Gynaecol. Res.* 2017. V. 43(7). P. 1139–1144.  
<https://doi.org/10.1111/jog.13329>
16. *Ponomarenko I., Reshetnikov E., Polonikov A. et al.* Candidate genes for age at menarche are associated with endometrial hyperplasia // *Gene.* 2020. V. 757. P. 144933.  
<https://doi.org/10.1016/j.gene.2020.144933>
17. *Reshetnikov E.A., Akulova L.Y., Dobrodomova I.S. et al.* The insertion-deletion polymorphism of the *ACE* gene is associated with increased blood pressure in women at the end of pregnancy // *J. Renin Angiotensin Aldosterone Syst.* 2015. V. 16(3). P. 623–632.  
<https://doi.org/10.1177/1470320313501217>
18. *Ponomarenko I., Reshetnikov E., Polonikov A. et al.* Candidate genes for age at menarche are associated with uterine leiomyoma // *Front. Genet.* 2021. V. 11. P. 512940.  
<https://doi.org/10.3389/fgene.2020.512940>
19. *Москаленко М.И., Миланова С.Н., Пономаренко И.В. и др.* Исследование ассоциаций полиморфизма генов матриксных металлопротеиназ с развитием артериальной гипертензии у мужчин // *Кардиология.* 2019. Т. 59. № 7S. С. 31–39.  
<https://doi.org/10.18087/cardio.2598>
20. *Полоников А.В., Клёсова Е.Ю., Азарова Ю.Э.* Биоинформатические инструменты и интернет-ресурсы для оценки регуляторного потенциала полиморфных локусов, установленных полногеномными ассоциативными исследованиями мультифакториальных заболеваний (обзор) // *Научные результаты биомедицинских исследований.* 2021. Т. 7. № 1. С. 15–31.  
<https://doi.org/10.18413/2658-6533-2020-7-1-0-2>
21. *Zagajewska K., Piątkowska M., Goryca K. et al.* GWAS links variants in neuronal development and actin remodeling related loci with pseudoexfoliation syndrome without glaucoma // *Exp. Eye Res.* 2018. V. 168. P. 138–148.
22. *Aung T., Ozaki M., Mizoguchi T. et al.* A common variant mapping to *CACNA1A* is associated with susceptibility to exfoliation syndrome // *Nat. Genet.* 2015. V. 47. P. 387–392.
23. *Ponomarenko I., Reshetnikov E., Polonikov A. et al.* Candidate genes for age at menarche are associated with endometriosis // *Reprod. Biomed. Online.* 2020. V. 41(5). P. 943–956.  
<https://doi.org/10.1016/j.rbmo.2020.04.016>
24. *Minyaylo O., Ponomarenko I., Reshetnikov E. et al.* Functionally significant polymorphisms of the *MMP-9* gene are associated with peptic ulcer disease in the Caucasian population of Central Russia // *Sci. Rep.* 2021. V. 11(1). P. 13515.  
<https://doi.org/10.1038/s41598-021-92527-y>
25. *Golovchenko O., Abramova M., Ponomarenko I. et al.* Functionally significant polymorphisms of *ESR1* and *PGR* and risk of intrauterine growth restriction in population of Central Russia // *Eur. J. Obstet. Gynecol. Reprod. Biol.* 2020. V. 253. P. 52–57.  
<https://doi.org/10.1016/j.ejogrb.2020.07.045>
26. *Litovkina O., Nekipelova E., Dvornyk V. et al.* Genes involved in the regulation of vascular homeostasis determine renal survival rate in patients with chronic glomerulonephritis // *Gene.* 2014. V. 546(1). P. 112–116.  
<https://doi.org/10.1016/j.gene.2014.04.020>
27. *Moskalenko M., Ponomarenko I., Reshetnikov E. et al.* Polymorphisms of the matrix metalloproteinase genes are associated with essential hypertension in a Caucasian population of Central Russia // *Sci. Rep.* 2021. V. 11(1). P. 5224.  
<https://doi.org/10.1038/s41598-021-84645-4>
28. *Jaimes M., Rivera-Parra D., Miranda-Duarte A. et al.* Prevalence of high-risk alleles in the *LOXL1* gene and its association with pseudoexfoliation syndrome and exfoliation glaucoma in a Latin American population // *Ophthalmic Genet.* 2012. V. 33. P. 12–17.

29. Pasutto F., Zenkel M., Hoja U. et al. Pseudoexfoliation syndrome-associated genetic variants affect transcription factor binding and alternative splicing of *LOXL1* // Nat. Commun. 2017. V. 8. P. 15466.
30. Álvarez L., García M., González-Iglesias H. et al. *LOXL1* gene variants and their association with pseudoexfoliation glaucoma (XFG) in Spanish patients // BMC Med. Genet. 2015. V. 16. P. 72.
31. Aragon-Martin J.A., Ritch R., Liebmann J. et al. Evaluation of *LOXL1* gene polymorphisms in exfoliation syndrome and exfoliation glaucoma // Mol. Vis. 2008. V. 14. P. 533–541.
32. Yaz Y., Yildirim N., Aydın Yaz Y. et al. Three single nucleotide polymorphisms of *LOXL1* in a Turkish population with pseudoexfoliation syndrome and pseudoexfoliation glaucoma // Turk. J. Ophthalmol. 2018. V. 48(5). P. 215–220.  
<https://doi.org/10.4274/tjo.83797>
33. Kobakhidze N., Tabagari S., Chichua G. *LOXL1* gene variants in association with exfoliation syndrome in georgian population // Georgian Med. News. 2019. V. 286. P. 32–36.
34. Tang J.Z., Wang X.Q., Ma H.F. et al. Association between polymorphisms in lysyl oxidase-like 1 and susceptibility to pseudoexfoliation syndrome and pseudoexfoliation glaucoma // PLoS One. 2014. V. 9(3). P. e90331.
35. Chen L., Jia L., Wang N. et al. Evaluation of *LOXL1* polymorphisms in exfoliation syndrome in a Chinese population // Mol. Vis. 2009. V. 15. P. 2349–2357.
36. Hauser M.A., Aboobakar I.F., Liu Y. et al. Genetic variants and cellular stressors associated with exfoliation syndrome modulate promoter activity of a lncRNA within the *LOXL1* locus // Hum. Mol. Genet. 2015. V. 24. P. 6552–6563.
37. Cissé Y., Bai L., Meng T. lncRNAs in genetic basis of glaucoma // BMJ Open Ophthalmol. 2018. V. 3. P. e000131.
38. Aboobakar I.F., Qin X., Stamer W.D. et al. A lncRNA in the *LOXL1* locus modulates the expression of genes relevant to exfoliation glaucoma pathobiology // Invest Ophthalmol. Vis. Sci. 2016. V. 57. P. 790.

## Analysis of Associations of Polymorphic Loci of the *LOXL1* Gene with the Development of Primary Open-Angle Glaucoma in Women of the Central Chernozem Region of Russia

N. V. Eliseeva<sup>a</sup>, I. V. Ponomarenko<sup>a</sup>, and M. I. Churnosov<sup>a, \*</sup>

<sup>a</sup>Belgorod State University, Belgorod, 308015 Russia

\*e-mail: [churnosov@bsu.edu.ru](mailto:churnosov@bsu.edu.ru)

A replicative study of associations of the *LOXL1* gene polymorphism with primary open-angle glaucoma (POAG) in the female population of the Central Chernozem region of Russia was performed. The work was done in the design “patients-control”. The study sample consisted of 290 women with POAG and 220 women in the control group. Three polymorphic loci of the *LOXL1* gene (rs2165241, rs4886776, rs893818) were genotyped by PCR (Tag-Man probe technology). The study revealed that SNPs of the *LOXL1* gene (rs2165241, rs4886776, rs893818) are associated with POAG in women of the Central Chernozem region of Russia: allele C rs2165241 (OR = 0.33–0.45 at  $p_{\text{perm}} \leq 0.0005$ ), alleles A rs4886776 (OR = 0.62–0.63  $p_{\text{perm}} \leq 0.031$ ) and rs893818 (OR = 0.53–0.62  $p_{\text{perm}} \leq 0.007$ ) and the CAA haplotype rs2165241–rs4886776–rs893818 (OR = 0.56  $p_{\text{perm}} = 0.022$ ) is protective for the development of the disease, and the TGG haplotype (OR = 2.19  $p_{\text{perm}} = 0.001$ ) is associated with an increased risk of developing POAG in women.

**Keywords:** *LOXL1*, associations, primary open-angle glaucoma, polymorphism, women.

УДК 577.29

## МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКАЯ ИДЕНТИФИКАЦИЯ БЕЗЫМЯННЫХ ЗАХОРОНЕНИЙ ПЕРВОЙ ПОЛОВИНЫ XVI в. ИЗ НЕКРОПОЛЯ ВОЗНЕСЕНСКОГО СОБОРА МОСКОВСКОГО КРЕМЛЯ

© 2022 г. И. В. Корниенко<sup>1, 2, 4, \*</sup>, Т. Д. Панова<sup>3</sup>, Т. Г. Фалеева<sup>1, 2, 4, 5</sup>,  
О. Ю. Арамова<sup>1, 2, 4</sup>, Ю. С. Сидоренко<sup>1, 6</sup>

<sup>1</sup>Южный научный центр Российской академии наук, Ростов-на-Дону, 344006 Россия

<sup>2</sup>Калмыцкий научный центр Российской академии наук, Элиста, 358000 Россия

<sup>3</sup>Государственный историко-культурный музей-заповедник “Московский Кремль”, Москва, 103132 Россия

<sup>4</sup>Южный федеральный университет, Академия биологии и биотехнологии им. Д.И. Иванковского, Ростов-на-Дону, 344090 Россия

<sup>5</sup>111 Главный государственный центр судебно-медицинских и криминалистических экспертиз, филиал № 2, Ростов-на-Дону, 344000 Россия

<sup>6</sup>Национальный медицинский исследовательский центр онкологии, Ростов-на-Дону, 344037 Россия

\*e-mail: ikornienko@yandex.ru

Поступила в редакцию 08.06.2021 г.

После доработки 15.07.2021 г.

Принята к публикации 29.07.2021 г.

Некрополь русских великих княгинь и цариц, погребенных в период 1407–1731 гг. в Вознесенском соборе Московского Кремля, претерпел значительные разрушения, многие саркофаги оказались безымянными, что привело к затруднениям в установлении личности находящихся в нем скелетов. Настоящая работа посвящена идентификации неустановленных останков, предположительно принадлежащих потомкам великой княгини Московской Софьи Палеолог (1455–1503), – ее дочери Евдокии Ивановне и внучке Анастасии Петровне. В исторических летописях сохранились сведения о захоронении в этом храме в первой половине XVI в. двух представительниц великокняжеской семьи. Генетический анализ позволил успешно выполнить непрямую идентификацию двух захоронений периода русского Средневековья из некрополя Вознесенского монастыря Московского Кремля и установить родственные взаимосвязи погребенных. Исследование полиморфных микросателлитных локусов аутосомной ДНК, а также гипервариабельных сегментов митохондриальной ДНК (мтДНК) трех женских скелетов позволило выявить их близкую родственную связь. Анализ гипервариабельных локусов мтДНК показал, что все три скелета, включая останки великой княгини Софьи Палеолог, относятся к одной митогруппе – HV0, которая наиболее распространена среди европейских популяций. Результаты настоящей работы стали крайне важной составляющей для реконструкции истории формирования в XVI в. одной из двух самых престижных усыпальниц Московского Кремля. Полученные данные могут помочь в реконструкции семейных захоронений высшей знати средневекового Русского государства.

**Ключевые слова:** некрополь Вознесенского монастыря, древние кости, древняя ДНК, аутосомные STR-локусы, митотипы.

**DOI:** 10.31857/S0016675822020072

Проект по изучению захоронений некрополя русских великих княгинь и цариц Вознесенского монастыря Московского Кремля стал заметным явлением в российской исторической науке [1]. Вознесенский собор служил местом погребения жен и дочерей московских великих князей и русских царей в период с 1407 по 1731 г. (рис. 1). Постройки этой обители и ее храм-усыпальница были разобраны в 1929–1931 гг. Сотрудникам музея “Оружейная палата” удалось спасти и сохранить белокаменные саркофаги с останками представи-

тельниц правящих династий России: они находятся в Кремле, в подземной палате возле южной стены Архангельского собора. Но внезапное уничтожение этого женского монастыря стало причиной возникновения многих вопросов, которые приходится решать при реконструкции его истории.

Это относится и к захоронениям некрополя Вознесенского собора. Так, при исследовании группы безымянных саркофагов XVI в. (это были

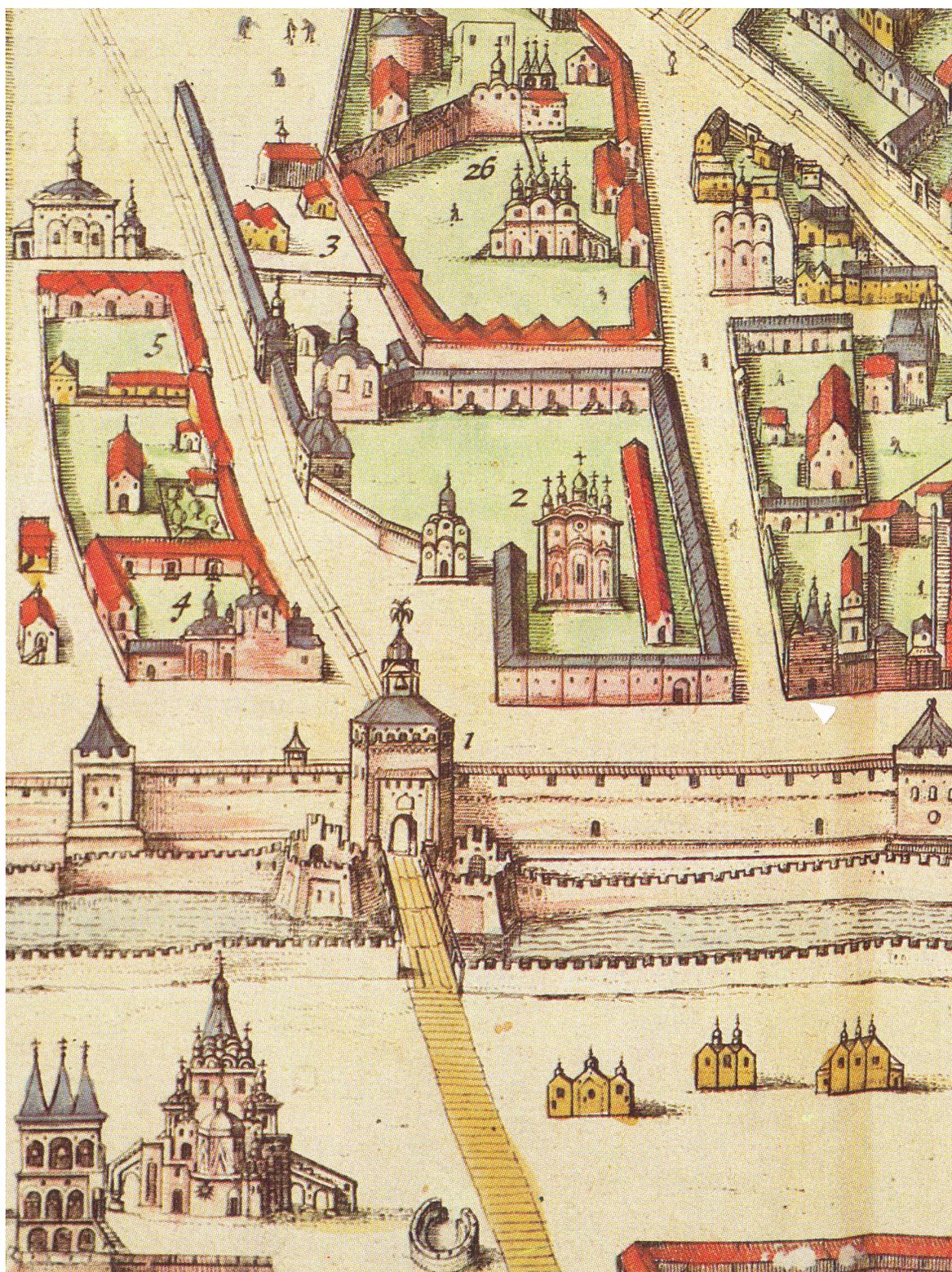


Рис. 1. “Кремленаград”. Вознесенский монастырь на плане Московского Кремля начала XVII в. (под № 26). Музеи Московского Кремля.

взрослые и детские погребения) выяснилось, что все они компактно размещались возле южных дверей храма-усыпальницы. Об этом свидетельствовали данные “Дневника вскрытия”, который вели осенью 1929 г. сотрудники музеев Кремля, фотоматериалы и некоторые зарисовки архитектора П.Н. Максимова того же года (рис. 2).

Схематические планы П.Н. Максимова показывают, что саркофаги XVI в. помещались на значительной глубине, так как сверху позднее были совершены захоронения XVII столетия. Это хорошо видно и на ряде фотографий, сделанных в 1929 г. (рис. 3).

Среди этих безымянных саркофагов три взрослых относятся к первой половине XVI в. (в одном из них скелет почти полностью утрачен). Крышки этих гробов не имеют эпитафий, что позволяет исключить их из числа захоронений великих княгинь и цариц. Тем более что погребения всех первых лиц государства XVI в. присутствуют в некрополе Вознесенского собора, места их могил известны, а крышки их каменных гробов начиная с 1538 г. (захоронение великой княгини Елены Глинской) снабжены надписями. Следовательно, безымянные погребения должны быть отнесены к числу женщин второго ряда — это великие княжны или удельные княгини.

В письменных источниках сохранились сведения о захоронении в этом храме в первой половине XVI в. двух представительниц великокняжеской семьи. В 1513 г. в феврале “преставилась благоверная великая княжна Евдокея, Петра царевича жена, в вторник на Федоровой недели; и положиша тело ея в церкви Възнесения на Москве внутри града” [2]. Речь в летописи идет о дочери великого князя Ивана III Васильевича и его второй жены гречанки Софьи Палеолог. Эта девочка родилась в семье московского государя в 1483 г. [1], хотя в исторической литературе приводится иногда и другая дата — 1492 г. В январе 1506 г. великую княжну Евдокию выдали замуж за крещеного в православие казанского царевича Кудайкула, принявшего имя Петр. В семье, просуществовавшей недолго, родились две дочери, но точные даты их появления на свет остаются неизвестными.

Под 1540 г. (иногда под 1541 г.) в русских летописях отмечены смерть и погребение в некрополе Вознесенского собора внучки великого князя Ивана III Анастасии Петровны — дочери великой княжны Евдокии Ивановны. Она умерла “лета 7049... тое же осени, декабря в 17 ден, в пяток... а положена внутри града на Москве у Вознесения, возле матери ея” [3]. Эта княжна родилась между 1506 и 1513 гг., но неизвестно, была ли она старшей или младшей дочерью великой княжны Евдокии (сестру ее звали Анной). Когда был заключен брак Анастасии с выходцем из Литвы служилым князем

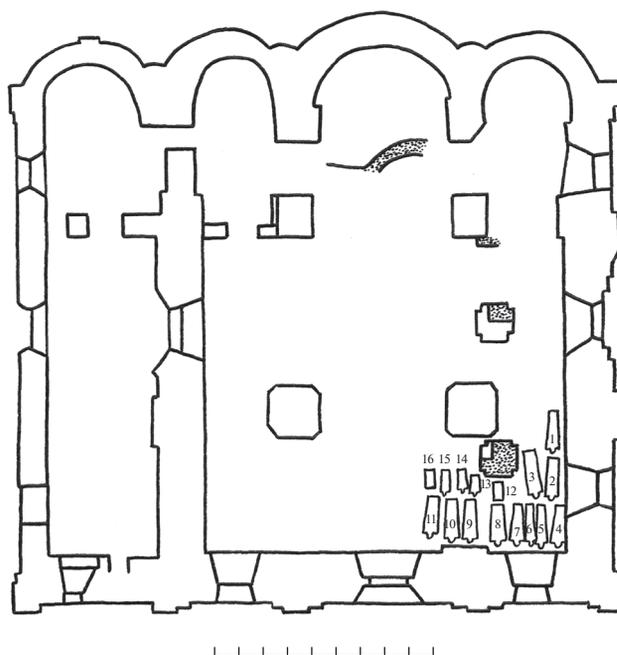


Рис. 2. План части некрополя Вознесенского собора с захоронениями. Архитектор П.Н. Максимов. 1929 г. Государственный научно-исследовательский музей архитектуры им. А.В. Щусева.

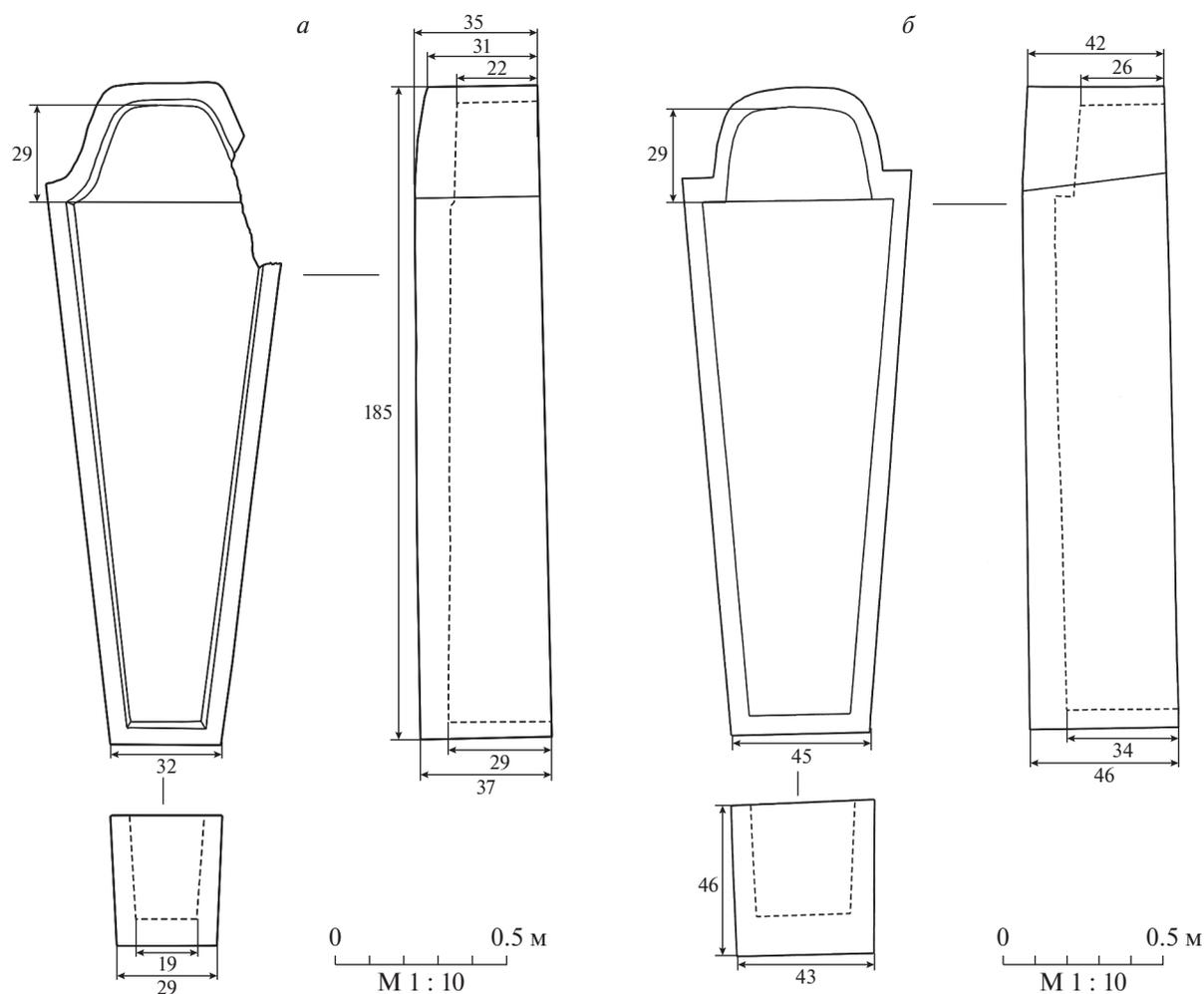
Федором Михайловичем Мстиславским мы не знаем: муж ее скончался летом 1540 года [4].

