

ISSN 0016-6758

Том 60, Номер 12

Декабрь 2024



# ГЕНЕТИКА



НАУКА

— 1727 —

# СОДЕРЖАНИЕ

---

Том 60, номер 12, 2024

---

## Обзорные и теоретические статьи

Номенклатура маркеров и контроль ошибок первого и второго рода при проведении генетических экспертиз достоверности происхождения крупного рогатого скота с использованием фрагментного анализа

*М. В. Модоров, И. В. Ткаченко, А. А. Клещева, М. Ю. Севостьянов*

3

Индукция тренированного иммунитета с помощью БЦЖ: недавние данные и мнения

*И. В. Алексеенко, Л. Г. Кондратьева, И. П. Чернов, Е. Д. Сверлов*

16

## Генетика микроорганизмов

Репаративная ДНК полимераза eta (Polη) играет ключевую роль в мутагенезе при низких дозах УФ-облучения у дрожжей *Saccharomyces cerevisiae*

*Е. А. Алексеева, Т. А. Евстюхина, И. И. Скобелева, В. Т. Пешехонов, В. Г. Королев*

25

## Генетика растений

Транскрипционная активность *CCA1* у растений *Arabidopsis thaliana* северной популяции в условиях измененного светового режима

*М. В. Зарецкая, О. М. Федоренко*

36

Гены партеногенеза кукурузы: сравнительный анализ мутаций

*Е. М. Мусеева, В. В. Фадеев, Ю. В. Фадеева, С. И. Мазилов, А. Ю. Колесова, М. И. Чумаков*

43

## Генетика животных

Генетическая изменчивость и филогенетические взаимоотношения популяций обыкновенной щуки (*Esox lucius L.*) из некоторых рек России

*Л. Т. Бачевская, В. В. Переверзева, А. А. Примак, Г. А. Агапова*

54

Оценка полиморфизма генов иммунной системы *TNF* и *TLR2* у крупного рогатого скота

*Л. А. Калашникова, И. Е. Багаль, Н. Е. Муругина, В. Е. Калашников, Р. Ю. Сенина*

66

Характеристика по STR-маркерам серого украинского скота, разводимого в Российской Федерации

*А. С. Мокеев, Н. Н. Фурса, С. В. Бекетов, Г. Р. Свищева, А. А. Онохов, Ю. А. Столповский*

75

## Генетика человека

Популяционно-генетическая структура населения Сибири по данным о частотах полиморфных вариантов генов транспорта и рецепции витамина D

*Б. А. Тхоренко, А. В. Майер, Г. В. Вавин, Д. О. Имекина,  
М. В. Ульянова, Ф. А. Лузина, М. Б. Лавряшина*

83

---

**Математические модели и методы**

О стабилизации темпа дивергенции изонимии

*В. П. Пасеков*

94

---

**Краткие сообщения**Паритетное соотношение самцов и самок в выводках красавки *Anthropoides virgo*  
в Калмыкии и Бурятии*Е. А. Мудрик, Е. И. Ильяшенко, К. Д. Кондракова, А. А. Абушин,  
Л. Д. Базаров, Ц. З. Доржиев, Д. В. Политов*

110

Влияние генотипа по гену алкогольдегидрогеназы *ADH1B* на уровень биомаркера  
 злоупотребления алкоголем CDT*А. А. Ким, А. С. Гуреев, А. В. Рубанович, С. А. Боринская*

115

---

**Персоналии**Поздравление с 90-летним юбилеем академика РАН  
Шестакова Сергея Васильевича

121

---

**Поправки**

122

---

ОБЗОРНЫЕ И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

---

УДК 636.082.21

## НОМЕНКЛАТУРА МАРКЕРОВ И КОНТРОЛЬ ОШИБОК ПЕРВОГО И ВТОРОГО РОДА ПРИ ПРОВЕДЕНИИ ГЕНЕТИЧЕСКИХ ЭКСПЕРТИЗ ДОСТОВЕРНОСТИ ПРОИСХОЖДЕНИЯ КРУПНОГО РОГАТОГО СКОТА С ИСПОЛЬЗОВАНИЕМ ФРАГМЕНТНОГО АНАЛИЗА

© 2024 г. М. В. Модоров<sup>1,\*</sup>, И. В. Ткаченко<sup>1</sup>, А. А. Клещева<sup>1</sup>, М. Ю. Севостьянов<sup>1</sup>

<sup>1</sup>Уральский федеральный аграрный научно-исследовательский центр Уральского отделения

Российской академии наук, Екатеринбург, 620142 Россия

\*e-mail: mmodorov@gmail.com

Поступила в редакцию 13.05.2024 г.