Учитывая возможную близкородственную связь типа “мать—дочь”, потребовалась идентификация захоронений этих двух женщин, с именами которых на первом этапе предположительно связали безымянные саркофаги, условно обозначенные как “ПН-36” и “ПН-48”. Сразу отметим, что на других участках усыпальницы Вознесенского собора безымянных захоронений XV—первой половины XVI в. не было. Поэтому в ходе исследований стало ясно, что упомянутые в хрониках могилы двух представительниц великокняжеской семьи были размещены в наиболее престижной части некрополя и входят в группу безымянных саркофагов, попытка идентификации которых и была предпринята в ходе настоящего исследования.

Для уточнения датировки безымянных саркофагов ПН-36 и ПН-48 в первую очередь уделили внимание их форме. Они отличались от белокаменных гробов XV столетия, так как не имели сужений в ножной части — это имитировало форму человеческого тела (рис. 4). Важно также отметить, что остальные саркофаги XVI столетия (начиная с 1538 г.) имели резные эпитафии на крышках и по-этому проблем с их идентификацией не было. При изучении погребальных сооружений изначально предположили, что саркофаг ПН-48 относится к началу XVI столетия, а гроб ПН-36 может датироваться временем ближе к середине этого века.



**Рис. 3.** Погребения некрополя Вознесенского собора в ходе разборки храма. Фотография 1929 г. Видны саркофаги XVII в. (сверху) и под ними – погребения XVI столетия. Музей Московского Кремля.



**Рис. 4.** Белокаменные саркофаги XV и XVI в. Планы и разрезы. Рисунок Т.Д. Пановой. *a* – саркофаг предположительно великой княжны Евдокии Ивановны; *b* – саркофаг предположительно великой княжны Анастасии Петровны.

Далее для определения принадлежности этих двух захоронений выбрали два метода исследования останков – антропологический и генетический анализ. Для их осуществления были привлечены также костные останки великой княгини Софьи Палеолог, так как она являлась матерью великой княжны Евдокии Ивановны и бабушкой Анастасии Петровны.

Антропологический метод показал, что обе женщины, погребенные в них, умерли в возрасте до тридцати лет: они прожили на свете от двадцати трех до двадцати пяти или чуть более лет. При этом анализе костных останков важные признаки родственных связей между женщинами из двух безымянных саркофагов и великой княгиней Софьей Палеолог выявила одонтолог Н.И. Халдеева. В своем заключении она отметила высокую степень сходства по формам зубных дуг на нижних челюстях великой княгини Софьи Палеолог и индивида из саркофага ПН-36, предварительно определенного как княжна Анастасия Петровна

[5]. О сходстве верхних зубов индивидов из гробов ПН-48 (Евдокия Ивановна) и ПН-36 (Анастасия Петровна) также свидетельствуют особенности некоторых элементов на верхних первых правых премолярах, верхних вторых правых премолярах и других зубах [6]. В работе С.В. Васильева и соавт. [5] отмечены одновременные редкие одонтологические элементы на зубах рассмотренных останков. Все эти наблюдения позволяют предположить возможные генетические связи всех трех женщин – великой княгини Софьи Палеолог, ее дочери и внучки.

В итоге комплекс косвенных свидетельств (наблюдения за формой саркофагов и общей топографией некрополя Вознесенского собора), а также натурные исследования антропологических материалов позволили предположительно идентифицировать два безымянных захоронения и восстановить состав могил возле южных дверей храму усыпальницы русских великих княгинь и цариц. Однако эти предварительные выводы потребовали



**Рис. 5.** Внешний вид костей скелетов ПН-6, ПН-48 и ПН-36. *а* – левая бедренная кость великой княгини Софьи Палеолог; *б* – правая большеберцовая кость великой княгини Московской Софьи Палеолог; *в* – правая бедренная кость скелета предположительно великой княжны Евдокии Ивановны; *г* – правая бедренная кость скелета предположительно великой княжны Анастасии Петровны.

дополнительного подтверждения. Таким важным и убедительным доводом могли стать результаты молекулярно-генетических идентификационных исследований останков, что и было проведено в настоящей работе.

## МАТЕРИАЛЫ И МЕТОДЫ

Для генетического исследования Государственным историко-культурным музеем-заповедником “Московский Кремль” были предоставлены левая бедренная и правая большеберцовая кости скелета великой княгини Софьи Палеолог (год смерти 1503), обозначенного как “ПН-6”; правая бедренная кость скелета предположительно великой княжны Евдокии Ивановны (год смерти 1513), обозначенного как “ПН-48”; правая бедренная кость скелета предположительно великой княжны Анастасии Петровны (год смерти 1541), обозначенного как “ПН-36” (рис. 5).

### *Подготовка помещения для работы с костными объектами*

Пробоподготовку костных объектов проводили в помещении, которое в течение двух дней непрерывно подвергали облучению с помощью пятилампового облучателя-рециркулятора закрытого типа “Дезар-7” (суммарная мощность излучения составляла 100 Вт). Далее полы и стены помещения обрабатывали 10%-ным коммерческим раствором “Белизна”, содержащим активный хлор в виде гипохлорита натрия, и оставляли еще на одни сутки с включенным облучателем-рециркулятором “Дезар-7”.

### *Пробоподготовка костных объектов и выделение ДНК*

Пробоподготовку костных объектов и фенол-органическую экстракцию ДНК из полученного костного порошка проводили согласно протоколу, описанному ранее [7]. Выделение ДНК из образцов № ПН-6 (левая бедренная кость), ПН-6

(правая большеберцовая кость), ПН-48 и ПН-36 проводили в двух независимых параллелях.

*Исследование STR-локусов аутосомной ДНК по системе IdentifilerPlus*

Генотипирование образцов ДНК проводили с использованием тест-системы AmpFLSTR Identifiler Plus PCR Amplification Kit (ThermoFisher Scientific, Applied Biosystems), включающей в себя 15 микросателлитных локусов: *D8S1179*, *D21S11*, *D7S820*, *CSF1PO*, *D3S1358*, *TH01*, *D13S317*, *D16S539*, *D2S1338*, *D19S433*, *vWA*, *TPOX*, *D18S51*, *D5S818* и *FGA*. Амплификацию участков ДНК осуществляли методом ПЦР с помощью термоциклера GeneAmp PCR System 9700 (Applied Biosystems) в режиме эмульсии 1°C/с в течение 30 циклов. Электрофорез образцов проводили с помощью ДНК-анализатора ABI PRISM 3130xl (Applied Biosystems). Инъекция образцов осуществлялась при напряжении 2.0 кВ в течение 20 с. Обработку результатов электрофореза и идентификацию аллелей проводили с помощью программ Gene Mapper ID (версия 3.2) и Gene Mapper ID-X (версия 1.5).

Типирование препаратов ДНК, полученных в ходе независимых экстракций из костных порошков, проводили в четырех параллелях для объектов № ПН-48 и ПН-36, а для объекта № ПН-6 (левая бедренная и правая большеберцовая кости) – в трех параллелях.

Для оценки специфичности реакции амплификации проводили типирование положительно-го (проба с ДНК 9947А) и отрицательного (проба без ДНК) контролей.

Идентификацию аллелей проводили относительно размерного стандарта LIZ-600 и аллельного леддера, прилагаемого к тест-системе AmpFISTR PCR Identifiler Plus Amplification Kit.

*Исследование STR-локусов аутосомной ДНК по системе GlobalFiler*

Генотипирование образцов ДНК проводили с использованием тест-системы GlobalFiler PCR Amplification Kit (ThermoFisher Scientific, Applied Biosystems), включающей в себя 21 микросателлитный локус: *D3S1358*, *vWA*, *D16S539*, *CSF1PO*, *TPOX*, *D8S1179*, *D21S11*, *D18S51*, *D2S441*, *D19S433*, *TH01*, *FGA*, *D22S1045*, *D5S818*, *D13S317*, *D7S820*, *SE33*, *D10S1248*, *D1S1656*, *D12S391*, *D2S133*. Амплификацию участков ДНК осуществляли методом ПЦР с помощью термоциклера GeneAmp PCR System 9700 (Applied Biosystems), используя режим эмульсии “Max ramping mode”, в течение 30 циклов. Электрофорез образцов проводили с помощью ДНК-анализатора ABI PRISM 3500 (Applied Biosystems). Инъекция образцов осуществлялась при напряжении 2.0 кВ в течение 18 с. Обработку результатов электрофореза и идентификацию ал-

лелей проводили с помощью программы Gene Mapper ID-X (версия 1.5). Типирование всех препаратов ДНК проводили в трех параллелях.

Для оценки специфичности реакции амплификации проводили типирование положительно-го (проба с ДНК 007) и отрицательного (проба без ДНК) контролей.

Идентификацию аллелей проводили относительно размерного стандарта LIZ-600 и аллельного леддера, прилагаемого к тест-системе GlobalFiler PCR Amplification Kit.

*Вероятностные расчеты степени родства по результатам типирования аутосомных микросателлитных локусов*

В расчетах были использованы численные значения аллельных частот, выложенные на электронном ресурсе <https://strbase.nist.gov> [8]. Отношения правдоподобия (likelihood ratio, LR) для таких сценариев не прямой идентификации, как родство типа “мать–дочь” и родство типа “бабушка–внучка”, рассчитывались отдельно для каждого STR-локуса на основе стандартных алгоритмов [9] с использованием соответствующих компьютерных программ EasyPA\_In\_1\_Minute и EasyPAnt\_In\_1\_Minute [10–12]. Далее, предполагая независимое наследование аллелей для исследованных полиморфных аутосомных STR-локусов (отсутствие сцепления), отношения правдоподобия были перемножены и рассчитан комбинированный индекс родства (Combined Kinship Index, SKI).

*Амплификация участков митохондриальной (мт) ДНК*

Энзиматическую амплификацию гипервариабельных (*HV1* и *HV2*) участков мтДНК исследуемых объектов проводили в двух параллелях с помощью минипраймеров, представленных в табл. 1.

Конечные концентрации  $MgCl_2$  и *SynTaq* ДНК-полимеразы (Синтол) в ПЦР-смеси составляли соответственно 2 мМ и 0.2 ед./мкл. Энзиматическую амплификацию проводили при помощи термоциклера GeneAmp PCR System 9700 (Applied Biosystems) в режиме эмульсии скорости нагрева-охлаждения, равной 1°C/с. Условия ПЦР приведены в табл. 2.

Продукты амплификации очищали от избытка дезоксинуклеотидтрифосфатов и праймеров с помощью набора реактивов на магнитных частицах CleanMag DNA (Евроген).

*Секвенирование локусов MPS1A, MPS1B, MPS2A, MPS2B, MPS3A, MPS3B, MPS4A, MPS4B*

Циклическое секвенирование амплифицированных участков мтДНК исследуемых объектов

**Таблица 1.** Минипраймеры, используемые для энзиматической амплификации гипервариабельных (*HVI* и *HV2*) участков мтДНК исследуемых объектов

Локус	Наименования праймеров	Последовательности минипраймеров
<i>MPS1A</i>	F15989	5'-CCCAAAGCTAAGATTCTAAT-3'
	R16158	5'-TACTACAGGTGGTCAAGTAT-3'
<i>MPS1B</i>	F16112	5'-CACCATGAATATTGTACGGT-3'
	R16251	5'-GGAGTTGCAGTTGATGT-3'
<i>MPS2A</i>	F16190	5'-CCCCATGCTTACAAGCAAGT-3'
	R16322	5'-TGGCTTTATGTACTATGTAC-3'
<i>MPS2B</i>	F16268	5'-CACTAGGATACCAACAAACC-3'
	R16410-M19	5'-GAGGATGGTGGTCAAGGGA-3'
<i>MPS3A</i>	F34	5'-GGGAGCTCTCCATGCATTTGGTA-3'
	R159	5'-AAATAATAGGATGAGGCAGGAATC-3'
<i>MPS3B</i>	F109	5'-GCACCCTATGTCGCAGTATCTGTC-3'
	R240	5'-TATTATTATGTCCTACAAGCA-3'
<i>MPS4A</i>	F151	5'-СТАТАТТТТТАТСГСАСТ-3'
	R292	5'-АТТТТТТТГТТАТГАТГТСТ-3'
<i>MPS4B</i>	F220	5'-TGCTTGTAGGACATAATAAT-3'
	R389	5'-СТГГТТАГГСТГГТГТТАГГ-3'

**Таблица 2.** Условия энзиматических амплификаций гипервариабельных участков мтДНК исследуемых объектов

Локус мтДНК	Параметры ПЦР		
	этап 1, предварительная инкубация	этап 2, параметры цикла	число циклов
<i>MPS1A</i>	95°C 300 с	95°C 20 с, 50°C 30 с, 72°C 30 с	38
<i>MPS1B</i>	95°C 300 с	95°C 20 с, 50°C 30 с, 72°C 30 с	42
<i>MPS2A</i>	95°C 300 с	95°C 20 с, 46°C 30 с, 72°C 30 с	38
<i>MPS2B</i>	95°C 300 с	95°C 20 с, 48°C 30 с, 72°C 30 с	42
<i>MPS3A</i>	95°C 300 с	95°C 20 с, 54°C 30 с, 72°C 30 с	38
<i>MPS3B</i>	95°C 300 с	95°C 20 с, 46°C 30 с, 72°C 30 с	42
<i>MPS4A</i>	95°C 300 с	95°C 20 с, 45°C 30 с, 72°C 30 с	38
<i>MPS4B</i>	95°C 300 с	95°C 20 с, 46°C 30 с, 72°C 30 с	42

проводили с помощью тест-набора для секвенирования DNA Sequencing Kit, BigDye Terminator Cycle Sequencing Ready Reaction, версии 1.1 (Applied Biosystems) с одним из вышеуказанных праймером (табл. 1) с помощью термоциклера GeneAmp PCR System 9700 (Applied Biosystems). Продукты секвенирования очищали на колонках CentriSep (Applied Biosystems). Идентификацию ампликонов проводили с помощью ДНК-анализатора ABI PRISM 3130xl (Applied Biosystems). Инъекция образцов осуществлялась при напряжении 1.5 кВ в течение 20 с. Обработку результатов секвенирования проводили с помощью программы Sequencing Analysis

(версия 5.3). Завершающий этап анализа нуклеотидных последовательностей мтДНК выполняли с помощью программы BioEdit Sequence Alignment Editor (версия 7.0.5.2). Полиморфизм нуклеотидных последовательностей *MPS1A*, *MPS1B*, *MPS2A*, *MPS2B*, *MPS3A*, *MPS3B*, *MPS4A*, *MPS4B* определяли в сравнении с Кембриджской референсной последовательностью (CRS) [13].

Статистический анализ выявленных митотипов проводили с использованием базы данных The EDNAP Mitochondrial DNA Population Database (EMPOP) (<https://empop.online/>). Вероят-

**Таблица 3.** Обобщенные результаты типирования 21 STR-локуса аутосомной ДНК костей скелетов ПН-6, ПН-48 и ПН-36 (тест системы Identifiler Plus и GlobalFiler)

Локус	ПН-6	ПН-48	ПН-36
<i>D8S1179</i>	9, 10	10, 13	10, 13
<i>D21S11</i>	29, 31.2	31.2, ?	30, 31
<i>D7S820</i>	–	11, 11	11, 12
<i>CSFIPO</i>	9, 10	10, 10	10, 11
<i>D3S1358</i>	15, 17	17, 18	17, 18
<i>TH01</i>	9.3, 9.3	8, 9.3	7, 8
<i>D13S317</i>	8, 11	8, 12	8, 12
<i>D16S539</i>	11, ?	11, ?	9, 11
<i>D2S1338</i>	–	24, ?	18, 25
<i>D19S433</i>	14, 15.2	13, 15.2	13, 15.2
<i>vWA</i>	17, 17	16, 17	14, 17
<i>TPOX</i>	–	8, 9	9, 9
<i>D18S51</i>	–	16, 17	16, 18
<i>D5S818</i>	11, 12	11, 12	12, 13
<i>FGA</i>	23, ?	–	22, 23
<i>D2S441</i>	10, 11	10, 11	10, 14
<i>D22S1045</i>	–	15, ?	15, 16
<i>SE33</i>	–	–	17, 25.2
<i>D10S1248</i>	–	13, 15	15, 16
<i>D1S1656</i>	15, ?	–	15, 16
<i>D12S391</i>	–	23, ?	20, 23
<i>Amelogenin</i>	X, X	X, X	XX

Примечание. Аллели обозначены номерами в соответствии с принятой номенклатурой; символ “–” обозначает, что идентифицировать генетический профиль локуса не представилось возможным; символ “?” обозначает, что при типировании локуса возможен эффект выпадения аллелей (allelic drop-out).

ность принадлежности митотипов к одной матрице рассчитывали по формуле

$$W = 1/(1 + 1/LR) \times 100\%,$$

где  $LR = 1/p_m$ , где  $p_m$  – частота митотипа по базе EMPOP.

## РЕЗУЛЬТАТЫ

Результаты исследования аутосомных STR-локусов скелетов великой княгини Софьи Палеолог (условное обозначение ПН-6), предположительно великой княжны Евдокии Ивановны (условное обозначение ПН-48) и предположительно великой княжны Анастасии Петровны (условное обозначение ПН-36) приведены в табл. 3.

Результаты типирования локуса *Amelogenin* подтвердили женский пол всех трех скелетов (табл. 3).

Практически все полученные препараты ДНК имели признаки деградации. Только для одного из трех скелетов – ПН-36 (предположительно великой княжны Анастасии Петровны) удалось по-

лучить полный генетический профиль по всем микросателлитным локусам.

В ходе энзиматической амплификации ряда локусов матричная активность ДНК объектов ПН-6 и ПН-48 либо отсутствовала, либо наблюдались невоспроизводимые в параллелях аллельные варианты. В связи с низкой концентрацией активной ДНК-матрицы устойчивые результаты генотипирования скелетов ПН-6 и ПН-48 получены соответственно только по 10 и 13 аутосомным STR-локусам (табл. 3). Сравнительный анализ результатов типирования этих аутосомных микросателлитных локусов показал, что по девяти из них (*D8S1179*, *CSFIPO*, *D3S1358*, *TH01*, *D13S317*, *D19S433*, *vWA*, *D5S818* и *D2S441*) выявлены совпадения в аллельных парах ПН-6 и ПН-48 (табл. 4). Совпадение генетических признаков не исключает родства типа “мать–дочь” скелетов ПН-6 и ПН-48, а вероятность такого типа родства составляет 99.84% (табл. 4).

Следующим этапом нашей работы была проверка гипотезы о наличии близких родственных отношений между скелетами ПН-48 и ПН-36, что

**Таблица 4.** Расчет комбинированного индекса и вероятности материнства великой княгини Софьи Палеолог по отношению к женщине, чей скелет исследован под условным обозначением ПН-48

Локус	ПН-6	ПН-48	Частоты совпадающих аллелей	Формула расчета родства типа “мать–дочь”	Индекс родства $KI(LR)$
<i>D3S1358</i>	15, 17	17, 18	$p_{17} = 0.2105$	$1/(4 \times p_{17})$	1.1876
<i>vWA</i>	17, 17	16, 17	$p_{17} = 0.2839$	$1/(2 \times p_{17})$	1.7612
<i>CSFIPO</i>	9, 10	10, 10	$p_{10} = 0.2202$	$1/(2 \times p_{10})$	2.2707
<i>D8S1179</i>	9, 10	10, 13	$p_{10} = 0.1025$	$1/(4 \times p_{10})$	2.4390
<i>D2S441</i>	10, 11	10, 11	$p_{10} = 0.2105, p_{11} = 0.3435$	$1/(4 \times p_{10}) + 1/(4 \times p_{11})$	1.9155
<i>D19S433</i>	14, 15.2	13, 15.2	$p_{15.2} = 0.0360$	$1/(4 \times p_{15.2})$	6.9444
<i>TH01</i>	9.3, 9.3	8, 9.3	$p_{9.3} = 0.3449$	$1/(2 \times p_{9.3})$	1.4497
<i>D5S818</i>	11, 12	11, 12	$p_{11} = 0.3560, p_{12} = 0.3878$	$1/(4 \times p_{11}) + 1/(4 \times p_{12})$	1.3469
<i>D13S317</i>	8, 11	8, 12	$p_8 = 0.1205$	$1/(4 \times p_8)$	2.0747
Совокупный индекс родства SKI					624.16
Оценка вероятности родства					99.84%

косвенно бы указывало на принадлежность скелета ПН-36 великой княжне Анастасии Петровне, дочери великой княжны Евдокии Ивановны и внучки великой княгини Софьи Палеолог. Для этого был проведен сравнительный анализ генотипов, который показал совпадение генетических признаков объектов ПН-48 и ПН-36, как минимум, по одному из пары аллелей по 13 аутосомным STR-локусам (*D8S1179*, *D7S820*, *CSFIPO*, *D3S1358*, *TH01*, *D13S317*, *D19S433*, *vWA*, *TPOX*, *D18S51*, *D5S818*, *D2S441* и *D10S1248*) (табл. 3). Таким образом, было установлено, что генетический материал ПН-36 мог быть получен по наследству от ПН-48. Относительно высокая вероятность родства (99.994%) подтверждает материнство объекта ПН-48 (предположительно Евдокии Ивановны) в отношении объекта ПН-36 (предположительно Анастасии Петровны) (табл. 5).

Результаты исследования гипервариабельных локусов мтДНК показали, что все три скелета относятся к одной митогруппе – HV0 (табл. 6).

Митотип, представленный в табл. 6, в основном встречается среди европейских популяций (<https://empor.online/>), где его средняя частота составляет 1.22% (доверительный интервал по методу Клоппера–Пирсона – от 1.03 до 2.14%). Наиболее широко митотип (*16298C 72C 263G 315.1C*) распространен в Польше (2.24%), Швеции (1.89%), Греции (1.71%) и Испании (1.70%).

Таким образом, вероятность события, что все три скелета (ПН-6, ПН-48 и ПН-36) действительно относятся к одной материнской линии, а не случайно подходят к ней по своему профилю мтДНК, составляет 98.79% ( $1/(1 + 1/LR) \times 100\%$ , где  $LR = 1/0.0122$ ).

Учитывая независимый характер наследования аутосомных и митохондриальных ДНК-маркеров для каждой пары (ПН-6 и ПН-48, ПН-48 и ПН-36), был рассчитан совокупный индекс родства типа “мать–дочь”, как произведение индексов родства, рассчитанных отдельно для аутосомных STR-локусов и митотипов [14].

Для пары скелетов ПН-6 и ПН-48 совокупный индекс родства составил ( $LR_a \times LR_m$ )  $624.16 \times 81.97 = 51152.4$ ; где  $LR_a$  – комбинированный индекс родства, рассчитанный для аутосомных STR-маркеров (табл. 4),  $LR_m$  – индекс родства, рассчитанный по частоте совпадающих митотипов (*16298C 72C 263G 315.1C*) в базе EMPOR (табл. 6). Полученные результаты с высокой вероятностью (99.998%) указывают на наличие родства типа “мать–дочь” между великой княгиней Софьей Палеолог и объектом ПН-48, что также косвенно подтверждает принадлежность скелета, обозначенного как ПН-48, великой княжне Евдокии Ивановне, дочери великой княгини Софьи Палеолог.

Величина вероятности материнства объекта ПН-48 по отношению к объекту ПН-36, рассчитанная через совокупный индекс родства (1471434) составила 99.99993%, что подтверждает первоначальную гипотезу о принадлежности скелета ПН-36 великой княжне Анастасии Петровне. Костные останки ПН-36 по сравнению с ПН-48 визуально имели значительно лучшую сохранность [15], что согласуется с результатами ДНК-исследования всех скелетов (табл. 3).

**Таблица 5.** Расчет комбинированного индекса и вероятности материнства объекта ПН-48 (предположительно Евдокии Ивановны) по отношению к объекту ПН-36 (предположительно Анастасии Петровны)

Локус	ПН-48	ПН-36	Частоты совпадающих аллелей	Формула расчета родства типа "мать-дочь"	Индекс родства $KI(LR)$
<i>D3S1358</i>	17, 18	17, 18	$p_{17} = 0.2105, p_{18} = 0.1510$	$1/(4 \times p_{17}) + 1/(4 \times p_{18})$	2.8433
<i>vWA</i>	16, 17	14, 17	$p_{17} = 0.2839$	$1/(4 \times p_{17})$	0.8806
<i>CSF1PO</i>	10, 10	10, 11	$p_{10} = 0.2202$	$1/(2 \times p_{10})$	2.2707
<i>TPOX</i>	8, 9	9, 9	$p_9 = 0.1274$	$1/(2 \times p_9)$	3.9246
<i>D8S1179</i>	10, 13	10, 13	$p_{10} = 0.1025, p_{13} = 0.3296$	$1/(4 \times p_{10}) + 1/(4 \times p_{13})$	3.1975
<i>D18S51</i>	16, 17	16, 18	$p_{16} = 0.1468$	$1/(4 \times p_{16})$	1.7030
<i>D2S441</i>	10, 11	10, 14	$p_{10} = 0.2105$	$1/(4 \times p_{10})$	1.1876
<i>D19S433</i>	13, 15.2	13, 15.2	$p_{13} = 0.2548, p_{15.2} = 0.0360$	$1/(4 \times p_{13}) + 1/(4 \times p_{15.2})$	7.9256
<i>TH01</i>	8, 9.3	7, 8	$p_8 = 0.0956$	$1/(4 \times p_8)$	2.6151
<i>D5S818</i>	11, 12	12, 13	$p_{12} = 0.3878$	$1/(4 \times p_{12})$	0.6447
<i>D13S317</i>	8, 12	8, 12	$p_8 = 0.1205, p_{12} = 0.2687$	$1/(4 \times p_8) + 1/(4 \times p_{12})$	3.0051
<i>D7S820</i>	11, 11	11, 12	$p_{11} = 0.2050$	$1/(2 \times p_{11})$	2.4390
<i>D10S1248</i>	13, 15	15, 16	$p_{15} = 0.1967$	$1/(4 \times p_{15})$	1.2710
Совокупный индекс родства CPI					17951.5
Оценка вероятности родства					99.994%

**Таблица 6.** Результаты исследования гипервариабельных регионов (*HV1* и *HV2*) мтДНК представленных образцов

Скелет	Локусы мтДНК		Частоты митотипов в базе EMPOP			
	<i>HV2</i> (72–340)	<i>HV1</i> (16024–16365)	Европа	Азия	Америка	Африка
ПН-6	<i>72C 263G 315.1C</i>	<i>16298C</i>	$1.219 \times 10^{-2}$	$6.130 \times 10^{-4}$	$3.446 \times 10^{-3}$	$2.002 \times 10^{-3}$
ПН-48	<i>72C 263G 315.1C</i>	<i>16298C</i>	$1.219 \times 10^{-2}$	$6.130 \times 10^{-4}$	$3.446 \times 10^{-3}$	$2.002 \times 10^{-3}$
ПН-36	<i>72C 263G 315.1C</i>	<i>16298C</i>	$1.219 \times 10^{-2}$	$6.130 \times 10^{-4}$	$3.446 \times 10^{-3}$	$2.002 \times 10^{-3}$

## ОБСУЖДЕНИЕ

В процессе генетического исследования установлено, что скелет ПН-48 с высокой долей вероятности принадлежит дочери великого князя Московского Ивана III (1440–1505) и великой княгини Московской Софьи Палеолог (1455–1503) – великой княжне Евдокии Ивановне (год смерти 1513). Также была показана принадлежность скелета ПН-36 великой княжне Анастасии Петровне (год смерти 1541), дочери великой княжны Евдокии Ивановны (1492–1513) и казанского царевича Петра Кудайкула (?–1523).

В результате исследований впервые в российской науке с помощью комплексного молекулярно-генетического анализа полиморфных участков аутомсомной и митохондриальной ДНК проведена непрякая идентификация двух захоронений периода русского Средневековья из некрополя Вознесенского монастыря Московского Кремля и

установлены родственные взаимосвязи погребенных.

В ходе анализа данных письменных источников (летописей) и архивных материалов 1929 г. из фондов московских музеев удалось определить точное место могил великой княжны Евдокии Ивановны и ее дочери Анастасии Петровны [16]. Сегодня их саркофаги (и захоронения в целом) перешли из разряда безымянных, не несущих эпитафий на крышках, в число имеющих точную принадлежность. Это делает биологический материал из погребений великой княжны Евдокии Ивановны и ее дочери Анастасии важным объектом при изучении антропологического типа представительниц русской правящей династии начала XVI столетия, наследственных семейных патологий высшей знати средневекового Русского государства.

Полученные генетические результаты стали крайне важной составляющей для реконструкции истории формирования в XVI в. одной из двух самых престижных усыпальниц Московского Кремля. Результаты комплексного изучения останков великой княгини Софьи Палеолог, ее дочери Евдокии и внучки Анастасии, полученные в настоящей работе, представят значительный интерес для специалистов разных направлений в российской науке.

Пробоподготовка биологических образцов, исследование мтДНК, а также аутосомных STR-локусов по системе Identifiler Plus выполнены в рамках реализации госзадания ЮНЦ РАН, № гр. Проекта 01201363186. Исследование аутосомных микросателлитных локусов по системе GlobalFiler выполнено при финансовой поддержке гранта Правительства Российской Федерации № 075-15-2019-1879 “От палеогенетики до культурной антропологии: комплексное междисциплинарное исследование традиций народов трансграничных регионов: миграции, межкультурное взаимодействие и картина мира”.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. *Панова Т.Д.* Предполагаемое захоронение великой княжны Евдокии Ивановны. 1513 год. Биографический очерк. Некрополь русских великих княгинь и цариц в Вознесенском монастыре Московского Кремля: Материалы исследований. В 4 томах. Т. 3. Погребения XVI—начала XVII века. Часть 1. М.: Музеи Московского Кремля, 2018. С. 66–68.
2. Русские летописи. Т. 3: Воскресенская летопись. Рязань: Узорочье, 1998. Т. 3. С. 334.
3. Русские летописи. Т. 4: Львовская летопись. Рязань: Узорочье, 1999. Т. 4. С. 500.
4. *Панова Т.Д.* Предполагаемое захоронение великой княжны Анастасии Петровны. 1541 год. Биографический очерк. Некрополь русских великих княгинь и цариц в Вознесенском монастыре Московского Кремля: Материалы исследований. В 4 томах. Т. 3. Погребения XVI—начала XVII века. Часть 1. М.: Музеи Московского Кремля, 2018. С. 204–205.
5. *Васильев С.В., Боруцкая С.Б., Халдеева Н.И.* Предполагаемое захоронение великой княжны Евдокии Ивановны. 1513 год. Антропологическое исследование останков. Некрополь русских великих княгинь и цариц в Вознесенском монастыре Московского Кремля: Материалы исследований. В 4 томах. Т. 3. Погребения XVI—начала XVII века. Часть 1. М.: Музеи Московского Кремля, 2018. С. 54–55.
6. *Васильев С.В., Боруцкая С.Б., Халдеева Н.И.* Предполагаемое захоронение великой княжны Евдокии Ивановны. 1513 год. Антропологическое исследование останков. Некрополь русских великих княгинь и цариц в Вознесенском монастыре Московского Кремля: Материалы исследований. В 4 томах. Т. 3. Погребения XVI—начала XVII века. Часть 1. М.: Музеи Московского Кремля, 2018. С. 56.
7. *Корниенко И.В., Фалеева Т.Г., Шурп Т.Г. и др.* Y-гаплогруппы костных останков из курганных погребений хазарского времени на территории юга России // Генетика. 2021. Т. 57. № 4. С. 464–477. <https://doi.org/10.31857/S0016675821040044>
8. <https://strbase.nist.gov/NISTpop.htm> (NIST 1036 Revised U.S. Population Dataset (July 2017)), Excel file of 1036 revised Allele Frequencies.
9. *Evetts I.W., Weir B.S.* Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists. Sunderland, MA: Sinauer Associates, Inc., 1998. 278 p.
10. *Fung W.K.* User-friendly programs for easy calculations in paternity testing and kinship determinations // Forensic Sci. Int. 2003. V. 136. P. 22–34. [https://doi.org/10.1016/S0379-0738\(03\)00218-4](https://doi.org/10.1016/S0379-0738(03)00218-4)
11. *Fung W.K., Hu Y.Q., Chung Y.K.* On statistical analysis of forensic DNA: Theory, methods and computer programs // Forensic Sci. Int. 2006. V. 162. P. 17–23. <https://doi.org/10.1016/j.forsciint.2006.06.025>
12. *Fung W.K., Hu Y.Q.* Statistical DNA Forensics: Theory, Methods and Computation. Sussex, U.K.: John Wiley & Sons, Ltd., 2008. 240 p.
13. *Anderson S., Bankier A.T., Barrell B.G. et al.* Sequence and organization of the human mitochondrial genome // Nature. 1981. V. 290. P. 457–465. <https://doi.org/10.1038/290457a0>
14. *Корниенко И.В., Вакуленко М.Ю., Фалеева Т.Г. и др.* Успешный опыт генетической идентификации останков военнослужащего Красной Армии времен Второй мировой войны // Генетика. 2018. Т. 54. № 4. С. 480–486. <https://doi.org/10.7868/S0016675818040100>
15. *Васильев С.В., Боруцкая С.Б., Халдеева Н.И.* Предполагаемое захоронение великой княжны Анастасии Петровны. 1541 год. Антропологическое исследование останков. Некрополь русских великих княгинь и цариц в Вознесенском монастыре Московского Кремля: Материалы исследований. В 4 томах. Т. 3. Погребения XVI—начала XVII века. Часть 1. М.: Музеи Московского Кремля, 2018. С. 180.
16. *Панова Т.Д.* Топография некрополя Вознесенского собора в XVI—начале XVII века. Некрополь русских великих княгинь и цариц в Вознесенском монастыре Московского Кремля: Материалы исследований. В 4 томах. Т. 3. Погребения XVI—начала XVII века. Часть 1. М.: Музеи Московского Кремля, 2018. С. 12–13.

## DNA Identification of Nameless Burials of the First Half of the 16th Century from the Necropolis of the Ascension Cathedral of the Moscow Kremlin

I. V. Kornienko<sup>a, b, d, \*</sup>, T. D. Panova<sup>c</sup>, T. G. Faleeva<sup>a, b, d, e</sup>, O. Yu. Aramova<sup>a, b, d</sup>, and Yu. S. Sidorenko<sup>a, f</sup>

<sup>a</sup>Federal Research Center the Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don, 344006 Russia

<sup>b</sup>Kalmyk Scientific Center of the Russian Academy of Sciences, Elista, 358000 Russia

<sup>c</sup>State Historical and Cultural Museum and Heritage Site "Moscow Kremlin," Moscow, 103132 Russia

<sup>d</sup>Southern Federal University, Academy of Biology and Biotechnology named after D.I. Ivanovsky, Rostov-on-Don, 344090 Russia

<sup>e</sup>111th Main State Center of Medical Forensic and Criminalistics Examinations, Branch no. 2, Rostov-on-Don, 344000 Russia

<sup>f</sup>National Medical Research Centre for Oncology, Rostov-on-Don, 344037 Russia

\*e-mail: ikornienko@yandex.ru

The necropolis of Russian Grand Duchesses and Tsarinas, buried between 1407 and 1731, in the Ascension Cathedral of the Moscow Kremlin has suffered significant destruction, many sarcophagi turned out to be nameless, which led to difficulties in identifying the skeletons inside them. This paper is devoted to the identification of unidentified remains, presumably belonging to the descendants of the Grand Duchess of Moscow Sophia Palaiologina (1455–1503) – her daughter Evdokia Ivanovna and granddaughter Anastasia Petrovna. Historical chronicles have preserved information about burial of two representatives of the grand ducal family in the Ascension Cathedral in the first half of the 16th century. Genetic analysis made it possible to successfully carry out indirect identification of two burials of the Russian Middle Ages from the necropolis of the Ascension Cathedral in the Moscow Kremlin and to determine the kinship of the buried. The study of polymorphic autosomal microsatellite loci and hypervariable segments of mitochondrial DNA (mtDNA) from three female skeletons revealed their close relationship. Hypervariable loci analysis of mtDNA found in all three of the skeletons, including the remains of Grand Duchess Sophia Palaiologina, belong to the same mtDNA haplogroup – HV0, which is most common in European populations. The results of this work became an extremely important component for the reconstruction of the history of the formation of one of two most prestigious burial vaults in the Moscow Kremlin in the 16th century. The data obtained can help in reconstruction of family burials of the highest nobility of medieval Russia.

**Keywords:** necropolis of the Ascension Cathedral, ancient bones, ancient DNA, autosomal STR loci, mitotypes.

## АЛГОРИТМ ФИЛЬТРАЦИИ НАБОРОВ ДАННЫХ ДНК-ШТРИХКОДИРОВАНИЯ ПРИРОДНЫХ СООБЩЕСТВ

© 2022 г. А. Ю. Краснопеев<sup>1</sup>, \*, Ю. С. Букин<sup>1</sup>, \*\*, С. А. Потапов<sup>1</sup>, О. И. Белых<sup>1</sup>

<sup>1</sup>Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, 664033 Россия

\*e-mail: andrewkrasnopeev@gmail.com

\*\*e-mail: bukinyura@mail.ru

Поступила в редакцию 21.04.2021 г.

После доработки 11.08.2021 г.

Принята к публикации 28.09.2021 г.

Определение таксономического разнообразия сообществ является важным этапом многих экологических и мониторинговых исследований экосистем. В основном для таких работ используют методы ампликонной метагеномики (метабаркодинга). В особенности широко метагеномный подход применяется при анализе сообществ микроорганизмов. Представленность определенного таксона в метагеномной пробе при этом определяется количеством (покрытием) расшифрованных нуклеотидных последовательностей. Несмотря на большие размеры выборок расшифрованных нуклеотидных последовательностей, получаемых на платформах секвенирования нового поколения, открытым остается вопрос о статистической достоверности оцениваемых показателей представленности таксонов. Главным образом это касается минорной компоненты исследуемых сообществ. Для решения задачи по оценке статистической репрезентативности выборок расшифрованных последовательностей ДНК для метабаркодинговых исследований нами разработан алгоритм, основанный на использовании бутстреп-метода. Применение данного алгоритма позволяет оценить достоверительные интервалы, абсолютные и относительные ошибки показателей представленности таксонов в сообществах. Исходный код, примеры работы и рабочие файлы доступны по адресу: [https://github.com/Krasnopeev/otu\\_error\\_filtering](https://github.com/Krasnopeev/otu_error_filtering).

**Ключевые слова:** ДНК-штрихкодирование, биоразнообразие, 16S рРНК, бутстреп-семплирование.

**DOI:** 10.31857/S0016675822020084

Определение таксономического разнообразия сообществ является важным этапом многих экологических и мониторинговых исследований экосистем. В основном для таких работ применяют методы “shotgun” метагеномики (метод дробовика – секвенирования случайных фрагментов) и секвенирования “ампликонов” (ПЦР-продуктов, полученных в результате проведения полимеразной цепной реакции) – ДНК-штрихкодирование [1, 2] с применением технологий секвенирования нового поколения (NGS, Next Generation Sequencing). В задачи такого анализа входят идентификация таксонов организмов, генетический материал которых находился в смеси ДНК, и определение показателей представленности таксонов – количества или доли прочтений, приходящихся на каждый таксон в исследуемом образце.

В “ампликонной” метагеномике (ДНК-штрихкодировании) на этапе пробоподготовки производится ПЦР-амплификация какого-либо широко используемого в филогении и ДНК-штрихкодировании генетического маркера, например, варибельного фрагмента гена 16S или

18S рРНК, фрагмента митохондриального гена *COX1*, ITS1 или ITS2 (internal transcribed spacers) – внутренних спейсеров 1 и 2 ядерного рибосомного кластера генов соответственно. Фрагменты ДНК из смеси продуктов амплификации расшифровываются с помощью методов высокопроизводительного секвенирования. Анализ выборок расшифрованных последовательностей выполняется в два этапа. Первый этап предполагает кластеризацию последовательностей с применением ДНК k-мерного анализа [9, 10] или кластеризацию по генетическим расстояниям, рассчитанным после процедуры выравнивания [11]. На втором этапе последовательности из кластеров сравниваются с референсными базами данных [12–14]. Во многих случаях исследователи оперируют не идентифицированными до вида группами последовательностей, а идентификаторами кластеров ОТЕ – оперативными таксономическими единицами, выделенными на генетических расстояниях уровня вида. Например для бактериальных сообществ при использовании маркерного

гена 16S рРНК порог кластеризации на уровне вида оставляет 3% нуклеотидных замен [9, 11].

Итоговый результат таксономической идентификации расшифрованных фрагментов ДНК при анализе серии образцов в рамках одного исследования – таблица, столбцы которой характеризуют отдельные образцы, а строки содержат информацию о количестве последовательностей (показатели представленности), приходящихся на идентифицированные таксоны. Полученные таблицы представленности таксонов являются отправной точкой при анализе и сравнении сообществ различными методами вычислительной экологии: расчет индексов  $\alpha$ -разнообразия, например, индексы Chao1 [15] и ACE [16] (Abundance Coverage Estimator), которые позволяют оценить реальное (максимальное) количество таксонов в образце, и  $\beta$ -разнообразия – различные виды кластерного анализа (UPGMA, UniFrac), метрические и неметрические статистики, оценка влияния различных биотических и абиотических факторов на состав и структуру сообщества методами многомерного шкалирования (MDS, PCA) и т.д. Все эти подходы широко применяют при анализе бактериальных сообществ [17, 18], например при исследованиях сезонной динамики численности разнообразных групп микроорганизмов [19].

Несмотря на большие размеры выборок расшифрованных нуклеотидных последовательностей, получаемых на платформах секвенирования нового поколения, открытым остается вопрос о статистической достоверности оцениваемых показателей представленности таксонов. В основном это касается минорной компоненты исследуемых сообществ. Для того чтобы избежать ложных выводов при использовании статистик вычислительной экологии, исследователь должен оценить статистическую ошибку определения показателей представленности таксонов в разных метагеномных образцах, связанную с ограничением размера выборки расшифрованных последовательностей ДНК. Задача сводится к расчетам доверительных интервалов в абсолютных и относительных величинах для показателей представленности каждого таксона в пробе, т.е. таксоны, показатели представленности которых определены с большой относительной ошибкой, должны удаляться из анализа перед использованием методов вычислительной экологии.

Одним из наиболее простых методов, который может быть использован для определения абсолютных и относительных ошибок показателей представленности таксонов в пробах при работе с данными секвенирования по методу дробовика и ампликонной метагеномики, является непараметрический бутстреп-анализ [20]. Принцип бутстреп-семплирования заключается в создании случайной выборки методом Монте-Карло для

оценки различных статистик (дисперсия, доверительные интервалы, систематические ошибки и т.д.) анализируемой модели.

Бутстреп-метод широко применяется в вычислительной экологии. Так, например используют бутстреп для определения доверительных интервалов для индекса Шеннона – меры  $\alpha$ -разнообразия сообщества [21].

Цель нашего исследования – разработка универсального алгоритма для расчета абсолютных и относительных ошибок показателей представленности таксонов в выборке данных, полученных при анализе серии проб с помощью методов метагеномики. Анализ предполагает использование непараметрического бутстрепа с реализацией в свободно распространяемой среде программирования R. Одной из задач алгоритма является автоматизированная фильтрация набора данных, связанная с удалением таксонов, показатели представленности которых оценены с большой относительной ошибкой. Исходный код алгоритма, примеры работы и рабочие файлы доступны по адресу: [https://github.com/Krasnopeev/otu\\_error\\_filtering](https://github.com/Krasnopeev/otu_error_filtering).