После доработки 18.06.2024 г.

Принята к публикации 09.07.2024 г.

Ошибочные записи родословных снижают качество племенной работы с крупным рогатым скотом, поэтому проведение генетических экспертиз достоверности происхождения стало неотъемлемой частью племенной работы. В течение многих лет на территории Российской Федерации экспертизы проводили с использованием иммуногенетических маркеров, однако совершенствование технологий и ужесточение требований регулятора привело к тому, что в последние годы на смену иммуногенетике приходят микросателлитные маркеры. В настоящее время отечественный протокол, четко регламентирующий процедуру проведения генетической экспертизы достоверности происхождения крупного рогатого скота с использованием микросателлитных локусов, отсутствует, что затрудняет работу испытательных лабораторий. В частности, требования к числу и номенклатуре генетических маркеров, которые должны быть использованы при проведении экспертиз, имеются только для скота, эмбрионов и спермопродукции, перемещаемых между государствами-членами Евразийского экономического союза. Отсутствуют требования к контролю ошибок первого (ложноположительные результаты) и второго (ложноотрицательные результаты) рода, учет которых необходим при формировании заключений экспертизы. В настоящей работе мы рассмотрим подходы к решению этих вопросов, предложенные Международным обществом генетики животных (ISAG), Международным комитетом регистрации животных (ICAR), Коллегией Евразийской экономической комиссии, а также отечественными нормативными документами, регламентирующими производство судебно-медицинских экспертиз, связанных с установлением родительства. По результатам обзора будут предложены номенклатура микросателлитных маркеров и протокол экспертизы, в который заложен контроль ошибок первого и второго рода. Особое внимание будет уделено описанию источников ошибок второго рода и необходимости их контроля при выдаче заключений.

**Ключевые слова:** *Bos taururs*, микросателлиты, селекция, ISAG, STR, индексы исключения родителей.

**DOI:** 10.31857/S0016675824120012 **EDN:** WAWXIF

К середине XX в. искусственное осеменение крупного рогатого скота прочно вошло в практику животноводства. При этом возросло количество ошибочных записей “отец – потомок”, что вносит значительное искажение как в оценку индексов наследования хозяйственно ценных признаков, так и в оценку качества быков по потомству [1, 2]. Для преодоления возникших трудностей в практику работы с племенным скотом вошли генетические тесты, использование которых позволило значительно снизить число ошибочных записей в

родословных. Вначале подобные тесты проводили с использованием иммуногенетических маркеров (групп крови). В настоящее время в качестве генетических маркеров используют STR-локусы (short tandem repeat, микросателлиты) либо SNP-локусы (single nucleotide polymorphism, однонуклеотидные полиморфизмы).

На территории РФ работа лабораторий, проводящих подтверждение достоверности происхождения скота с использованием генетических маркеров, в настоящее время требует получения

Свидетельства о регистрации в государственном племенном регистре. Ситуация изменилась с вступлением в силу Приказа Минсельхоза России № 336 от 2 июня 2022 г. “Об утверждении требований к видам племенных хозяйств”. Согласно пункту 8 данного Приказа к лабораториям молекулярно-генетической экспертизы в качестве дополнительного требования устанавливается “наличие документа об аккредитации в национальной системе аккредитации”.

К настоящему моменту отечественный ГОСТ либо методические указания, содержащие обязательные условия к процедуре подтверждения достоверности происхождения скота с использованием фрагментного анализа, отсутствуют. Это вносит значительную неопределенность при подготовке лабораторий к аккредитации, в связи с чем интерес представляет анализ требований, предъявляемых к данному типу анализов за рубежом, а также положений отечественных нормативных документов, регламентирующих производство судебно-медицинских экспертиз, связанных с установлением родительства. В настоящей работе мы рассмотрим вопросы, касающиеся научной составляющей проведения исследований достоверности происхождения крупного рогатого скота (обсуждение требований к лабораторным помещениям и организации менеджмента качества испытательных лабораторий в наши задачи не входит), а именно:

1) генетическая идентификация (выбор генетических маркеров, их числа и номенклатуры);

2) достоверность происхождения (сопоставление микросателлитных профилей с учетом контроля ошибок первого и второго рода, т. е. ложноположительных и ложноотрицательных результатов теста).

Цель настоящей работы – разработка протокола сопоставления генетических профилей предполагаемых родителей и потомка, в котором учтены интересы основных участников и выгодоприобретателей проводимых экспертиз, а именно племенной службы, сельхозпроизводителей, поставщиков спермопродукции и лабораторий, проводящих экспертизы.

#### *Обзор требований, предъявляемых к проведению фрагментного анализа, связанного с подтверждением родительства*

*ISAG*. Наиболее влиятельной международной организацией, определяющей тип, число и номенклатуру генетических маркеров, используемых в тестах для подтверждения достоверности происхождения сельскохозяйственных животных, является Международное общество генетики животных (International Society for Animal Genetics, ISAG). В частности, ISAG разрабатывает руководства по

управлению генетическими ресурсами сельскохозяйственных животных в интересах FAO (Food and Agriculture Organization, Продовольственная и сельскохозяйственная организация ООН) [3, 4].

К 2006 г. для проведения тестов на достоверность происхождения крупного рогатого скота ISAG было рекомендовано девять микросателлитных локусов: BM1824, INRA23, BM2113, SPS115, ETH10, TGLA122, ETH225, TGLA126 и TGLA227 [5]. В 2008 г. к ним было добавлено еще три локуса, а именно: BM1818, ETH3 и TGLA53 [6]. По состоянию на 2021-й год минимальное число STR-локусов, рекомендованных ISAG для подтверждения достоверности происхождения крупного рогатого скота, не изменилось, оно остается равным 12 [7]. Раз в два года ISAG разрабатывает правила и проводит сравнительные испытания (comparison tests), в ходе которых лаборатории получают пробирки с выделенной ДНК нескольких пород скота. Задачей лабораторий, добровольно участвующих в данном испытании, является корректная расшифровка неизвестных генотипов по 12 локусам, т. е. генетическая идентификация скота.

*ICAR*. Помимо ISAG, сертификацию лабораторий, проводящих подтверждение достоверности происхождения крупного рогатого скота, проводит ICAR (International Committee for Animal Recording, Международный комитет регистрации животных). ICAR рекомендует использовать тот же набор STR-маркеров, что и ISAG. Для аккредитации в ICAR лаборатории необходимо соответствовать более жестким требованиям, чем в ISAG, в частности, иметь сертификат аккредитации по ISO 17025. Кроме того, лабораториям необходимо проводить расчеты индексов исключения родителей (exclusion probability) как для каждого тестируемого локуса, так и для всей панели маркеров. Популяция, для которой проводят расчет индексов, должна быть охарактеризована по породной принадлежности и числу проанализированных голов, которых должно быть не менее 150 [8].

*Решение Коллегии Евразийской экономической комиссии*. Русскоязычным документом, определяющим номенклатуру локусов, используемых для подтверждения достоверности происхождения крупного рогатого скота, и требования к проводящим тесты лабораториям, является Решение Коллегии Евразийской экономической комиссии от 2 июня 2020 г. № 74 “Об утверждении Положения о проведении молекулярной генетической экспертизы племенной продукции государств – членов Евразийского экономического союза” [9]. Согласно данному документу, проведение молекулярной генетической экспертизы осуществляется лабораториями, аккредитованными в государственных (национальных) системах аккредитации

государств – членов Евразийского экономического союза (далее государства – члены) либо ICAR (пункт 3).