## МАТЕРИАЛЫ И МЕТОДЫ

### *Описание алгоритма*

В качестве входных данных для работы используется таблица представленности таксонов, полученная в ходе анализа результатов секвенирования ДНК. Она представляет собой матрицу, в которой по вертикали расположены наименования проб, по горизонтали – таксоны. Соответственно в каждой ячейке отмечается количество последовательностей, приходящихся на данный таксон в образце (показатель представленности таксона в пробе). Для удобства обозначим таксон в массиве данных как ОТЕ – оперативная таксономическая единица. Всего в рамках анализа исследуется  $n$  проб (образцов), в которых в совокупности выделено  $k$  ОТЕ. Индекс пробы обозначим  $i$ ,  $i$  принимает значения от 1 до  $n$ . Индекс ОТЕ обозначим  $j$ ,  $j$  принимает значения от 1 до  $k$ . ОТЕ $_{ij}$  – количество расшифрованных последовательностей, приходящихся на таксон с индексом  $j$  в  $i$  пробе (показатель представленности таксона в пробе).

На первом этапе анализа для каждой пробы  $i$  формируется символьный вектор  $S_i$  следующего содержания:

$$S_i = \{(\text{“ОТЕ1”}, \text{“ОТЕ1”}, \dots, \text{“ОТЕ1”}) - \text{ОТЕ}_{i1} \text{ раз}, \\ (\text{“ОТЕ2”}, \text{“ОТЕ2”}, \dots, \text{“ОТЕ2”}) - \text{ОТЕ}_{i2} \text{ раз}, \\ (\text{“ОТЕ3”}, \text{“ОТЕ3”}, \dots, \text{“ОТЕ3”}) - \text{ОТЕ}_{ik} \text{ раз}, \\ (\text{“ОТЕ}j\text{”}, \text{“ОТЕ}j\text{”}, \dots, \text{“ОТЕ}j\text{”}) - \text{ОТЕ}_{ij} \text{ раз}, \\ (\text{“ОТЕ}k\text{”}, \text{“ОТЕ}k\text{”}, \dots, \text{“ОТЕ}k\text{”}) - \text{ОТЕ}_{ik} \text{ раз}\}.$$

**Таблица 1.** Пример таблицы данных с указанием величины относительной ошибки представленности ОТЕ в пробе, полученной в результате работы предложенного алгоритма

	S8	S10	S11	S12	...	$S_n$
ОТЕ1	0.034*	1	0.015*	0.074*	...	0.05*
ОТЕ2	0.045*	0.55	0.22	0.886	...	0.94
ОТЕ3	0.0564*	0.32	0.01*	0.13*	...	1
ОТЕ4	1	1	1	0.02*	...	1
ОТЕ5	0.167*	1	0.178*	1	...	0.01*
...	...	...	...	...	...	0.04*
ОТЕ $i$	0.87	0.01*	0.21	0.44	...	0.03*

Примечание. S с номером – название пробы, ОТЕ с ошибкой определения ниже 0.2 отмечены звездочкой.

Количество элементов  $N_i$  в векторе  $S_i$  задается формулой:

$$N_i = \sum_{j=1}^k \text{ОТЕ}_{ij}. \quad (2)$$

На следующем этапе каждый вектор, как описано в работе [21], подвергается бутстреп-семплированию с применением непараметрического бутстрепа с возвратом. Всего проводится  $h$  циклов бутстреп-семплирования. В результате для пробы  $i$  после каждой итерации получается вектор  $(S_i)_l$  – бутстреп реплика набора с индексом  $l$  ( $l$  принимает значения от 1 до  $h$ ).

На основе векторов  $(S_i)_l$  формируется бутстреп-выборка значений показателей представленности  $(\text{ОТЕ}_{ij})_l$  каждого таксона в каждой пробе. Значения  $(\text{ОТЕ}_{ij})_l$  определяются как количество записей элемента “ОТЕ $j$ ” в векторе  $(S_i)_l$ . Для того, чтобы найти доверительный интервал уровня вероятности  $P$  согласно процедуре бутстреп-анализа [21], значения в векторах  $(\text{ОТЕ}_{ij})_l$  сортируются в порядке возрастания. В отсортированном векторе элемент с индексом  $h(1 - P)/2$  будет  $XMIN_{ij}$  левой границей доверительного интервала, а элемент с индексом  $h(P + 1)/2$   $XMAX_{ij}$  – правой границей доверительного интервала. Далее, рассчитывается  $D_{ij}$  – относительная ошибка определения показателя представленности  $\text{ОТЕ}_{ij}$  по формуле:

$$D_{ij} = \frac{XMAX_{ij} - XMIN_{ij}}{XMAX_{ij}}. \quad (3)$$

По умолчанию в программной реализации алгоритма количество бутстреп-реplik  $h = 1000$ , ширина доверительного интервала задается уровнем вероятности  $P = 0.95$ . Но допустимы изменения этих параметров пользователем.

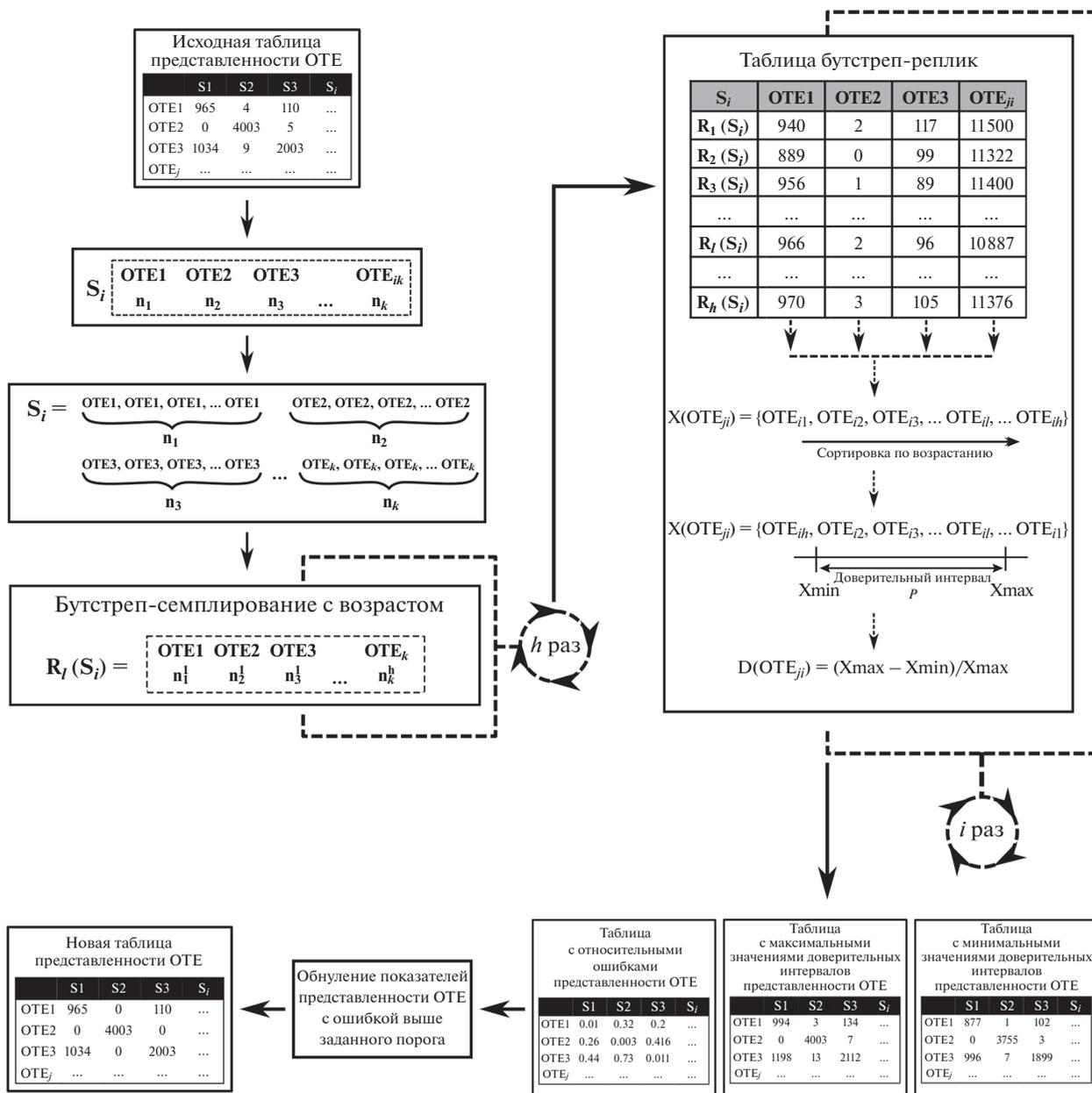
Далее, производится фильтрация – удаление ОТЕ, относительная ошибка определения которых превышает значение порога в 0.2 – минимальный порог достоверности данных, приемле-

мый для биологических исследований [22], заданный по умолчанию в алгоритме. Такие ОТЕ исключаются из дальнейшего анализа (табл. 1). Параметр порога относительной ошибки может меняться пользователем алгоритма. Фильтрация – удаление ОТЕ по уровню ошибки может быть отключена, тогда будет рассчитываться только относительная и абсолютная ошибка определения показателей представленности ОТЕ в пробах.

Соответственно, оставшаяся часть ОТЕ с относительной ошибкой меньше либо равной 0.2 является достоверно представленной. После этой процедуры значения ошибки и количество последовательностей достоверно оцененных ОТЕ переносятся в новую таблицу. В итоге, полученный список таблиц объединяется в одну общую матрицу, имеющую вид оригинальной таблицы представленности, но содержащую в себе только те ОТЕ, которые были оценены с заданным размером ошибки. Отсутствующие значения, которые отброшены в ходе работы алгоритма, заменяются нулями в соответствующих ячейках новой таблицы. Блок-схема работы алгоритма фильтрации данных представлена на рис. 1.

Для работы с программой пользователю необходимо загрузить в рабочую среду R свою таблицу представленности таксонов. Входная функция, выполняющая первичную обработку данных, это `getSub (data, value = 0.95, cutoff = 1, subsample = FALSE, subsize = NULL)` с параметрами по умолчанию, где `data` – таблица представленности таксонов (ОТЕ), `value` – размер выборки от общего числа (по умолчанию 0.95–95% доминирующий пул всех ОТЕ), `cutoff` – значение нижней границы численности анализируемых ОТЕ (по умолчанию – 2, т.е. удаляем из набора все ОТЕ с суммарной численностью 2 и менее), `subsample` – логический оператор, позволяет указать нужно ли выполнять семплирование образцов с размером выборки `subsize` (на усмотрение пользователя).

Далее происходит генерация векторов, имитирующих сообщества в пробе с помощью функции `getOtuSample (data, trimmedby = 0)`, где `data` – это



**Рис. 1.** Блок-схема работы алгоритма. S – идентификатор пробы. Значения показателей представленности в ячейках показаны для примера. R – бутстреп-реплика конкретной пробы. Исходная таблица представленности разбивается на отдельные таблицы соответственно каждой пробе и преобразуется в список. Каждый элемент списка преобразуется в вектор повторяющихся идентификаторов OTE<sub>j</sub> в зависимости от числа последовательностей, принадлежащих этой OTE<sub>j</sub> в рассматриваемой пробе. Далее производится бутстреп-семплирование (генерация случайной выборки) с возвратом для каждого вектора (т.е. пробы) в количестве h штук (значение h устанавливается пользователем, по умолчанию h = 1000). Реплики векторов трансформируются в таблицы представленности соответственно пробам. После этого рассчитываются относительные ошибки представленности ОТЕ в образце и генерируется новая таблица, содержащая в себе достоверные, согласно алгоритму, ОТЕ, а также отдельные таблицы с минимальными и максимальными значениями доверительных интервалов, заданных пользователем.

образец-столбец из таблицы представленности, а `trimmedby` (по умолчанию равно 0) – дополнительный вариант фильтрации уже отдельного образца по аналогии с `cutoff` из функции `getSub`.

Полученный список векторов передается в функцию, которая выполняет непосредственно бут-

стреп-семплирование – `getReplica (data, n = 1000, scaling = "int")`, где `data` – вектор, полученный с помощью функции `getOtuSample`, `n` – число бутстреп-семплирований (по умолчанию 1000), `scaling` – преобразование полученных количественных значений в доли представленности – “percent”,

приведение к натуральному логарифму по основанию 10 – “log”, оригинальные количественные показатели представленности – “int” (по умолчанию).

Последним шагом является расчет относительных ошибок представленности ОТЕ с помощью функции `repError (data, rate = 0.2, filterError = TRUE, p = 0.95)`, где `data` – набор реплик конкретного образца, `rate` – максимальный размер ошибки представленности (задается на усмотрение пользователя), рассчитанной по формуле (3), `filterError` – логический оператор, отвечающий за фильтрацию ОТЕ по значению `rate` (по умолчанию `TRUE`), `p` – доверительный интервал уровня вероятности (по умолчанию 0.95).

В результате работы последней функции пользователь получит список таблиц: (i) таблица представленности ОТЕ, скорректированная с учетом значения `rate` и переменной `filterError`, (ii) таблица с показателями представленности ОТЕ, соответствующих верхнему пределу доверительного интервала, и (iii) таблица с показателями представленности ОТЕ нижнего предела доверительного интервала уровня вероятности. Для удобства пользователя значения в таблицах (ii) и (iii) конвертированы в относительные доли представленности.

Пример пошаговой работы алгоритма можно также найти в файле `README.md` в репозитории по адресу [https://github.com/Krasnopeev/otu\\_error\\_filtering](https://github.com/Krasnopeev/otu_error_filtering).

### *Тестирование алгоритма*

В качестве тестовых данных использовали результаты секвенирования с помощью ДНК-штрихкодирования 28 образцов воды (табл. 2), отобранных в озере Байкал в течение 2017–2018 гг. Образцы получены на разных станциях и глубинах во всех трех котловинах озера как в подледный период (март), так и в период открытой воды (июнь–сентябрь). Исследуемые пробы различались по составу и количественным показателям развития фито- и бактериопланктона, в том числе по доминирующим группам первичных продуцентов. Кроме того, места отбора проб характеризовались различным уровнем антропогенной нагрузки [23–25]. Можно ожидать, что исследованные образцы будут достоверно различаться по составу таксонов и видовому богатству прокариотических микроорганизмов.

Пробы воды фильтровали через поликарбонатные фильтры с диаметром пор 0.2 мкм (Sartorius, Германия). Суммарную ДНК выделяли методом ферментативного лизиса с использованием лизоцима (Roche, Швейцария), протеиназы К (Thermo Scientific, США) и додецилсульфата натрия (VWR Life Science AMRESCO, США) с последую-

щей экстракцией фенолом (Медиген, Россия) хлороформом (Экос-1, Россия) [26]. Библиотеки ампликонов фрагмента V3–V4-гена 16S рРНК получены с помощью пары универсальных праймеров 343F (5'-CTCCTACGGRRSGCAGCAG-3') и 806R (5'-GGACTACNVGGGTWTCTAAT-3') [27] на платформе Illumina MiSeq (Illumina, США) в Евроген (Москва, Россия) с использованием протокола парно-концевых прочтений 2 × 250 пн. Полученные библиотеки ампликонов депонированы в базу данных NCBI GenBank с доступом по номеру PRJNA637453.

Первичную обработку данных секвенирования проводили с использованием программного пакета `Trimmomatic` [28] с параметрами `MINLEN:200`. Для удаления праймеров и адаптеров использовали `cutadapt v1.14` [29]. Полученные в результате первичной обработки парные последовательности сшиты с помощью `FLASH` [30] с параметрами по умолчанию. Удаление химер, прочтений короче 400 и длиннее 435 пн, а также кластеризацию последовательностей в ОТЕ проводили при помощи программного пакета `mothur v. 1.40.5` (<https://mothur.org/>) с порогом кластеризации в 97% гомологии согласно стандартному протоколу для наборов данных, полученных с секвенаторов Illumina [31]. Выходные данные оформлены в виде таблицы представленности ОТЕ. Для каждого образца рассчитывали индексы  $\alpha$ -разнообразия Шеннона, Симпсона и `Chao1`. Показаны результаты корреляционного анализа матриц расстояний по метрике Брея–Кёртиса [32] и Евклидовых расстояний. Корреляционный анализ между матрицами расстояний выполнен с использованием непараметрического критерия Спирмена. Алгоритм реализован при помощи средств языка программирования R [33, 34].

### РЕЗУЛЬТАТЫ

Всего для 28 тестовых проб (образцов) получено 2 109 405 парно-концевых прочтений. В среднем на каждый образец приходилось по 75 336 последовательностей (от 20 155 до 195 476 прочтений). Средняя длина фрагмента составила 418 пн.

До проведения процедуры фильтрации предложенным алгоритмом количество ОТЕ составляло от 824 до 5338 (рис. 2, а). Значения индекса Шеннона варьировали в диапазоне 3.18–4.54. Индекс Симпсона для исходного набора данных составлял от 0.83 до 0.97. Индекс `Chao1` принимал значения от 634 до 3262 (табл. 3).

С помощью разработанного алгоритма в наборе данных идентифицированы ОТЕ, ошибка определения представленности которых была меньше 0.2. Для наборов, состоявших из достоверно представленных ОТЕ, рассчитаны новые значения индексов Шеннона и Симпсона для сравне-

Таблица 2. Характеристика образцов тестового набора

Название пробы	Глубина, м	Станции отбора	Дата отбора 2018 г.
S8	0–5	7 км от п. Листвянка	Март
S10	20–25	7 км от п. Листвянка	Март
S11	50–100	7 км от п. Листвянка	Март
S12	0–5	Ц. р. п. Листвянка–п. Танхой	Июнь
S13	10–15	Ц. р. п. Листвянка–п. Танхой	Июнь
S14	20–25	Ц. р. п. Листвянка–п. Танхой	Июнь
S15	50–100	Ц. р. п. Листвянка–п. Танхой	Июнь
S16	0–5	Ц. р. м. Ухан–м. Тонкий	Июнь
S17	10–15	Ц. р. м. Ухан–м. Тонкий	Июнь
S18	20–25	Ц. р. м. Ухан–м. Тонкий	Июнь
S19	50–100	Ц. р. м. Ухан–м. Тонкий	Июнь
S20	0–5	Пролив Малое Море	Июнь
S22	0–5	Баргузинский залив	Июнь
S23	0–5	Ц. р. м. Елохин–п. Давша	Июнь
S24	10–15	Ц. р. м. Елохин–п. Давша	Июнь
S25	20–25	Ц. р. м. Елохин–п. Давша	Июнь
S26	50–100	Ц. р. м. Елохин–п. Давша	Июнь
S27	0–50	Ц. р. п. Листвянка–п. Танхой	Сентябрь
S28	0–50	Ц. р. м. Ухан–м. Тонкий	Сентябрь
S29	0–50	Ц. р. м. Елохин–п. Давша	Сентябрь
S30	0	Пролив Малое Море	Август
S31	0	Пролив Малое Море	Август
S32	0	Пролив Малое Море	Август
S35	0–5	Ц. р. п. Листвянка–п. Танхой	Август
S36	10–15	Ц. р. п. Листвянка–п. Танхой	Август
S37	25–50	Ц. р. п. Листвянка–п. Танхой	Август
S38	0	м. Лиственичный, около пос. Листвянка	Июнь
S40	0	м. Елохин	Июнь

Примечание. Ц. р. – центр разреза, м. – мыс.

ния с исходными значениями. Несмотря на значительный объем библиотек ампликонов, после использования предложенного алгоритма оценки количество ОТЕ сократилось в несколько раз относительно изначального значения в среднем – с 2854 до 33 ОТЕ (рис. 2). Значения индексов  $\alpha$ -разнообразия Шеннона и Симпсона в большинстве проб до и после фильтрации с помощью алгоритма изменились незначительно (табл. 3) (в среднем значения индексов сократились на 5–10%). Это говорит о том, что после фильтрации наборов данных информация о сложности устройств сообществ практически не была потеряна. Количество потерянных прочтений (рис. 2) при удалении ОТЕ с большими ошибками определения показателей представленности в большинстве случаев составило небольшую долю от общего

числа исходных расшифрованных последовательностей в пробе. В небольшом количестве проб (рис. 2, табл. 3) доля потери прочтений в результате фильтрации достигала 30%, однако и в них значения индексов  $\alpha$ -разнообразия Шеннона и Симпсона изменились в незначительных пределах. Можно сделать вывод, что отфильтрованные последовательности и ОТЕ сами по себе не вносили достаточно весомый вклад в  $\alpha$ -разнообразие исследованных проб (табл. 3).

Ожидаемое резкое изменение характера кривых насыщения также наблюдали в отфильтрованном (рис. 3,б) наборе данных по сравнению с исходным (рис. 3,а). Такое сильное изменение показателей величин насыщения, очевидно, было вызвано значительным уменьшением количества ОТЕ в отфильтрованном наборе, что, в це-

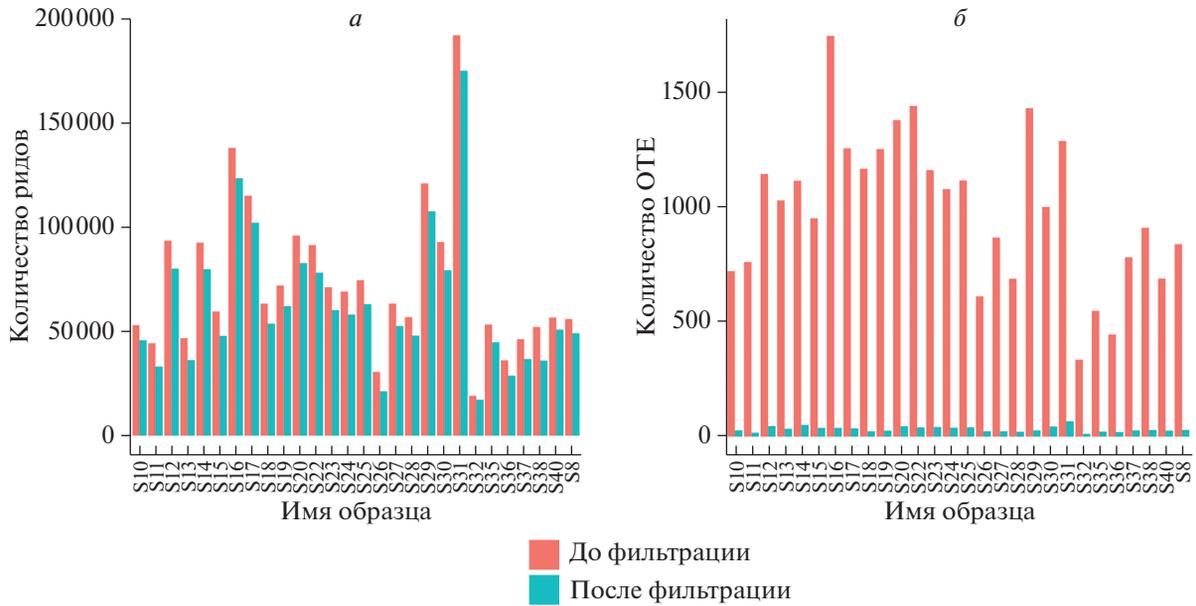


Рис. 2. Гистограммы сравнения количества прочтений – а, ОТЕ – б, до и после фильтрации.

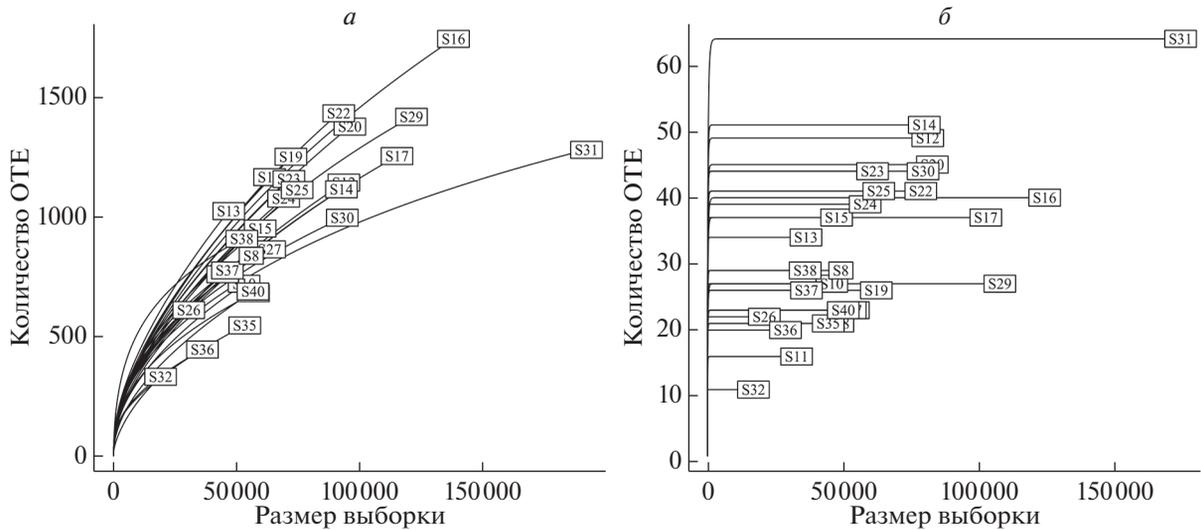


Рис. 3. Кривые насыщения для образцов до (а) и после (б) применения алгоритма фильтрации.

лом, упрощает дальнейший анализ данных для исследователя.

Для сравнения проб по таксономическому составу и уровню представленности различных ОТЕ были рассчитаны матрицы попарных расстояний Брея–Кёртиса и Евклидовых расстояний. Матрицы рассчитаны для наборов данных до фильтрации и после фильтрации – удаления недостоверно представленных ОТЕ (ОТЕ, для которых уровень относительной ошибки определения показателей представленности в пробах был больше заданного по умолчанию порога в 0.2), согласно предложенному в работе алгоритму

оценки. Для сравнения матриц расстояний проведен корреляционный анализ, который показал, что прослеживалась строгая линейная зависимость между матрицами расстояний до и после фильтрации (рис. 4). Это было подтверждено высоким значением коэффициента Спирмена 0.9912 ( $p\text{-value} = 0$ ) для матриц по метрике Брея–Кёртиса и 0.9936 ( $p\text{-value} = 0$ ) для матриц Евклидовых расстояний.

Также при сравнении UPGMA-дендрограм кластеризации образцов на основе матриц расстояний Брея–Кёртиса до и после применения алгоритма фильтрации (рис. 5) показано, что то-

**Таблица 3.** Сравнительная таблица наборов данных до и после фильтрации. Simpson – индекс Симпсона, Shannon – индекс Шеннона, Chao1 – индекс Chao1, N OTE – количество OTE, N прочтений – количество прочтений в образце

Название пробы	До фильтрации					После фильтрации					
	N прочтений	N OTE	Shannon	Simpson	Chao1	N прочтений	N OTE	Shannon	Simpson	доля потеря прочтений	Chao1
S8	57355	2246	3.55	0.92	1914	49098	29	2.66	0.89	14.4	29
S10	54086	1780	3.58	0.93	1547	45769	27	2.7	0.9	15.4	27
S11	45230	1675	3.71	0.92	1666	33095	17	2.24	0.86	26.8	17
S12	95652	3261	4.18	0.96	2732	80117	46	3.31	0.95	16.2	46
S13	49298	3529	4.54	0.97	2547	36243	34	3.16	0.94	26.5	34
S14	94799	3339	4.28	0.97	2558	79801	51	3.43	0.95	15.8	51
S15	61176	2578	4.3	0.96	2133	47860	38	3.2	0.94	21.8	38
S16	141690	5338	3.28	0.83	3262	123465	38	2.34	0.77	12.9	38
S17	117288	3426	3.22	0.83	2455	102160	36	2.34	0.77	12.9	36
S18	65364	3174	3.37	0.85	2937	53721	23	2.19	0.78	17.8	23
S19	74259	3481	3.27	0.83	2752	62143	26	2.14	0.75	16.3	26
S20	99253	4630	4.1	0.95	3090	82802	45	3.1	0.93	16.6	45
S22	94560	4514	3.96	0.93	3194	78175	40	2.88	0.89	17.3	40
S23	74318	4244	4.37	0.97	2871	60195	42	3.3	0.95	19	42
S24	72027	3983	4.26	0.96	2228	58050	39	3.15	0.93	19.4	39
S25	77453	4008	4.27	0.96	2550	63102	41	3.23	0.94	18.5	41
S26	30990	1056	4.18	0.96	1032	21246	23	2.84	0.92	31.4	23
S27	64182	1667	3.68	0.94	1529	52583	23	2.66	0.91	18.1	23
S28	57482	1339	3.47	0.92	1235	47990	21	2.54	0.89	16.5	21
S29	124128	4503	3.49	0.91	2385	107617	27	2.57	0.88	13.3	27
S30	94676	2762	4.11	0.95	1765	79355	44	3.23	0.93	16.2	44
S31	195476	4599	3.94	0.94	1914	175031	67	3.28	0.92	10.5	67
S32	20155	1339	3.18	0.91	634	17232	12	2.24	0.87	14.5	12
S35	53855	1109	3.25	0.89	1179	44775	22	2.33	0.84	16.9	22
S36	36494	824	3.59	0.93	958	28715	20	2.59	0.89	21.3	20
S37	47400	1932	4.07	0.96	1789	36749	27	2.93	0.93	22.5	27
S38	53134	1888	4.48	0.96	1381	35979	29	2.92	0.92	32.3	29
S40	57625	1685	3.29	0.91	1476	50859	26	2.58	0.89	11.7	26

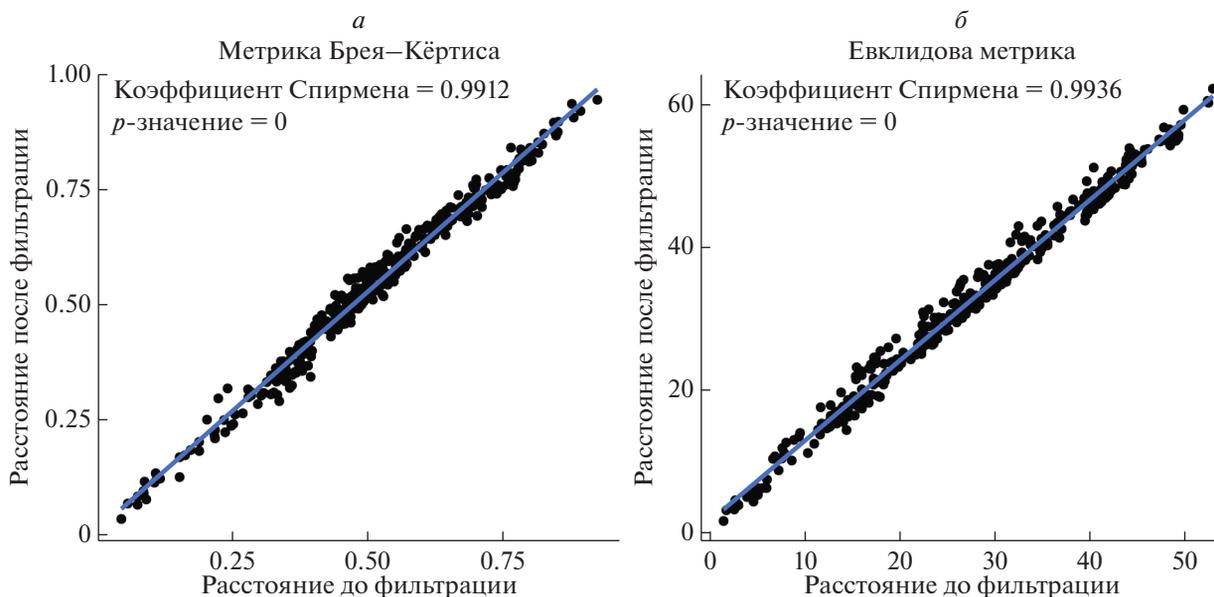


Рис. 4. Корреляционный анализ между матрицами расстояний анализируемых образцов до и после фильтрации. *а* – сравнение матриц по метрике Брея–Кёртиса, *б* – сравнение матриц расстояний по Евклидовым расстояниям.

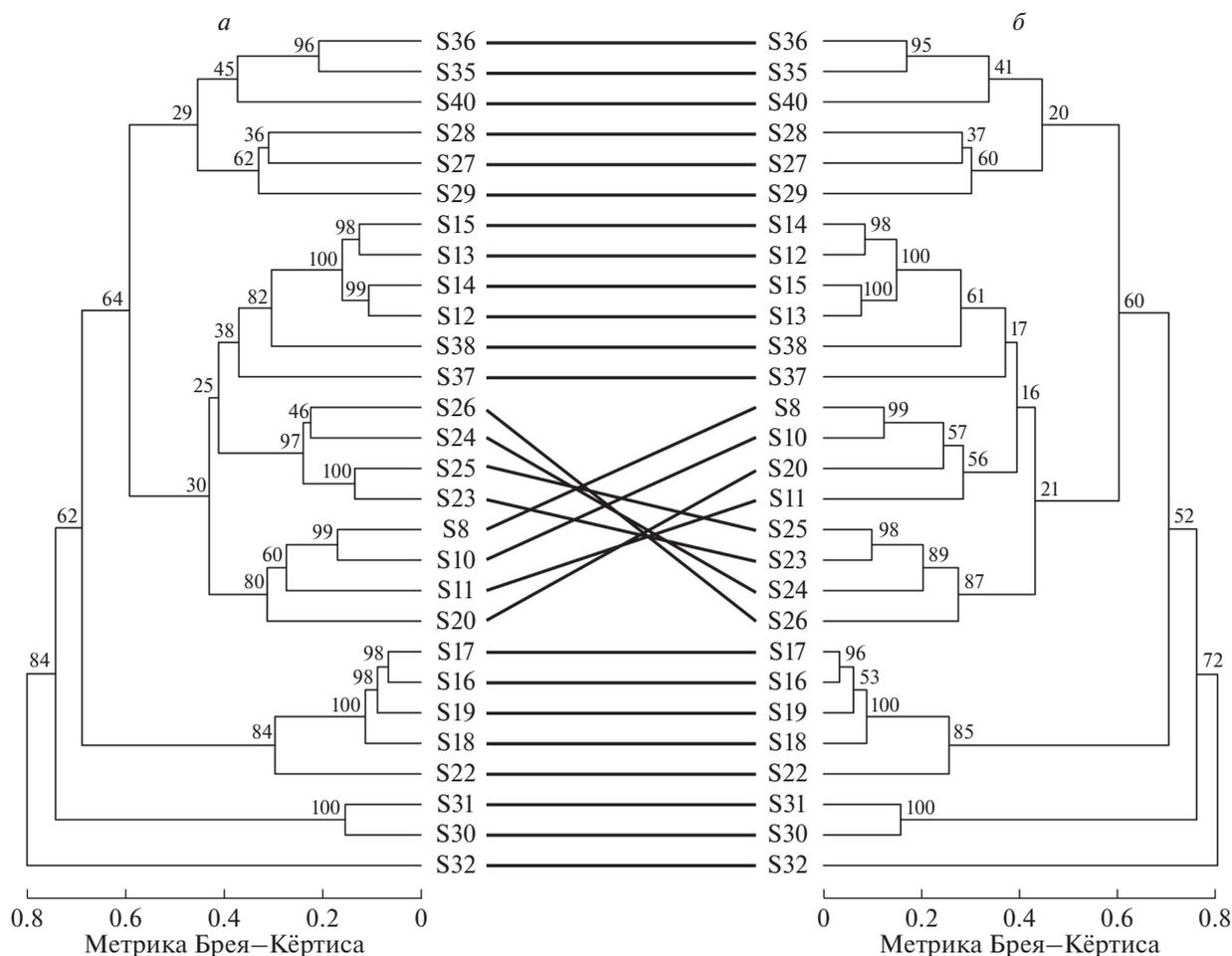
тология дендрограм в обоих случаях сохранялась для большей части образцов. Наблюдалось изменение топологии некоторых ветвей, которые изначально имели меньшие значения бутстреп-поддержек, однако внутри узла топология сохраняла свою структуру. Такое распределение показывает, что вклад, вносимый отфильтрованной частью данных, был незначителен.

При сравнении проб с различными индексами  $\alpha$ -разнообразия была обнаружена интересная зависимость между величиной индекса Шеннона и количеством достоверно определенных ОТЕ в образцах. На рис. 6 показаны графики, отражающие тенденцию к уменьшению достоверно определенных ОТЕ с уменьшением индекса Шеннона. Для примера мы выделили пробы с максимальным, средним и минимальным индексом Шеннона соответственно. В каждой из этих проб был выбран пул доминирующих ОТЕ в количестве 50 штук. Следует отметить, что с падением показателей индекса Шеннона в пробе также увеличивалось количество ОТЕ, определенных с ошибкой выше порога в 0.2. Такое быстрое увеличение значения ошибки могло быть вызвано недостаточной глубиной прочтения. Согласно графикам (рис. 6, *а–в*), при высоком значении индекса Шеннона доля достоверно определенных ОТЕ была гораздо выше, чем при низком значении, что позволило нам использовать в анализе гораздо большую выборку ОТЕ из этого набора.

## ОБСУЖДЕНИЕ

Предложенный в работе алгоритм позволяет решать несколько важных задач, связанных с анализом данных исследований методом ДНК-штрихкодирования. Информация об уровне ошибок определения показателей представленности видов в сообществах, анализируемых с помощью ДНК-штрихкодирования, требуется при проведении мониторинговых и скрининговых работ [35, 36]. В таких исследованиях часто ставится задача определить, меняется ли представленность каких-либо таксонов в сообществе. Сделать выводы о таких изменениях можно только тогда, когда доверительные интервалы показателей представленности вида в двух сравниваемых пробах не пересекаются. Предлагаемый в работе алгоритм помогает решить задачу такого сопоставления, представляя исследователю данные о доверительных интервалах определения показателей представленности всех таксонов в сравниваемых пробах.

Часто для сравнения сообществ, информация о составе которых получена методом ДНК-штрихкодирования, применяют методы многомерной статистики (кластерный анализ на основе различных метрик расстояний, метод главных компонент – РСА, многомерное шкалирование и т.п.) [37]. С помощью таких методов сопоставляют различные пробы на основе показателей представленности входящих в них видов. Разрешающие способности и точности методов многомерной статистики падают, если в массивах информации присутствуют неточные – “зашумленные данные” (данные о видах, показатели представленности которых определены с большими ошибками).



**Рис. 5.** Сравнение UPGMA-дендрограм на основе дистанций Брея–Кёртиса до (а) и после (б) использования предложенного алгоритма. Цифрами на ветвях обозначены значения рассчитанных бутстреп-поддержек.

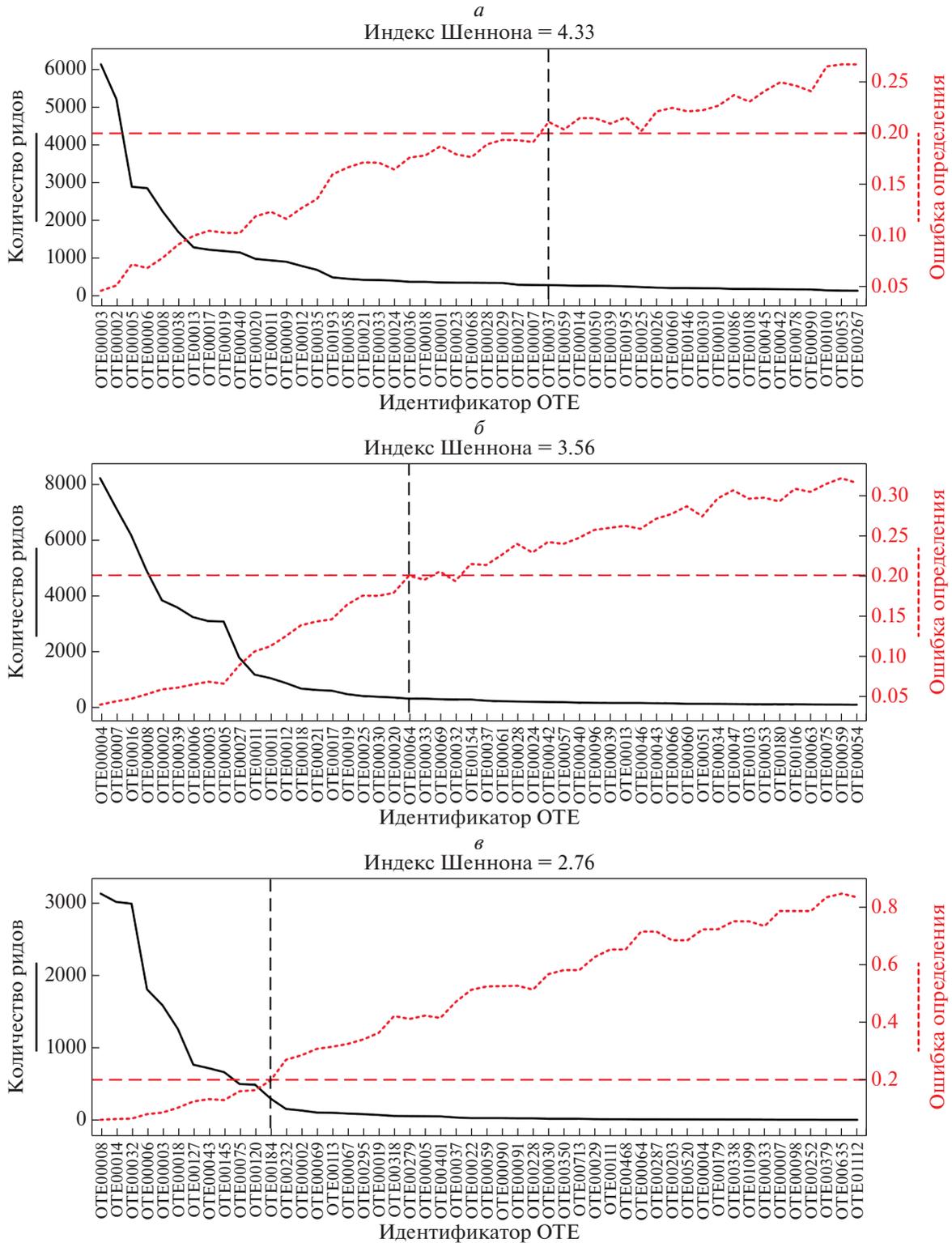
Применение нашего алгоритма позволяет отфильтровать такие данные, сделав массив более пригодным для многомерного статистического анализа.

В работах по ДНК-штрихкодированию различных сообществ [38, 39] для определения глубины прочтения ампликонов часто применяют методы, позволяющие оценить количество недоучтенных (потерянных видов) из-за недостаточно количества расшифрованных последовательностей ДНК. К таким методам относятся расчеты индексов Chao1 [15], ACE [16] и др. Если доля недоучтенных видов по сравнению с оцененным разнообразием сообщества мала, то массив данных пригоден для анализа. Предложенный в работе подход, в отличие от индексов ACE, Chao1, может определить достаточность размера выборки расшифрованных последовательностей ДНК для оценки показателей представленности каждого вида в отдельности. Если для каких-либо видов (таксонов) ошибка оценки показателей представленности высока и не позволяет сделать правиль-

ных выводов, то исследователь может принять решение о дополнительном секвенировании пробы, даже если индексы ACE, Chao1 говорят о достаточности размера выборки.

Таким образом, предложенный нами алгоритм фильтрации наборов данных может существенно облегчить работу специалистов-микробиологов и ученых, занимающихся экологической статистикой и сравнительным анализом различных сообществ. Разработанный подход позволяет отфильтровать так называемый шум из данных, который может влиять на результаты как многомерной статистики, так и при учете индексов видового разнообразия.

Авторы выражают благодарность сотрудникам Иркутского Суперкомпьютерного Центра СО РАН за предоставление доступа к вычислительным ресурсам кластера «Академик В.М. Матросов». Работа выполнена в рамках темы госзадания ЛИН СО РАН № 0279-2021-0015 (разработка алгоритма и анализ данных), а также при финансовой поддержке проектов РФФИ № 18-34-00513 мол\_а



**Рис. 6.** Графики отношения между количеством и представленностью ОТЕ (черная линия) от величины ошибки их определения (красная линия). Проба с наибольшим показателем индекса Шеннона – *а*, проба со средним показателем индекса Шеннона – *б*, проба с наименьшим показателем индекса Шеннона – *в*. ОТЕ ранжированы в порядке убывания представленности в сообществе.

(пробоподготовка и выделение ДНК) и РФФИ № 18-54-05005 Арм\_а (секвенирование).