Обязательной молекулярной генетической экспертизе подлежат племенные производители сельскохозяйственных животных (крупный рогатый скот, лошади, овцы, козы, свиньи, олени, верблюды), перемещаемые между государствами-членами, а также племенные производители и доноры эмбрионов сельскохозяйственных животных, спермопродукция и эмбрионы которых перемещаются между государствами-членами (п. 6). Молекулярная генетическая экспертиза племенной продукции осуществляется методом ДНК-типовирования с использованием методик, разработанных с учетом рекомендаций ISAG, в соответствии с областью аккредитации лаборатории (п. 7). Проведение молекулярной генетической экспертизы с целью подтверждения достоверности происхождения племенной продукции осуществляется методом генотипирования STR-маркеров или SNP-маркеров (п. 9). Перечень STR-маркеров должен составлять не менее 12 STR-маркеров, включенных в базовую STR-панель ISAG.

*В Приказе Минсельхоза России № 336 от 2 июня 2022 г. “Об утверждении требований к видам племенных хозяйств” номенклатура генетических маркеров, требуемых для проведения генетических тестов на достоверность происхождения скота, отсутствует [10]. Однако этот документ определяет группы крупного рогатого скота, для которых генетическая экспертиза достоверности происхождения должна быть проведена. Это скот, содержащийся в племенных заводах, племенных репродукторах и генофондных хозяйствах; молочные породы скота (100% ремонтного молодняка для реализации, 100% матерей быков-производителей), мясные породы скота (100% быков-производителей и коров быкопроизводящей группы).*

*Нормативная база, регламентирующая производство судебно-медицинских экспертиз, связанных с установлением родительства. В настоящее время (весна 2024 г.) на территории РФ действует Приказ Министерства здравоохранения и социального развития Российской Федерации от 12 мая 2010 г. № 346н “Об утверждении Порядка организации и производства судебно-медицинских экспертиз в государственных судебно-экспертных учреждениях Российской Федерации” [11]. Особенности порядка производства генетической экспертизы описаны в 84-м разделе данного документа. Согласно ему генетическая экспертиза по поводу спорного происхождения детей отвечает на вопросы:*

1) исключается или не исключается отцовство, материнство данного индивидуума в отношении данного ребенка (плода);

2) если отцовство, материнство не исключается, то какова вероятность того, что полученный результат не является следствием случайного совпадения индивидуализирующих признаков неродственных лиц.

Данный документ не регламентирует число и номенклатуру используемых локусов, однако предъявляет требования к вероятностной оценке неисключенного родства (пп. 84.12.5 и 84.12.6). Принимается, что совпадение условно отцовских (нематеринских) аллелей в генотипе ребенка с аллелями, присутствующими в геноме предполагаемого отца, не означает доказанного отцовства. Вероятностная оценка неисключенного родства строго обязательна. Искомое условное значение вероятности отцовства соответствует байесовской постериорной вероятности при 50%-ной априорной вероятности отцовства. Уровень доказательности экспертного исследования в случае неисключения отцовства, материнства должен составлять для полного трио (мать – ребенок – предполагаемый отец) при условии, что истинность другого родителя считается бесспорной, не ниже 99.90% (рассчитываемый как байесовская вероятность отцовства/материнства); не ниже 1000 (рассчитываемый как индекс отцовства PI). Для дуэта (ребенок – предполагаемый отец) в отсутствие другого родителя – не ниже 99.75% (рассчитываемый как байесовская вероятность отцовства/материнства); не ниже 400 (рассчитываемый как индекс отцовства PI). Согласно п. 84.12.7 при количественном анализе результата для расчета вероятности отцовства (материнства) следует учитывать этническую принадлежность обследуемых лиц. Кроме этого, в Приказе предъявляются требования к отрицательному заключению (п. 84.12.4). Для обоснованного вывода о безусловном исключении отцовства, материнства аллели ребенка, не свойственные ни одному из указанных родителей, должны быть зарегистрированы как минимум в двух несцепленных локусах.

С 1 сентября 2024 г. на смену Приказу Министерства здравоохранения и социального развития Российской Федерации от 12 мая 2010 г. № 346н приходит Приказ № 491н от 25 сентября 2023 г. “Об утверждении Порядка проведения судебно-медицинской экспертизы” [12]. Приложение № 11 нового приказа посвящено “Правилам организации деятельности отделения генетической экспертизы и проведения генетической экспертизы”. Согласно этому документу, для вывода об исключении отцовства или материнства аллели ребенка, не свойственные ни одному из предполагаемых родителей, регистрируются сразу в нескольких несцепленных локусах. Для вывода о том, какое лицо является родителем данного ребенка, необходима вероятностная оценка наблюдаемого

совпадения/несовпадения индивидуальных профилей полиморфизма ДНК сравниваемых биологических объектов в рамках интерпретационной модели противоположных версий: по умолчанию – версии несовпадения аллелей по причине нарушения базовых закономерностей их совпадения и версии аллельного несовпадения у не состоящих в родстве лиц. Расчетная вероятностная оценка рассматриваемых версий спорного родительства (исключение/неисключение отцовства или материнства) соответствует условной байесовской апостериорной вероятности при 50%-ной априорной вероятности родительства. При количественном анализе результата для расчета вероятности отцовства (материнства) учитывается этническая принадлежность обследуемых лиц. Методическое содержание и объем исполнения завершенного исследования спорного происхождения детей предусматривают в типовом случае анализа генотипирование у каждого из обследуемых лиц не менее 16 несцепленных полиморфных tandemных локусов-маркеров аутосомной ДНК.

Проанализированные материалы позволяют говорить о том, что к настоящему времени существует консенсус в номенклатуре микросателлитных маркеров, используемых для проведения экспертиз на достоверность происхождения крупного рогатого скота, а именно использование не менее 12 локусов базовой панели ISAG. В то же время в отличие от требований, предъявляемых к проведению судебно-медицинских экспертиз, контроль ошибок первого и второго рода у крупного рогатого скота не регламентирован.

#### *Источники ошибок при сравнении генотипов предполагаемых родителей и потомка*

При проведении экспертизы достоверности происхождения крупного рогатого скота необходимо контролировать ошибки первого и второго рода. Ошибка первого рода (ложноположительное заключение) – ситуация, в которой подтверждается родительство особи, которая родителем не является. Ошибка второго рода (ложноотрицательное заключение) – выдача отрицательного заключения на настоящего родителя. При проведении судебно-медицинских экспертиз для контроля ошибки первого рода вводятся пороговые значения индексов байесовской вероятности отцовства/материнства либо индекса отцовства (PI). Кроме этого, в Приказе № 491н от 25 сентября 2023 г. вводится требование к минимальному числу локусов, используемых для анализа. Для контроля ошибки второго рода вывод об исключении отцовства или материнства может быть сделан в случае наличия у ребенка аллелей, не свойственных ни одному из предполагаемых родителей, сразу в нескольких несцепленных локусах (минимум двух).

В исследованиях крупного рогатого скота для контроля ошибки первого рода ICAR требует рассчитывать индексы исключения родителей, однако в просмотренных документах нам не удалось найти порогового значения индексов, превышение которого необходимо для оформления заключений. Контроль ошибки второго рода у крупного рогатого скота при использовании панели STR-маркеров, по-видимому, не проводится. Однако при использовании для анализа SNP-маркеров допускается, что подтверждение родительства происходит при отсутствии общих аллелей в двух из 186 протестированных локусов, при несовпадении аллелей в трех – пяти локусах предполагаемый родитель получает статус “сомнительный” (doubtful) и только при несовпадении более пяти локусов родитель исключается [8].

#### *Индексы исключения родителей (контроль ошибки первого рода или выдача ложноположительных заключений)*

Наличие общих аллелей по каждому исследованному локусу у предполагаемого родителя и потомка может быть объяснено двумя способами:

- 1) потомок происходит от предполагаемого родителя;
- 2) аллели у анализируемых индивидов совпали случайно.

Высокая вероятность случайного совпадения аллелей у предполагаемого родителя и потомка приведет к высокой частоте ложноположительных экспертных заключений, подтверждающих происхождение крупного рогатого скота (т. е. к статистической ошибке первого рода). Для контроля этой ошибки при проведении судебно-медицинских экспертиз требуется рассчитать индекс байесовской вероятности отцовства/материнства либо индекс отцовства (PI). Эти индексы рассчитываются с учетом частоты встречаемости каждого обнаруженного аллеля в популяции (этносе) [13–15]. Таким образом, при использовании единой панели генетических маркеров в каждой проведенной экспертизе эта вероятность требует отдельного расчета.