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. *Zakharenko A.S., Galachyants Y.P., Morozov I.V. et al.* Bacterial communities in areas of oil and methane seeps in pelagic of Lake Baikal // *Microbial Ecology*. 2019. V. 78. № 2. P. 269–285. <https://doi.org/10.1007/s00248-018-1299-5>
2. *Potapov S.A., Tikhonova I.V., Krasnopeev A.Yu. et al.* Metagenomic analysis of viroplankton from the pelagic zone of Lake Baikal // *Viruses*. 2019. I. 11. P. 1–15. <https://doi.org/10.3390/v11110991>
3. *Wood D.E., Salzberg S.L.* Kraken: Ultrafast metagenomic sequence classification using exact alignments // *Genome Biol*. 2014. V. 15. P. 46. <https://doi.org/10.1186/gb-2014-15-3-r46>
4. *Meyer F., Paarmann D., D'Souza M. et al.* The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes // *BMC Bioinformatics*. 2008. V. 9. P. 386. <https://doi.org/10.1186/1471-2105-9-386>
5. *Buchfink B., Xie C., Huson D.* Fast and sensitive protein alignment using DIAMOND // *Nat. Methods*. 2015. V. 12. P. 59–60. <https://doi.org/10.1038/nmeth.3176>
6. *Roux S., Faubladiere M., Mahul A. et al.* Metavir: A web server dedicated to virome analysis // *Bioinformatics*. 2011. V. 27. I. 21. P. 3074–3075. <https://doi.org/10.1093/bioinformatics/btr519>
7. *Kang D.D., Li F., Kirton E. et al.* MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies // *PeerJ*. 2019. V. 7. P. e7359. <https://doi.org/10.7717/peerj.7359>
8. *Wu Y.-W., Simmons B., Singer S.* MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets // *Bioinformatics*. 2016. V. 32. I. 4. P. 605–607. <https://doi.org/10.1093/bioinformatics/btv638>
9. *Edgar R.* UPARSE: Highly accurate OTU sequences from microbial amplicon reads // *Nat. Methods*. 2013. V. 10. P. 996–998. <https://doi.org/10.1038/nmeth.2604>
10. *Callahan B., McMurdie P., Rosen M. et al.* DADA2: High-resolution sample inference from Illumina amplicon data // *Nat. Methods*. 2016. V. 13. P. 581–583. <https://doi.org/10.1038/nmeth.3869>
11. *Schloss P., Westcott S., Ryabin T. et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities // *Applied and Envir. Microbiology*. 2009. V. 75. P. 7537–7541. <https://doi.org/10.1128/AEM.01541-09>
12. *Quast C., Pruesse E., Yilmaz P. et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools // *Nucl. Ac. Res.* 2013. V. 41. I. D1. P. 590–596. <https://doi.org/10.1093/nar/gks1219>
13. *Wang Q., Garrity G., Tiedje J., Cole J.* Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy // *Applied and Envir. Microbiology*. 2007. V. 73. P. 5261–5267. <https://doi.org/10.1128/AEM.00062-07>
14. *DeSantis T., Hugenholtz P., Larsen N. et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB // *Applied and Envir. Microbiology*. 2006. V. 72. P. 5069–5072. <https://doi.org/10.1128/AEM.03006-05>
15. *Chao A.* Non-parametric estimation of the number of classes in a population // *Scand. J. Stat.* 1984. V. 11. № 4. P. 265–270. URL: <http://www.jstor.org/stable/4615964> (дата обращения: 15.09.2021).
16. *Chao A., Lee S.M.* Estimating the number of classes via sample coverage // *J. Amer. Stat. Association*. 1992. V. 87. I. 417. P. 210–217. [http://chao.stat.nthu.edu.tw/wordpress/paper/1992\\_JASA\\_87\\_P210.pdf](http://chao.stat.nthu.edu.tw/wordpress/paper/1992_JASA_87_P210.pdf) (дата обращения: 15.09.2021).
17. *Sorokin D.Y., Berben T., Melton E.D. et al.* Microbial diversity and biogeochemical cycling in soda lakes // *Extremophiles*. 2014. V. 18. P. 791–809. <https://doi.org/10.1007/s00792-014-0670-9>
18. *Feng C., Jia J., Wang C. et al.* Phytoplankton and bacterial community structure in two chinese lakes of different trophic status // *Microorganisms*. 2019. V. 7. P. 621. <https://doi.org/10.3390/microorganisms7120621>
19. *Eiler A., Heinrich F., Bertilsson S.* Coherent dynamics and association networks among lake bacterioplankton taxa // *ISME*. 2012. V. 6. P. 330–342. <https://doi.org/10.1038/ismej.2011.113>
20. *Efron B.* Bootstrap methods: Another look at the jack-knife // *The Annals of Statistics*. 1979. V. 7. P. 1–26.
21. *Pla L.* Bootstrap confidence intervals for the Shannon biodiversity index: A simulation study // *JABES*. 2004. V. 9. I. 42. P. 42–56. <https://doi.org/10.1198/1085711043136>
22. *Рокицкий П.Ф.* Биологическая статистика. Минск: “Вышэйш. шк.”, 1973. 320 с.
23. *Mincheva E.V., Peretolchina T.E., Bukin Yu.S. et al.* The composition of the communities of green algae and microeukaryotes at the sites of Lake Baikal contrasting in anthropogenic pressure // *Limnol. and Freshwater Biol.* 2020. I. 4. P. 729–730. <https://doi.org/10.31951/2658-3518-2020-A-4-729>
24. *Мальник В.В., Штыкова Ю.П., Сутурин А.Н., Тимошкин О.А.* Влияние населенных пунктов на санитарно-микробиологическое состояние малых притоков и прибрежных вод на примере залива Лиственничный (Южный Байкал) // *География и природные ресурсы*. 2019. № 4. С. 84–92. [https://doi.org/10.21782/GIPR0206-1619-2019-4\(84-92\)](https://doi.org/10.21782/GIPR0206-1619-2019-4(84-92))
25. *Timoshkin O.A., Klump J.V., Rudstam L.G. et al.* Recent research on Lake Baikal: The 2018 Conference, freshwater ecosystems-key problems // *J. of Great Lakes*

- Res. 2020. V. 46. I. 1. P. 1–3.  
<https://doi.org/10.1016/j.jglr.2019.12.004>
26. *Maniatis T., Fritsch E.F., Sambrook J.* Molecular Cloning. The Laboratory Manual. N.Y.: Cold Spring Harbor Laboratory, 1982. P. 545.  
[https://doi.org/10.1016/0307-4412\(83\)90068-7](https://doi.org/10.1016/0307-4412(83)90068-7)
  27. *Caporaso J., Lauber C., Walters W. et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and Metagenomic sequencing platform // ISME J. 2012. I. 6. P. 1621–1624.  
<https://doi.org/10.1038/ismej.2012.8>
  28. *Bolger A.M., Lohse M., Usadel B.* Trimmomatic: A flexible trimmer for Illumina sequence data // Bioinformatics. 2014. V. 30. P. 2114–2120.  
<https://doi.org/10.1093/bioinformatics/btu170>
  29. *Martin M.* Cutadapt removes adapter sequences from high-throughput sequencing reads // EMBnet J. 2011. I. 17. P. 10–12.  
<https://doi.org/10.14806/ej.17.1.200>
  30. *Magoč T., Salzberg S.* FLASH: Fast length adjustment of short reads to improve genome assemblies // Bioinformatics. 2011. V. 27. I. 21. P. 2957–2963.  
<https://doi.org/10.1093/bioinformatics/btr507>
  31. *Schloss P.* MiSeq SOP: Mothur. 2015. URL: [http://www.mothur.org/wiki/MiSeq\\_SOP](http://www.mothur.org/wiki/MiSeq_SOP) (дата обращения: 15.09.2021).
  32. *Bray J.R., Curtis J.T.* An ordination of the upland forest communities of southern Wisconsin // Ecol. Monographs. 1957. V. 27. I. 4. P. 325–349.  
<https://doi.org/10.2307/1942268>
  33. *R Core Team.* R: A language and environment for statistical computing. R Foundation for Statistical // Computing. 2019. URL: <http://www.R-project.org/> (дата обращения: 15.09.2021).
  34. *Wickham H.* The Split-Apply-Combine strategy for data analysis // J. Stat. Software. 2011. V. 40. I. 1. P. 1–29.  
<https://doi.org/10.18637/jss.v040.i01>
  35. *Galan M., Razzauti M., Bard E. et al.* 16S rRNA amplicon sequencing for epidemiological surveys of bacteria in wildlife // mSystems. 2016. V. 1. I. 4. e00032–16.  
<https://doi.org/10.1128/mSystems.00032-16>
  36. *Yoshitake K., Kimura G., Sakami T. et al.* Development of a time-series shotgun metagenomics database for monitoring microbial communities at the Pacific coast of Japan // Sci. Rep. 2021. V. 11.  
<https://doi.org/10.1038/s41598-021-91615-3>
  37. *Rudi K., Zimonja M., Trosvik P., Næs T.* Use of multivariate statistics for 16S rRNA gene analysis of microbial communities // Int. J. Food Microbiol. 2007. V. 120. № 2. P. 95–99.  
<https://doi.org/10.1016/j.ijfoodmicro.2007.06.004>
  38. *Sinclair L., Osman O.A., Bertilsson S., Eiler A.* Microbial community composition and diversity via 16S rRNA gene amplicons: evaluating the illumina platform // PLoS One. 2015. V. 10. e0116955.  
<https://doi.org/10.1371/journal.pone.0116955>
  39. *Bukin Y., Galachyants Y., Morozov I. et al.* The effect of 16S rRNA region choice on bacterial community metabarcoding results // Sci. Data. 2019. I. 6.  
<https://doi.org/10.1038/sdata.2019.7>

## Data Filtering Algorithm for DNA Metabarcoding Studies of Environmental Communities

A. Yu. Krasnopeev<sup>a, \*</sup>, Yu. S. Bukin<sup>a, \*\*</sup>, S. A. Potapov<sup>a</sup>, and O. I. Belykh<sup>a</sup>

<sup>a</sup>*Limnological Institute, Siberian Branch, Russian Academy of Science, Irkutsk, 664033 Russia*

*\*e-mail: andrewkrasnopeev@gmail.com*

*\*\*e-mail: bukinyura@mail.ru*

Despite the large sample sizes, the question of the statistical significance of the estimated indicators of taxonomic representativeness remains open. In particular, This mainly concerns the minor component of the communities. To solve the problem of estimating the statistical representativeness of sequenced samples in metabarcoding studies, we developed an algorithm based on the bootstrap method. This algorithm makes can estimate the confidence intervals and absolute and relative errors of the taxonomic representativeness in the studied communities. This study revealed that the proportion of reliable taxa in the communities is several times smaller than their original number. Regression analysis confirmed all obtained results. R code and workflow examples with test files are available at [https://github.com/Krasnopeev/otu\\_error\\_filtering](https://github.com/Krasnopeev/otu_error_filtering).

**Keywords:** metabarcoding, biodiversity, 16S rRNA, bootstrapping, DNA barcoding.

**ФИЛОГЕНЕТИЧЕСКОЕ ПОЛОЖЕНИЕ СИГОВ ГРУППЫ  
*Coregonus lavaretus* (L.) ИЗ ТЕЛЕЦКОГО ОЗЕРА (СИБИРЬ) НА ОСНОВЕ  
ПОЛНОГЕНОМНЫХ ДАННЫХ МИТОХОНДРИАЛЬНОЙ ДНК**

© 2022 г. М. М. Соловьев<sup>1, 2, \*</sup>, Н. А. Бочкарев<sup>1</sup>, Н. В. Орешкова<sup>3, 4</sup>,  
Е. Н. Кашина<sup>1</sup>, Е. П. Симонов<sup>1, 5</sup>

<sup>1</sup>Институт систематики и экологии животных Сибирского отделения  
Российской академии наук, Новосибирск, 630091 Россия

<sup>2</sup>Томский государственный университет, Томск, 634050 Россия

<sup>3</sup>Федеральный исследовательский центр, Красноярский научный центр Сибирского отделения  
Российской академии наук, Красноярск, 660036 Россия

<sup>4</sup>Сибирский федеральный университет, Красноярск, 660041 Россия

<sup>5</sup>Тюменский государственный университет, Тюмень, 625003 Россия

\*e-mail: yarmak85@mail.ru

Поступила в редакцию 10.04.2021 г.

После доработки 12.06.2021 г.

Принята к публикации 25.06.2021 г.

Систематика комплекса видов европейского сига *Coregonus lavaretus* (L.) чрезвычайно сложна. Благодаря широкому распространению от Центральной Европы до водоемов Чукотки и высокой биологической пластичности описаны многие формы с различным таксономическим статусом. В Телецком озере (Россия) обитает эндемичная пара сигов – сиг Правдина *Coregonus lavaretus pravdinellus* и телецкий сиг *C. smitti*. Ранее было отсекуено 106 митогеномов комплекса *C. lavaretus* из разных популяций Европы, в то время как для сигов из Сибири такие данные полностью отсутствуют. Нами получены митохондриальные геномы *C. smitti* и *C. l. pravdinellus* длиной 16738 пн с высоким покрытием. Митогеномы различаются по 46 позициям (0.27%). Совместный анализ митогеномов *C. lavaretus* из водоемов Европы и Телецкого озера подтверждает общее происхождение популяций сига восточной части Балтийского моря и водоемов Сибири. Кроме того, анализ полных митогеномов телецкого сига и сига Правдина не подтверждает их отдельный таксономический статус, по крайней мере, на основе мтДНК.

**Keywords:** Coregoninae, *Coregonus smitti*, *Coregonus lavaretus pravdinellus*, Алтай, митогеном, Сибирь.

**DOI:** 10.31857/S0016675822020126

Сиги рода *Coregonus* (семейство Salmonidae) имеют циркумполярное распространение и в течение длительного времени служат моделью для изучения различных аспектов эволюционной биологии [1–8]. Считается, что систематика комплекса *Coregonus lavaretus* s.l. чрезвычайно сложна, и благодаря высокой биологической пластичности комплексного вида, в нем было описано множество внутривидовых форм [9, 10]. Относительно описанных форм/видов существует точка зрения, согласно которой значительная часть разнообразия принадлежит к одной форме/виду и является синонимами [11–13]. В Южной Сибири представители комплекса распространены от Телецкого озера до притоков р. Лена, где, как и в Арктике, было описано большое число “экологических форм” и видов [10, 11].

Первое упоминание о сигах из Телецкого озера можно найти в статье С.И. Гуляева “О сельдях в Телецком озере”, опубликованной в 1865 г. в газете “Томские губернские ведомости” [14]. Однако в статье нет какого-либо деления сигов на формы или виды и только упоминается, что “похожи на каспийских и принадлежат к породе пресноводных сельдей” [14]. В 1900 г. в статье Н.А. Варпаховского “Рыбы Телецкого озера” опубликованы сведения о *C. lavaretus* из Телецкого озера, где он описан как *C. smitti* [15]. Позже, в 1916 г. Л.С. Берг отнес этого сига к ледовитоморским сигам *C. lavaretus pidschian* (Gmelin) natio Smitti Wapachowski, поскольку он не нашел четких различий между сигом из Телецкого озера и сигами из других водоемов Сибири [16]. Е.П. Радченко в своей работе, опубликованной в 1935 г., также использует название *C. l. pidschian* (Gmelin)

natio Smitti Warpachowski для обозначения телецкого сига – сельди при описании его биологии. В работе не отмечаются различия между формами, хотя приводятся данные по сигам со средней длиной 92.7–170.8 мм (по Смитту) [17]. В 1949 г. Г.Д. Дулькейт публикует статью “Ихтиофауна Телецкого озера и реки Бия”, в которой телецкий сиг упоминается как форма сига-пыжьяна *C. l. pid-schian natio smitti* Warpachowski и дает описание нового вида сига из Телецкого озера – сига Правдина *Coregonus pravdinellus* sp. nov. [18]. В 1962 г. А.Н. Гундризер понижает таксономический статус сига Правдина до подвидового – *Coregonus lavaretus pravdinellus* Dulkeit и приводит сравнительную морфологическую характеристику относительно телецкого сига [19]. В 1981 г. коллективом авторов была опубликована монография “Рыбы Телецкого озера”, где анализируются морфологические, биологические, экологические характеристики сигов и рассматривается их происхождение [20]. Позже были опубликованы еще несколько работ, касающихся морфологической и биологической изменчивости сига Правдина [21–23]. Основываясь на данных аллозимного анализа в 2010 г., С.Н. Балдина с соавт. и Д.В. Политов с соавт. из-за заметных отличий этого сига от *C. lavaretus* из других географических регионов предложили рассматривать телецкого сига как самостоятельный вид *C. smitti* [24, 25]. Однако исследования мтДНК (1930 пн *ND1* [26] и 1140 пн *Cytb* [27]) не выявили значительных отличий как между телецкими сигами, так и между ними и сигами *C. lavaretus* из других водоемов бассейна р. Обь. Также не было обнаружено морфологического и генетического соответствия между различными формами/видами симпатрических сигов Телецкого озера [26]. Такое состояние мтДНК может являться результатом, как формообразования, так и встречного расселения с последующей гибридизацией.

В настоящем исследовании телецкого сига мы будем называть как “*C. smitti*”, а сига Правдина как “*C. l. pravdinellus*”. Телецкий сиг может достигать длины по Смитту 20–25 см (в возрасте семи лет) и общей массы тела около 150–200 г, имеет в среднем (*Sp.br*) 27–28 тычинок на первой жаберной дуге (малотычинковая форма), в то время как сиг Правдина значительно меньше и относится к среднетычинковым сигам (*Sp.br* = 33–34) [28]. Обе формы/вида сигов Телецкого озера относятся к малочешуйчатым сигам (*ll* = 81–83), широко распространенным в водоемах Евразии. Спектр питания телецкого сига таксономически разнообразен и включает в себя личинок насекомых, моллюсков, гаммарид и др., в то время как сиг Правдина питается преимущественно планктонными ракообразными [22]. Подобные формы сигов в литературе часто упоминаются как нормальные (*normal*) (бентофаги) и карликовые (*dwarf*) (планктофаги) [5, 29].

В 2012 г. Якобсен с соавт. анализируя 106 митогеномов представителей комплекса *C. lavaretus* из разных популяций в Европе, достигли гораздо более разрешенной топологии дерева, чем это было возможно ранее на основе небольших фрагментов мтДНК [30]. Однако в этот анализ не были включены представители *C. lavaretus* из Сибири. Чтобы частично восполнить этот пробел, в настоящей работе мы предлагаем полные митогеномы телецких симпатрических сигов *C. smitti* и *C. l. pravdinellus*.

Сиги были пойманы в сентябре 2017 г. в северной части Телецкого озера (52.32° с.ш. 87.50° в.д.) (Республика Алтай, Западная Сибирь, Россия). Митогеномы были собраны из данных, полученных в эксперименте по секвенированию транскриптомов желудочно-кишечного тракта (RNA-seq). Тотальную РНК из кишечника по одной особи от каждого сига выделяли с использованием реагента ExtractRNA для выделения суммарной РНК (ЗАО “Евроген”, Москва) по инструкции производителя. Целостность выделенной РНК проверяли на электрофоретической системе Bioanalyzer 2100 (Agilent Technologies) с использованием Agilent RNA 6000 Nano Reagents (Agilent Technologies). Концентрацию РНК в образце измеряли с помощью флуориметра Qubit (Thermo Fisher Scientific, США). Библиотеку готовили с использованием набора RNA NEBNext Ultra II для Illumina (New England Biolabs) в соответствии с инструкциями производителя. Контроль качества подготовленной библиотеки проводили на приборе Bioanalyzer 2100 с использованием реагентов Agilent DNA 1000 (Agilent Technologies). Готовую библиотеку секвенировали на платформе Illumina NovaSeq 6000 (300 циклов, 2 × 150 пн). Качество полученных прочтений оценивалось с помощью FastQC 0.11.7 [31]. Удаление адаптеров и низкокачественных ридов выполнялись с помощью Cutadapt 1.9.1 [32]. Прочтения были картированы на митохондриальные геномы Coregoninae и собраны в отдельный файл с помощью bowtie2 2.3.5 (с параметрами по умолчанию) [33]. Затем эти чтения были собраны *de novo* в контиги с помощью инструмента “De Novo Assembly” в CLC Genomics Workbench с настройками по умолчанию. Полученные контиги были выровнены на ранее опубликованный митогеном *C. lavaretus* (NC\_002646.1) с помощью инструмента CLC “Map Reads to Reference”. Выровненные на митогеном контиги объединялись в консенсусную последовательность с помощью инструмента CLC “Extract Consensus Sequence”.

В качестве альтернативного подхода прочтения картировались на референсный митогеном с помощью инструмента CLC “Map Reads to Refer-

**Таблица 1.** Оптимальные схемы разделения выравнивания и модели нуклеотидных замен, использованные в филогенетическом анализе

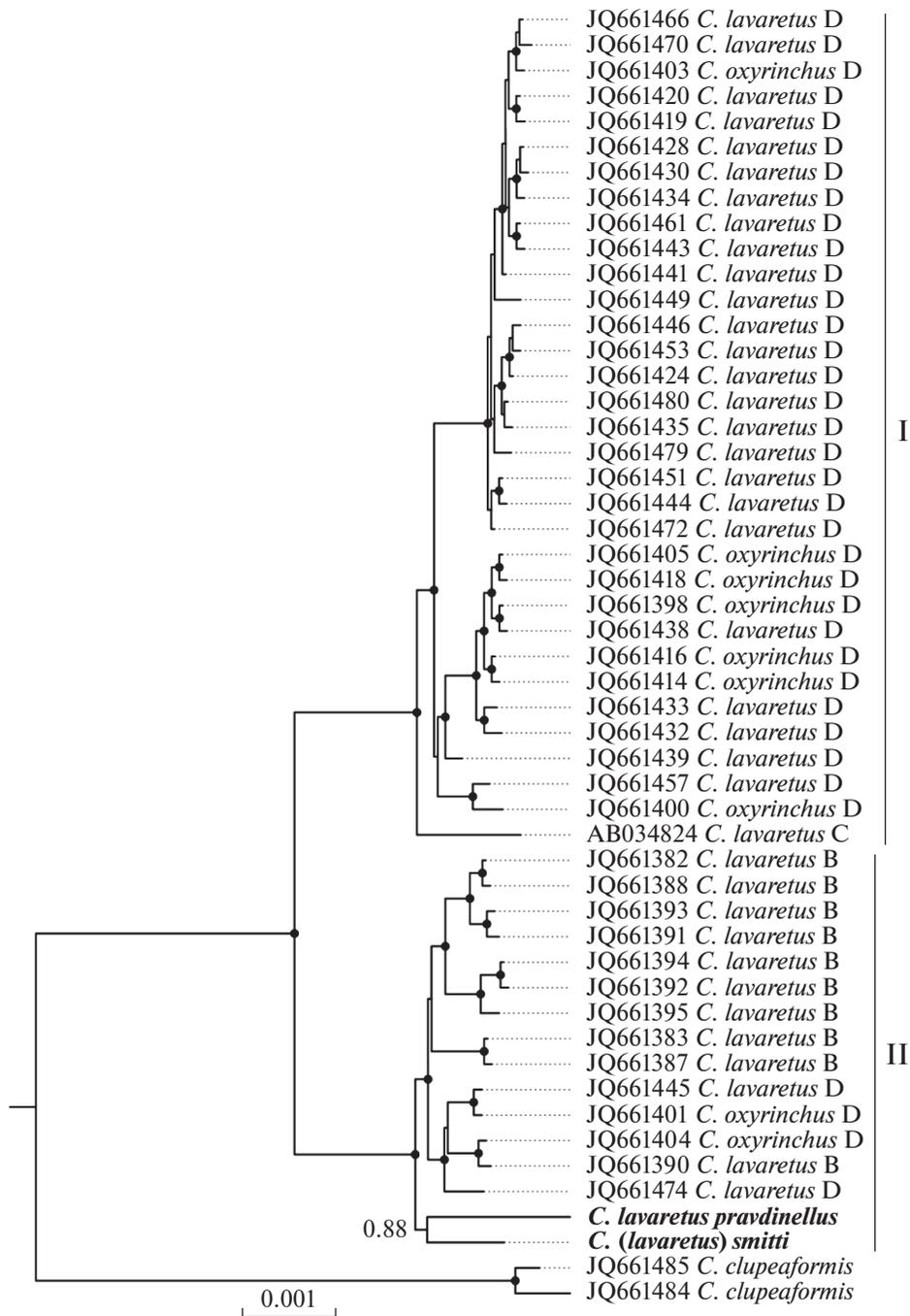
Анализ	Модель	Разделы
BI	HKY + I	tRNA-Ser, tRNA-His, tRNA-Tyr, tRNA-Ala, tRNA-Pro, tRNA-Glu, tRNA-Met, L-strand rep. origin, ATP8, tRNA-Leu, COX2, COX3, COX1, tRNA-Phe, control region, tRNA-Gln
	HKY	tRNA-Lys, tRNA-Asp, tRNA-Val, tRNA-Trp, 12S, 16S, tRNA-Ser, tRNA-Thr, tRNA-Arg, tRNA-Cys
	GTR	tRNA-Gly, tRNA-Asn, ND4l, ND1, ND3, CYTB, ATP6, ND2, ND4, ND5, tRNA-Ile, tRNA-Leu
	HKY	ND6
ML	TN + F	tRNA-Phe, 12S, tRNA-Val, 16S, tRNA-Trp, L-strand rep. origin, ND6, tRNA-Thr
	TN + F + I	tRNA-Leu, ND1, tRNA-Ile, tRNA-Gln, tRNA-Met, ND2, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, COX1, tRNA-Ser2, tRNA-Asp, COX2, tRNA-Lys, ATP8, ATP6, COX3, tRNA-Gly, ND3, tRNA-Arg, ND4l, ND4, tRNA-His, tRNA-Ser, tRNA-Leu2, ND5, tRNA-Glu, CYTB, tRNA-Pro, control region

ence” (настройки по умолчанию) с извлечением консенсусной последовательности. Затем обе консенсусные последовательности мтДНК, полученные двумя разными подходами, вручную выравнивались и сопоставлялись друг с другом в SeaView 4.5.4 [34]. Полученное выравнивание проверялось вручную, и несколько конфликтующих позиций были исправлены путем сопоставления с результатами обратного картирования ридов на полученные консенсусные последовательности. Автоматическая аннотация митохондриальных геномов выполнена с помощью MitoAnnotator [35], проверена и скорректирована вручную. Нами успешно получены митохондриальные геномы *C. smitti* и *C. l. pravdinellus* длиной 16738 пн со средним покрытием 9301× и 13187× соответственно (номера доступа в NCBI GenBank MK913369 и MW817086).

Реконструкция филогенетических отношений выполнялась при помощи методов максимального правдоподобия (maximum likelihood, ML) и байесовского анализа (Bayesian inference, BI). Митохондриальные геномы *C. lavaretus* были получены из публичных данных в GenBank (номера доступа даны на рис. 1), множественное выравнивание последовательностей выполнено с помощью Muscle 3.8.31 [36]. Выравнивание было разделено на отдельные партии согласно схеме аннотации мтДНК, поиск оптимальной схемы разделения и выбор моделей нуклеотидных замен проведен в PartitionFinder 2.1.1 [37] для BI, или в IQ-TREE 2.1.2 [38] с использованием ModelFinder [39] для ML-анализа (табл. 1). Программа

MrBayes 3.2.6 [40] использовалась для байесовского анализа со следующими параметрами: два одновременных запуска по  $10^6$  генераций с частотой записи параметров каждые 500 генераций. Первые 25% генераций были исключены как тренировочные (burn-in). ML-анализ выполнен в программе IQ-TREE, с оценкой устойчивости узлов филогенетических деревьев при помощи процедуры ультрабыстрого бутстрепа [41] с 1000 реплик. Некорректированные генетические дистанции между проанализированными митогеномами были рассчитаны в MEGA7 [42].

Полученные митогеномы различаются по 46 позициям (0.27%). Филогенетические деревья, реконструированные двумя методами, имеют одинаковую топологию и помещают обоих исследованных сигов в “европейскую северо-восточную” кладу *C. lavaretus* (согласно [30]) со 100%-ной поддержкой (на рис. 1 приведены только результаты BI). Среднее нескорректированное генетическое расстояние (*p*-distance) между членами клады составляет  $0.14 \pm 0.01\%$  (min–max: 0.006–0.29%), в то время как среднее расстояние до членов сестринской “Европейской юго-западной” клады (по [30]) составляет  $0.48 \pm 0.05\%$ . Таким образом, совместный анализ митогеномов *C. lavaretus* из водоемов Европы и Телецкого озера подтверждает общее происхождение популяций сига восточной части Балтийского моря и водоемов Сибири. Кроме того, анализ полных митогеномов не подтверждает их отдельный таксономический статус (как “*C. smitti*” и “*C. l. pravdinellus*”), по крайней мере на основе мтДНК.



**Рис. 1.** Филогенетическое дерево комплекса *C. lavaretus*, полученное при помощи байесовского филогенетического анализа на основе полных последовательностей мтДНК. Приведены апостериорные вероятности выше 0.7, черные кружки в узлах означают апостериорную вероятность больше или равную 0.98. Последовательности *C. clupeaformis* использованы в качестве внешней группы. Прочие обозначения: I – Европейская юго-западная клада, II – Европейская северо-восточная клада (по [30]), D – Дания, B – Балтийское море, C – Чехия.

Филогенетический анализ показал базальное положение сигов Телецкого озера в пределах их клады, то есть как наиболее раннюю ветвь, отделившуюся от самого недавнего общего предка “европейской северо-восточной” клады. В предыдущих работах с сибирскими популяциями сигов

показано, что сиги Телецкого озера имеют общие гаплотипы мтДНК по фрагменту ND1 (975 пн) и 16S-ND1 (1929 пн) с сигами из притоков нижнего течения р. Обь (Западная Сибирь), бассейнов рек Большой Енисей и Анабар, водоемов п-ова Таймыр (Восточная Сибирь) и водоемов Фенноскан-

дии [13, 26]. Что свидетельствует о широком распространении и расселении этой гаплогруппы сигов в водоемах Евразии.

Заселение сигами постледниковых водоемов Европы из водоемов Сибири скорее всего происходило двумя путями. В настоящее время известно, что сиги из рефугиумов Алтае-Саянской горной страны расселялись вниз по течению основной магистрали р. Обь, и мигрировали как на восток (в бассейны рек Оленек, Анабар, Пра-Енисей), так и в западном направлении, до бассейна Балтийского моря [13, 43]. Кроме того, морфологически и генетически дистантные ледниково-равнинные сиги (*C. l. pidschian* n. *glacialis*), вероятно арктического происхождения, мигрировали вдоль побережья в западном направлении [1]. Таким образом, разнообразие сигов с нижним и полунижним ртом в европейских водоемах обеспечивалось гибридизацией, по крайней мере четырех генетических линий сигов: сформировавшихся в европейских рефугиумах Альпийской и Североморской, и пришедшими с Востока Южно-Сибирской и Сибирской Арктической [1, 2].

Исследование поддержано грантом РНФ 17-74-10071.

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

Вклад авторов: М.М.С. и Е.П.С. внесли равный вклад в работу и подготовку рукописи.

## СПИСОК ЛИТЕРАТУРЫ

1. Østbye K., Bernatchez L., Næsje T.F. et al. Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers // *Mol. Ecol.* 2005. V. 14. P. 4371–4387. <https://doi.org/10.1111/j.1365-294X.2005.02737.x>
2. Østbye K., Amundsen P., Bernatchez L. et al. Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times // *Mol. Ecol.* 2006. V. 15. № 13. P. 3983–4001. <https://doi.org/10.1111/j.1365-294X.2006.03062.x>
3. Rogers S.M., Bernatchez L. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs // *Mol. Biol. Evol.* 2007. V. 24. № 6. P. 1423–1438. <https://doi.org/10.1093/molbev/msm066>
4. Lundsgaard-Hansen B., Matthews B., Vonlanthen P. et al. Adaptive plasticity and genetic divergence in feeding efficiency during parallel adaptive radiation of whitefish (*Coregonus* spp.) // *J. Evol. Biol.* 2013. V. 26. № 3. P. 483–498. <https://doi.org/10.1111/jeb.12063>
5. Sevellec M., Derome N., Bernatchez L. Holobionts and ecological speciation: the intestinal microbiota of lake whitefish species pairs // *Microbiome*. 2018. V. 6. № 47. <https://doi.org/10.1186/s40168-018-0427-2>
6. Solovyev M.M., Kashinskaya E.N., Bochkarev N.A. et al. The effect of diet on the structure of gut bacterial community of sympatric pair of whitefishes (*Coregonus lavaretus*): One story more // *PeerJ*. 2019. V. 7. P. e8005. <https://doi.org/10.7717/peerj.8005>
7. Vychenko O.S., Sukhanova L.V., Azhikina T.L. et al. Differences in brain transcriptomes of closely related Baikal Coregonid species // *Bio. Med. Res. Int.* 2014. V. 2014. P. 1–11. <https://doi.org/10.1155/2014/857329>
8. Teterina V., Sukhanova L., Smirnov V. et al. Complete mitochondrial genomes of Baikal endemic coregonids: Omul and lacustrine whitefish (Salmonidae: *Coregonus* sp.) // *Mitochondrial DNA Part B-Resources*. 2020. V. 5. P. 414–416. <https://doi.org/10.1080/23802359.2019.1703565>
9. Правдин И.Ф. Сиги водоемов Карело-Финской ССР. М., Л.: Изд-во АН СССР, 1954. 324 с.
10. Решетников Ю.С. Экология и систематика сиговых рыб. М.: Наука, 1980. 300 с.
11. Скрябин А.Г. Сиговые рыбы юга Сибири. Новосибирск: Наука, 1979. 229 с.
12. Романов В.И., Зуйкова Е.И., Бочкарев Н.А. Морфологическая и экологическая дифференциация симпатрических сигов рода *Coregonus* из оз. Таймыр // *Сиб. экологич. журн.* 2016. № 9. С. 290–296.
13. Бочкарёв Н.А., Пестрякова Л.А., Захаров Е.С. и др. Сиг-пыжьян (*Coregonus lavaretus pidschian*, Coregonidae) реки Анабар: морфогенетическая структура популяции // *Генетика*. 2018. Т. 54. № 9. С. 1057–1067. <https://doi.org/10.1134/S1022795418090041>
14. Гуляев С.И. О сельдях в Телецком озере // *Томские губернские ведомости (часть неофициальная)*. 1865. № 43. С. 1–4.
15. Варнаховский Н.А. Рыбы Телецкого озера // *Ежегодник зоол. музея АН*. 1900. Т. 5. С. 241–271.
16. Берг Л.С. Рыбы пресных вод Российской Империи. М.: Типография Т-ва Рябушинских, 1916. 563 с.
17. Радченко Е.П. О рыбном хозяйстве Телецкого озера (промыслово-биологический очерк) // *Тр. ВНИРО*. 1935. Т. 2. С. 64–100.
18. Дулькейт Г.Д. Ихтиофауна озера Телецкого и реки Бия // *Заметки по фауне и флоре Сибири*. Томск: Томский гос. ун-тет, 1949. № 8. С. 9–12.
19. Гундризер А.Н. К биологии сига Правдина из Телецкого озера и реки Бия // *Изв. Сиб. отд. АН СССР*. 1962. № 3. С. 111–119.
20. Гундризер А.Н., Иоганзен Б.Г., Кафанова В.В., Кривошеиков Г.М. Рыбы Телецкого озера. Новосибирск: Наука, 1981. 189 с.
21. Бочкарёв Н.А., Гафина Т.Э. Сравнительная характеристика телецкого сига и сига Правдина Телец-

- кого озера. (Алтайский край) // Сиб. биол. журн. 1993. № 2. С. 64–69.
22. Бочкарёв Н.А., Зуйкова Е.И. Морфобиологическая и экологическая дифференциация симпатрических сигов рода *Coregonus* из Телецкого озера // Зоол. журн. 2006. Т. 85. № 8. С. 950–958.
  23. Голубцов А.С., Малков Н.П. Очерк ихтиофауны Республики Алтай: систематическое разнообразие, распространение и охрана. М.: Тов-во науч. изд. КМК, 2007. 164 с.
  24. Балдина С.Н., Гордон Н.Ю., Политов Д.В. Генетическая дифференциация некоторых видов сиговых рыб Сибири // Биология, биотехника разведения и состояние запасов сиговых рыб. Матер. VII Международного научно-произв. совещ. / Под ред. Литвиненко А.И., Решетникова Ю.С. Тюмень: Госрыбцентр, 2010. С. 5–9.
  25. Политов Д.В., Гордон Н.Ю., Балдина С.Н. Филогенетика и филогеография сиговых – сравнение морфологического и генетического подходов // Биология, биотехника разведения и состояние запасов сиговых рыб. Матер. VII Международного научно-произв. совещ. / Под ред. Литвиненко А.И., Решетникова Ю.С. Тюмень: Госрыбцентр, 2010. С. 37–41.
  26. Bochkaev N.A., Zuykova E.I., Katokhin A.V. Morphology and mitochondrial DNA variation of the Siberian whitefish *Coregonus lavaretus pidschian* (Gmelin) in the upstream water bodies of the Ob and Yenisei rivers // *Evol. Ecol.* 2011. V. 25. P. 557–572. <https://doi.org/10.1007/s10682-010-9437-7>
  27. Sukhanova L.V., Smirnov V.V., Smirnova-Zalumi N.S. et al. Molecular phylogeography of Lake Baikal coregonid fishes // *Advanc. Limnol.* 2012. V. 63. P. 261–283. <https://doi.org/10.1127/advlim/63/2012/261>
  28. Бочкарёв Н.А. Межгодовая изменчивость биологических признаков у симпатрических сигов Телецкого озера // Вест. СВНЦ ДВО РАН. 2009. № 3. С. 40–45.
  29. Bernatchez L. Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (Salmoniformes) // *Evolution Illuminated: Salmon and Their Relatives*. N.Y.: Oxford Univ. Press, 2004. P. 175–207.
  30. Jacobsen M.M., Hansen M.M., Orlando L. et al. Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and ecologically distinct European whitefish (*Coregonus* spp.) // *Mol. Ecol.* 2012. V. 21. P. 2727–2742. <https://doi.org/10.1111/j.1365-294X.2012.05561.x>
  31. Andrews S. FastQC: A quality control tool for high throughput sequence data. [accessed 2019 March 10]. 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
  32. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads // *EMBnet.journal*. 2011. V. 17. № 1. P. 10–12. <https://doi.org/10.14806/ej.17.1.200>
  33. Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2 // *Nat. Methods*. 2012. V. 185. № 9. P. 357–354. <https://doi.org/10.1038/nmeth.1923>
  34. Gouy M., Guindon S., Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building // *Mol. Biol. Evol.* 2010. V. 27. P. 221–224. <https://doi.org/10.1093/molbev/msp259>
  35. Iwasaki W., Fukunaga T., Isagozawa R. et al. MitoFish and MitoAnnotator: A mitochondrial genome database of fish with an accurate and automatic annotation pipeline // *Mol. Biol. Evol.* 2013. V. 30. № 11. P. 2531–2540. <https://doi.org/10.1093/molbev/mst141>
  36. Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput // *Nucl. Ac. Res.* 2004. V. 32. P. 1792–1797. <https://doi.org/10.1093/nar/gkh340>
  37. Lanfear R., Frandsen P.B., Wright A.M. et al. Partition-Finder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses // *Mol. Biol. Evol.* 2016. V. 34. № 3. P. 772–773. <https://doi.org/10.1093/molbev/msw260>
  38. Minh B.Q., Schmidt H.A., Chernomor O. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era // *Mol. Biol. Evol.* 2020. V. 37. P. 1530–1534. <https://doi.org/10.1093/molbev/msaa015>
  39. Kalyaanamoorthy S., Minh B.Q., Wong T.K.F. et al. ModelFinder: Fast model selection for accurate phylogenetic estimates // *Nat. Methods*. 2017. V. 14. P. 587–589. <https://doi.org/10.1038/nmeth.4285>
  40. Ronquist F., Huelsenbeck J.P. MRBAYES 3: Bayesian phylogenetic inference under mixed models // *Bioinformatics*. 2003. V. 19. P. 1572–1574. <https://doi.org/10.1093/bioinformatics/btg180>
  41. Hoang D.T., Chernomor O., von Haeseler A. et al. UFBoot2: Improving the ultrafast bootstrap approximation // *Mol. Biol. Evol.* 2018. V. 35. P. 518–522. <https://doi.org/10.1093/molbev/msx281>
  42. Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets // *Mol. Biol. Evol.* 2016. V. 33. P. 1870–1874. <https://doi.org/10.1093/molbev/msw054>
  43. Бочкарёв Н.А., Сендек Д.С., Зуйкова Е.И. и др. Популяционная структура и происхождение некоторых экологических форм *Coregonus lavaretus pidschian* из р. Оленек // *Генетика*. 2021. Т. 57. № 7. С. 797–809. <https://doi.org/10.31857/S0016675821070043>

**Phylogenetic Position of Whitefish *Coregonus lavaretus* (L.)  
from Teletskoye Lake (Siberia) Based on Complete Mitochondrial DNA**

**M. M. Solovyev<sup>a, b, \*</sup>, N. A. Bochkarev<sup>a</sup>, N. V. Oreshkova<sup>c, d</sup>, E. N. Kashinskaya<sup>a</sup>, and E. P. Simonov<sup>a, e</sup>**

<sup>a</sup>*Institute of Systematics and Ecology of Animals of the Siberian Branch  
of the Russian Academy of Sciences, Novosibirsk, 630091 Russia*

<sup>b</sup>*Tomsk State University, Tomsk, 634050 Russia*

<sup>c</sup>*Federal Research Center “Krasnoyarsk Science Center” of the Siberian Branch of the Russian Academy of Sciences,  
Krasnoyarsk, 660036 Russia*

<sup>d</sup>*Siberian Federal University, Krasnoyarsk, 660041 Russia*

<sup>e</sup>*University of Tyumen, Tyumen, 625003 Russia*

*\*e-mail: yarmak85@mail.ru*

Taxonomy of the European whitefish *Coregonus lavaretus* (L.) species complex is known to be extremely intricate. Due to circumpolar distribution and high biological plasticity, many sympatric and allopatric forms/species of whitefishes have been described. Previously, 106 mitogenomes of *C. lavaretus* complex from different populations of the Baltic and North seas were sequenced but it not included any whitefish samples from Siberia. A sympatric pair of endemic whitefishes (*C. smitti* and *C. l. pravdinellus*) inhabit Teletskoye Lake. In the present article we have obtained the mitochondrial genomes of *C. smitti* and *C. l. pravdinellus* 16738 bp length with a high coverage. The mitogenomes were different by 46 positions (0.27%). Comparative analysis of *C. lavaretus* mitogenomes from water bodies of Europe and Teletskoye Lake confirms the common origin of the populations of the eastern part of the Baltic Sea and water bodies of Siberia. In addition, the analysis of the complete mitogenomes of the *C. smitti* and *C. l. pravdinellus* does not confirm their separate taxonomic status, at least based on mtDNA.

**Keywords:** Coregoninae, *Coregonus smitti*, *Coregonus lavaretus pravdinellus*, Altai Republic, mitogenome, Siberia.

## ПОЛИМОРФИЗМ NBS-LRR ГЕНОВ УСТОЙЧИВОСТИ ГИБРИДОВ САХАРНОЙ СВЕКЛЫ ПО ДАННЫМ NBS-ПРОФАЙЛИНГА

© 2022 г. А. А. Трифонова<sup>1</sup>, \*, Е. Р. Парадня<sup>1</sup>, К. В. Борис<sup>1</sup>, А. М. Кудрявцев<sup>1</sup>

<sup>1</sup>Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

\*e-mail: aichka89@mail.ru

Поступила в редакцию 19.04.2021 г.

После доработки 14.05.2021 г.

Принята к публикации 25.05.2021 г.

Проведено исследование 49 современных отечественных и зарубежных гибридов сахарной свеклы методом NBS-профайлинга. Уровень полиморфизма NBS-LRR генов устойчивости гибридов сахарной свеклы был достаточно высоким и составил 87.77% (74.21% для 15 отечественных гибридов и 83.80% для 34 зарубежных гибридов), для 11 гибридов были выявлены уникальные NBS-фрагменты. Кластерный анализ полученных данных выявил дифференциацию отечественных и зарубежных гибридов сахарной свеклы, что свидетельствует о различии в их наборах генов устойчивости. Кроме того была выявлена кластеризация ряда зарубежных гибридов в зависимости от селекционной компании.

**Ключевые слова:** *Beta vulgaris*, генетическое разнообразие, R-гены.

**DOI:** 10.31857/S0016675822010118

Сахарная свекла (*Beta vulgaris* L. subsp. *vulgaris*) – основная техническая культура, выращиваемая для производства сахара в регионах с умеренным климатом, в том числе в России. Одним из главных факторов, влияющих на производство сахарной свеклы, является пораженность посевов различными болезнями и вредителями, которые могут вызывать существенные потери урожая. В этой связи большое значение приобретает создание гибридов сахарной свеклы, устойчивых к фитопатогенам. В настоящее время до 80% открытых селекционных программ по созданию гибридов сахарной свеклы направленно на формирование устойчивости к различным биотическим и абиотическим стрессам [1]. При этом важную роль в селекционном процессе играют современные молекулярно-генетические методы и подходы, позволяющие значительно ускорить процесс создания гибридов и сократить затраты [2].

Устойчивость растений ко многим патогенам обуславливается семейством генов резистентности (гены устойчивости, Resistance genes, **R-гены**). В геноме сахарной свеклы было идентифицировано 715 аналогов R-генов (RGA – Resistance Gene Analogs) [3]. Одним из наиболее распространенных классов R-генов у *B. vulgaris* являются **NBS-LRR** (Nucleotide Binding Site, Leucine-Rich Repeats) гены, кодирующие соответствующие консервативные домены [3, 4]. Показано, что именно NBS-LRR гены обуславливают устойчивость к таким рас-

пространным патогенам сахарной свеклы как ризомания (Beet Necrotic Yellow Vein Virus) и цистовая нематода (*Heterodera schachtii* Schm.) [5–7].

Для оценки варибельности R-генов растений, кодирующих NBS-домен, был разработан метод NBS-профайлинга, основанный на амплификации фрагментов R-генов с помощью выжженных праймеров, комплементарных консервативным участкам NBS-домена (P-петля, киназа 2, GLPL) [8]. Данный метод является универсальным и был успешно применен для исследования генов устойчивости широкого спектра сельскохозяйственных культур [9–12]. Однако ранее NBS-профайлинг не применялся для изучения NBS-LRR генов устойчивости сахарной свеклы и других видов рода *Beta*.

Целью данной работы является оценка варибельности NBS-LRR генов устойчивости современных коммерческих гибридов сахарной свеклы отечественной и зарубежной селекции методом NBS-профайлинга.

Для проведения исследования были отобраны 15 отечественных гибридов сахарной свеклы: Рамонский МС 46, Рамонский МС 60, Рамоза, РМС 120, РМС 121, РМС 127, РМС 134, РО 117 (ВНИИСС), Гибриды 1146, 1182, 1132, 1177, 1143 (СоюзСемСвекла); Каскад (Львовские семена); Рубин (Первомайская СОС); и 34 зарубежных гибрида сахарной свеклы: Андромеда КВС, Баронесса КВС, Дубравка КВС, Маришка КВС, Оле-

**Таблица 1.** Показатели генетического разнообразия гибридов сахарной свеклы по данным NBS-профайлинга

Анализируемая выборка	Общее число фрагментов	Число полиморфных фрагментов	Число уникальных фрагментов	Уровень полиморфизма, %	Ожидаемая гетерозиготность, $H_e$	Диапазон значений коэф-та Дайса
Общая выборка, NBS2/ <i>Mse</i> I	196	167	10	85.2	—	—
Общая выборка, NBS7/ <i>Mse</i> I	172	156	4	90.7	—	—
Общая выборка, NBS2/ <i>Mse</i> I и NBS7/ <i>Mse</i> I	368	323	14	87.77	0.265	0.71–0.91
Отечественные гибриды, NBS2/ <i>Mse</i> I и NBS7/ <i>Mse</i> I	318	236	3	74.21	0.260	0.73–0.88
Зарубежные гибриды, NBS2/ <i>Mse</i> I и NBS7/ <i>Mse</i> I	358	300	11	83.80	0.260	0.72–0.91

сия КВС, Светлана КВС (KWS); BTC 320, BTC 590, BTC 845, BTC 980 (BETASEED); Баккара, Милорд, Оти, Урал (FLORIMOND); Кариока, Митика, Мишель, Портланд, Шаннон (LION SEEDS); Ангус, Бугги, Вентура, Гамильтон, Харлей (MARIBO); Волга, Неро, Риттер (DLF SEEDS); Каньон, Койот, Шайенн (SESVANDERHAVE); Борислав, Гулливер, Логан, Пушкин (STRUBE), предоставленных ООО «СоюзСемСвекла» в рамках подпрограммы «Развитие селекции и семеноводства сахарной свеклы в Российской Федерации». 42 из 49 образцов изученной выборки включены в Государственный реестр селекционных достижений, допущенных к использованию на 2020 г.

ДНК выделяли из свежих проростков (по одному проростку для каждого гибрида) с использованием набора ZR Plant/Seed DNA Mini Prep (Zymo Research) согласно инструкции производителя.

NBS-профайлинг проводили по стандартной методике [8], с небольшими модификациями. Для рестрикции 200 нг геномной ДНК использовали эндонуклеазу *Mse*I (NEB). Полимеразную цепную реакцию проводили с использованием набора реактивов производства «Диалат-ЛТД» и двух праймеров (NBS2 5'-GTWGTYYTTCYRA-ICCISSCAT-3' и NBS7 5'-ATTGTTGGRATGGG-MGGIMTIGG-3') [8], в термоциклере T100™ (Bio-Rad). Продукты реакции амплификации разделяли электрофорезом в денатурирующем

6%-ном полиакриламидном геле в 1× TBE буфере в камере Sequi-Gen® GT Sequencing Cell (Bio-Rad), затем гели окрашивали нитратом серебра [13].

Наличие/отсутствие продуктов амплификации отмечали визуально, учитывали только четко воспроизводимые фрагменты, результаты заносили в бинарную матрицу (1/0) в программе Microsoft Excel. Расчет коэффициента генетического сходства Дайса между парами образцов проводили с использованием программы PAST 3.16 [14]. Значения ожидаемой гетерозиготности ( $H_e$ ) рассчитывали с помощью программы GenAlEx6.41 [15]. Кластерный анализ проводили методом Neighbor-Joining (NJ) в программе MEGA 7 [16].

В ходе NBS-профайлинга 49 гибридов сахарной свеклы с использованием двух праймеров было выявлено 368 фрагментов, в том числе 323 полиморфных (табл. 1). Для каждого изученного образца сахарной свеклы был получен уникальный NBS-спектр, а для двух отечественных и девяти зарубежных гибридов были идентифицированы уникальные фрагменты. При этом для образцов BTC-320 (BETASEED), Шайенн (SESVANDERHAVE) и Рамоза (ВНИИСС) выявлено по два уникальных фрагмента. Присутствие уникальных фрагментов в NBS-спектрах отдельных образцов позволяет предположить наличие у них специфичного набора генов устойчивости.

Уровень генетического разнообразия изучаемой выборки оказался достаточно высоким и составил 87.77%, а показатель ожидаемой гетерозиготности ( $H_e$ ) – 0.265. При этом для 15 отечественных гибридов уровень полиморфизма составил 74.21%, а для 34 зарубежных – 83.80% (табл. 1). Несмотря на то, что NBS-профайлинг ранее не применялся для изучения представителей рода *Beta*, данный метод был успешно использован для изучения других сельскохозяйственных культур и родственных видов с использованием системы из двух комбинаций праймер/фермент, что позволяет сравнить полученные данные. Так, более низкий уровень варибельности (70.6%) был показан для 66 образцов картофеля отечественной и зарубежной селекции (NBS7/*MseI* и NBS9/*MseI*), а для одного сорта был выявлен уникальный фрагмент [12]. Существенно ниже был уровень полиморфизма NBS-LRR генов устойчивости 53 сортов твердой пшеницы (34%), выявленный с помощью рестриктазы *MseI* и пяти NBS-праймеров [9]. Уровень варибельности генов устойчивости 46 представителей рода *Malus* выявленный с использованием комбинаций праймер/фермент NBS2/*MseI* и NBS5/*MseI* составил 79%, при этом для изученных сортов яблони отечественной и зарубежной селекции данный показатель был ниже и составил 49% [11]. При исследовании межвидового полиморфизма NBS-LRR генов устойчивости 46 образцов рода *Cucumis* (NBS3/*MseI* и NBS5/*MseI*) уровень варибельности составил 89% [10]. Интересно, что уровень внутривидового разнообразия генов устойчивости изученных гибридов сахарной свеклы оказался сопоставим с выявленным в вышеперечисленных работах межвидовым уровнем разнообразия, и был выше, чем внутривидовой у преимущественно вегетативно размножаемых культур, таких как картофель и яблоня.

Таким образом, система NBS-профайлинга, впервые примененная для анализа генетического разнообразия гибридов сахарной свеклы оказалась эффективной несмотря на относительно небольшое число генов устойчивости в геноме сахарной свеклы [3], а уровень их полиморфизма был достаточно высоким. В целом, учитывая узость генетической базы сахарной свеклы, обусловленную ее происхождением и почти полным переходом на ЦМС-гибриды, уровень полиморфизма генома данной культуры, выявленный ранее с использованием различных типов маркеров достаточно велик. Так, SSR-анализ 40 гибридов сахарной свеклы показал значительный уровень полиморфизма не только между гибридами, но и внутри гибридов (до семи аллельных вариантов по одному локусу) [17]. А уровень генетического полиморфизма 43-х российских и иностранных гибридов сахарной свеклы, выявленный с помощью AFLP-анализа, составил 91% (75.8% – для

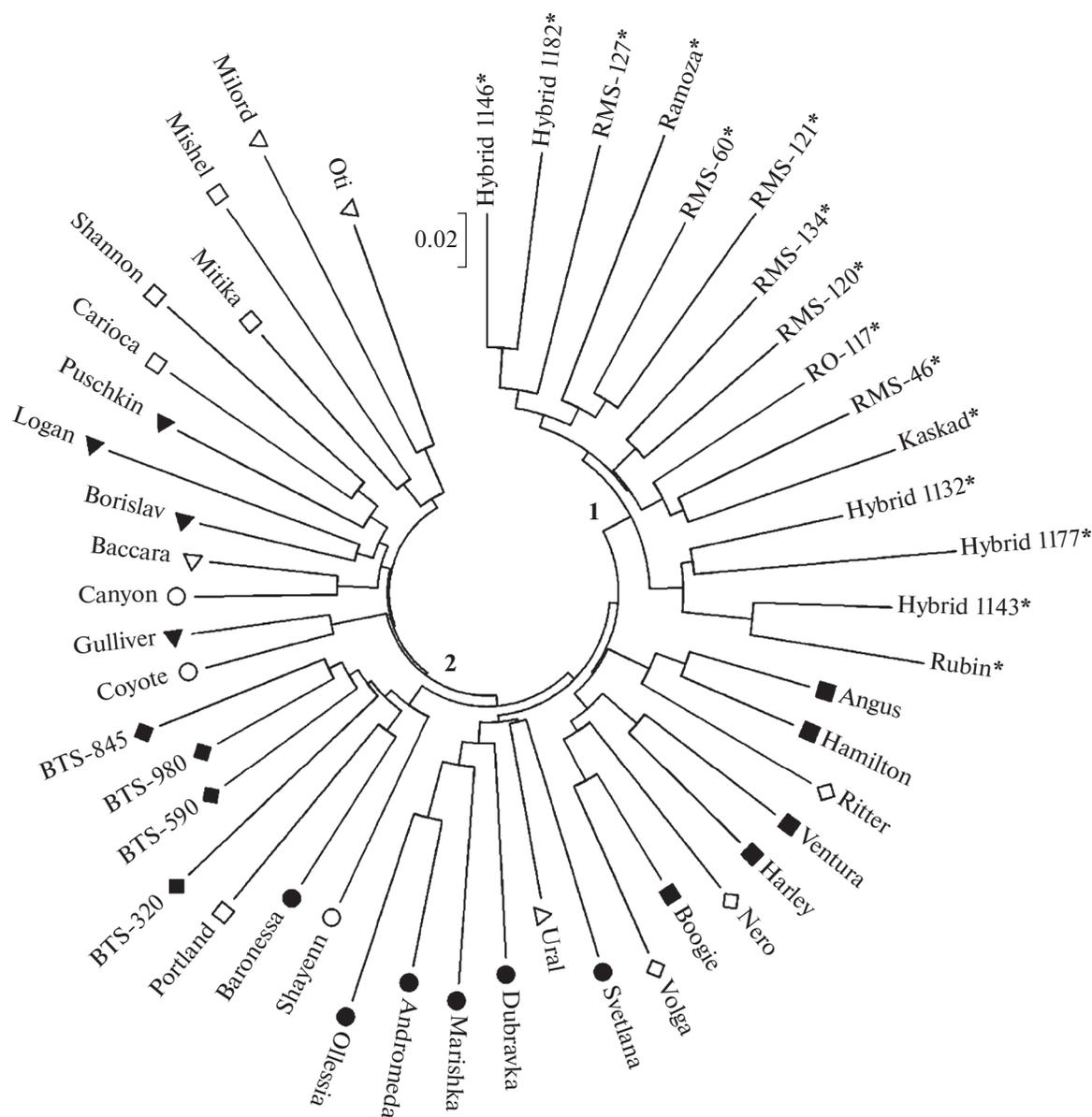
восьми отечественных и 87.6 – для 35 зарубежных гибридов) [18].

Среднее значение коэффициента попарного генетического сходства Дайса для всей изученной выборки составило 0.81, при максимальном значении 0.91 (между образцами Ангус и Гамильтон (MARIBO)), и минимальном – 0.71 (между образцами Мишель (LION SEEDS) и Гибрид 1182 (СоюзСемСвекла)). Минимальный уровень сходства между иностранными гибридами отмечен для образцов Мишель (LION SEEDS) и Олесия (KWS) (0.72), максимальный совпадает с показателем для общей выборки. Отечественные гибриды Рубин (Первомайская СОС) и Гибрид 1143 (СоюзСемСвекла) имели наибольшее сходство, а гибриды РМС-121 (ВНИИСС) и Гибрид 1177 (СоюзСемСвекла) – наименьшее.

На основе данных NBS-профайлинга была построена дендрограмма на которой отечественные и зарубежные гибриды сахарной свеклы формируют отдельные кластеры (рис. 1). Такое разделение может свидетельствовать о существенном различии в наборах генов устойчивости гибридов разного происхождения, которое, по-видимому, связано с использованием различных генетических баз для их создания. Ранее с помощью AFLP-анализа также была выявлена дифференциация зарубежных и отечественных гибридов [18]. При этом для отечественных гибридов, как правило, характерна более высокая адаптивность, устойчивость к корневым и кагатным гнилям и лежкость, что важно из-за особенностей технологии хранения и переработки сахарной свеклы в нашей стране. Однако в настоящее время на высокопродуктивные зарубежные гибриды, обладающие комплексной устойчивости к болезням, приходится более 90% посевов сахарной свеклы в России.

Отдельно была рассмотрена дифференциация зарубежных и отечественных гибридов. Среди проанализированных отечественных гибридов (преимущественно двух производителей) не наблюдается дифференциации по происхождению. Зарубежные гибриды формируют несколько подкластеров, объединяющих гибриды одной селекционной компании. Так, отдельный подкластер формируют восемь гибридов датских компаний MARIBO и DLF SEEDS, вместе группируются пять из шести изученных гибридов компании KWS, а также четыре гибрида компании BETASEED. Остальные зарубежные гибриды формируют смешанные подкластеры (рис. 1).

В исследованиях, проведенных ранее с помощью других методов, также наблюдалась дифференциация гибридов некоторых производителей. Так, из трех групп образцов, выявленных при анализе популяционной структуры 54 гибридов сахарной свеклы пяти американских компаний с



**Рис. 1.** Дендрограмма, построенная по данным NBS-профайлинга 49 образцов сахарной свеклы методом Neighbor-Joining. 1: \* – отечественные гибриды; 2: зарубежные гибриды ● – KWS; ■ – MARIBO; □ – LION SEEDS; △ – FLORIMOND; ○ – SESVANDERHAVE; ◇ – DLF SEEDS; ◆ – BETASEED; ▲ – STRUBE.

помощью SSR, DaT и SNP маркеров, одна группа была сформирована гибридами компании Holly Hybrids, а остальные имели смешанный состав [19]. По данным SSR-анализа 40 зарубежных гибридов сахарной свеклы тринадцати компаний не группировались в зависимости от производителя, кроме девяти гибридов компании KWS [17]. Дифференциация образцов производства компаний BETASEED и KWS была выявлена с помощью AFLP-анализа гибридов восьми различных зарубежных производителей [18].

Выявленная дифференциация гибридов сахарной свеклы может свидетельствовать о том,

что производители используют различные источники и разных доноров устойчивости при создании гибридов. Известно, что в селекции сахарной свеклы на устойчивость часто применяют близкородственные виды секции *Beta*, например, свеклу морскую (*B. vulgaris* ssp. *maritima*) как источник генов устойчивости к ризомании и церкоспорозу [20]. В отечественной селекции на устойчивость к биотическим стрессам использовались и представители секции *Corollinae*, такие как *B. corolliflora* и *B. trigyna* [2]. Разработанная система NBS-профайлинга в дальнейшем может быть использова-

на для изучения полиморфизма и поиска доноров генов устойчивости для селекции у видов рода *Beta*.

Таким образом, был впервые изучен полиморфизм NBS-LRR генов устойчивости современных гибридов сахарной свеклы методом NBS-профайлинга. Выявлен высокий уровень полиморфизма генов устойчивости как отечественных, так и зарубежных гибридов и показана их дифференциация в зависимости от происхождения и компании производителя.

Работа выполнена в рамках темы государственного задания ИОГен РАН “Разработка и применение систем молекулярных маркеров для создания отечественных гибридов сахарной свеклы”.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

#### СПИСОК ЛИТЕРАТУРЫ

1. McGrath J.M., Panella L. Sugar Beet Breeding // Plant Breeding Reviews. 2018. V. 42. P. 167–218. <https://doi.org/10.1002/9781119521358.ch5>
2. Корниенко А.В., Подвигина О.А., Жужжалова Т.П. и др. Приоритетные направления исследований по генетике и селекции сахарной свеклы (*Beta vulgaris* L.) в XXI веке // Генетика. 2014. Т. 50. № 11. С. 1286–1298. <https://doi.org/10.7868/S001667581411006X>
3. Dohm J.C., Minoche A.E., Holtgräwe D. et al. The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*) // Nature. 2014. V. 505. P. 546–549. <https://doi.org/10.1038/nature12817>
4. Funk A., Galewski P., McGrath J.M. Nucleotide-binding resistance gene signatures in sugar beet, insights from a new reference genome // The Plant J. 2018. V. 95. P. 659–671. <https://doi.org/10.1111/tpj.13977>
5. Cai D., Kleine M., Kifle S. et al. Positional cloning of a gene for nematode resistance in sugar beet // Science. 1997. V. 275. P. 832–834. <https://doi.org/10.1126/science.275.5301.832>
6. Lein J.C., Asbach K., Tian Y. et al. Resistance gene analogues are clustered on chromosome 3 of sugar beet and cosegregate with QTL for rhizomania resistance // Genome. 2007. V. 50. P. 61–71. <https://doi.org/10.1139/g06-131>
7. Capistrano-Gossmann G.G., Ries D., Holtgräwe D. et al. Crop wild relative populations of *Beta vulgaris* allow direct mapping of agronomically important genes // Nat. Commun. 2017. V. 8. 15708. <https://doi.org/10.1038/ncomms15708>
8. Van der Linden C.G., Wouters D.C., Mihalka V. et al. Efficient targeting of plant disease resistance loci using NBS profiling // Theor. Appl. Genet. 2004. V. 109. № 2. P. 384–393. <https://doi.org/10.1007/s00122-004-1642-8>
9. Mantovani P., van der Linden G., Maccaferri M. et al. Nucleotide-binding site (NBS) profiling of genetic diversity in durum wheat // Genome. 2006. V. 49. № 11. P. 1473–1480. <https://doi.org/10.1139/g06-100>
10. Горюнова С.В., Гашкова И.В., Косарева Г.А. Изменчивость и филогенетические связи вида *Cucumis sativus* L. по данным NBS-профайлинга и RAPD-анализа // Генетика. 2011. Т. 47. № 8. С. 1052–1063.
11. Савельева Е.Н., Борис К.В., Кочиева Е.З., Кудрявцев А.М. Полиморфизм NBS-LRR генов устойчивости сортов яблони (*Malus domestica* Borkh.) по данным NBS-профайлинга // Генетика. 2016. Т. 52. № 12. С. 1463–1468. <https://doi.org/10.7868/S0016675816120110>
12. Дьяченко Е.А., Кулакова А.В., Кочиева Е.З., Щенникова А.В. Вариабельность геномных RGA-локусов современных отечественных сортов картофеля: данные NBS-маркирования // С.-хоз. биология. 2021. Т. 56. № 1. С. 32–43. <https://doi.org/10.15389/agrobiology.2021.1.32rus>
13. Benbouza H., Jacquemin J.M., Baudoin J.P., Mergeai G. Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels // BASE. 2006. V. 10. № 2. P. 77–81.
14. Hammer O., Harper D.A.T., Ryan P.D. PAST: Paleontological Statistics software package for education and data analysis // Paleontologia Electronica. 2001. V. 4. № 1. P. 1–9.
15. Peakall R., Smouse P.E. GenALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research an update // Bioinformatics. 2012. V. 28. № 19. P. 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
16. Kumar S., Stecher G., Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets // Mol. Biol. and Evol. 2016. V. 33. № 7. P. 1870–1874. <https://doi.org/10.1093/molbev/msw054>
17. Smulders M.J.M., Esselink G.D., Everaert I. et al. Characterization of sugar beet (*Beta vulgaris* L. ssp. *vulgaris*) varieties using microsatellite markers // BMC Genetics. 2010. V. 11. № 41. P. 1–11. <https://doi.org/10.1186/1471-2156-11-41>
18. Коваль Е.В., Трифонова А.А., Дедова Л.В. и др. Генетическое разнообразие отечественных и зарубежных гибридов сахарной свеклы (*Beta vulgaris* L.) по данным AFLP-анализа // Генетика. 2018. Т. 54. № 13. С. 37–40. <https://doi.org/10.1134/S001667581813009X>
19. Simko I., Eujayl I., van Hintum T.J.L. Empirical evaluation of DArT, SNP, and SSR marker-systems for genotyping, clustering, and assigning sugar beet hybrid varieties into populations // Plant Science. 2012. V. 184. P. 54–62. <https://doi.org/10.1016/j.plantsci.2011.12.009>
20. Biancardi E., Panella L.W., Lewellen R.T. Beta Maritima: The Origin of Beets. New York: Springer, 2012. 294 p. <https://doi.org/10.1007/978-1-4614-0842-0>

## **NBS-LRR Resistance Genes Polymorphism of Sugar Beet Hybrids according to NBS-Profiling Data**

**A. A. Trifonova<sup>a, \*</sup>, E. R. Paradnya<sup>a</sup>, K. V. Boris<sup>a</sup>, and A. M. Kudryavtsev<sup>a</sup>**

*<sup>a</sup>Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, 119991 Russia*

*\*e-mail: aichka89@mail.ru*

NBS-profiling was used to study 49 modern Russian and foreign sugar beet hybrids. The detected NBS-LRR resistance genes polymorphism was rather high – 87.77% (74.21% for 15 Russian hybrids and 83.80% for 34 foreign hybrids), for 11 hybrids unique NBS-fragments were detected. Cluster analysis revealed the differentiation of Russian and foreign sugar beet hybrids, which indicates a difference in their sets of resistance genes. In addition, the clustering of a number of foreign hybrids was revealed depending on the seed company.

**Keywords:** *Beta vulgaris*, genetic diversity, R-genes.