ICAR следует иному подходу и для контроля ошибки первого рода предлагает рассчитывать индексы исключения родителей [8]. Формулы, позволяющие оценить эти индексы, приведены на сайте ISAG [16], более подробную информацию об этих формулах можно найти в работе Джеймсона и Тэйлора [17]. Эти индексы характеризуют локус (а не аллель), поэтому могут быть рассчитаны для определенной популяции единожды. По сути индексы исключения родителей показывают вероятность того, что у случайно отобранных в

популяции особей в тестируемом локусе не будет сходного аллеля [18].

Отметим, что значительную путаницу при использовании индексов исключения родителей вносит то, что в публикации Джеймесона и Тэйлора все они называются просто “P”, без какого-либо дополнительного буквенного или цифрового обозначения. В настоящей работе вслед за разработчиками программы GenAIEx 6.5 [19] мы будем называть эти индексы  $P_1$ ,  $P_2$ ,  $P_3$ , что соответствует последовательности написания формул в работе [17].

1. Формула, используемая для расчета индекса исключения родителей, определяется тем, какие особи были выбраны для сравнения генотипов. В наиболее простом случае проводят сравнение генотипов потомка и **одного** предполагаемого родителя (данные о втором родителе в этом teste не важны, формула 1, индекс  $P_2$ ).

$$P_2 = 1 - 4 \sum_{i=1}^n p_i^2 + 2 \left( \sum_{i=1}^n p_i^2 \right)^2 + 4 \sum_{i=1}^n p_i^3 - 3 \sum_{i=1}^n p_i^4 , \quad (1)$$

где  $p_i$  – частоты аллелей в тестируемом локусе.

2. Во второй ситуации тестируется происхождение потомка от двух предполагаемых родителей (индекс  $P_3$ , формула 2), т. е. вероятность отсутствия у потомка хотя бы одного аллеля, один из которых сходен с аллелем предполагаемого отца, а другой с аллелем предполагаемой матери:

$$P_3 = 1 + 4 \sum_{i=1}^n p_i^4 - 4 \sum_{i=1}^n p_i^5 - 3 \sum_{i=1}^n p_i^6 - 8 \left( \sum_{i=1}^n p_i^2 \right)^2 + \\ + 8 \left( \sum_{i=1}^n p_i^2 \right) \left( \sum_{i=1}^n p_i^3 \right) + 2 \left( \sum_{i=1}^n p_i^3 \right)^2 , \quad (2)$$

где  $p_i$  – частоты аллелей в тестируемом локусе.

В практической работе лабораторий, специализирующихся на проведении генетических экспертиз на достоверность происхождения крупного рогатого скота, абсолютное большинство тестов описывается двумя вышеупомянутыми формулами.

3. Формула 3 (индекс  $P_1$ ) применима в ситуации установления отцовства, когда проверке подлежит не два родителя, а только один (например, в происхождении ребенка от матери сомнений нет, а вот отец сомнения вызывает). При этом известны генотипы всей триады: потомок, известный родитель и предполагаемый родитель. В практической работе по проведению генетических экспертиз на достоверность происхождения крупного рогатого скота с подобной ситуацией мы не сталкивались,

однако при проведении судебно-медицинских экспертиз этот индекс может быть использован:

$$P_1 = 1 - 2 \sum_{i=1}^n p_i^2 + \sum_{i=1}^n p_i^3 + 2 \sum_{i=1}^n p_i^4 - 3 \sum_{i=1}^n p_i^5 - \\ - 2 \left( \sum_{i=1}^n p_i^2 \right)^2 + 3 \sum_{i=1}^n p_i^2 \sum_{i=1}^n p_i^3 , \quad (3)$$

где  $p_i$  – частоты аллелей в тестируемом локусе.

С ростом значения индекса растет его информативность для решения задач по проведению генетических экспертиз на достоверность происхождения крупного рогатого скота. Если в исследовании использовано несколько ( $n$ ) локусов, то для оценки индексов исключения родителей с использованием всей панели маркеров может быть применена формула:

$$P = 1 - \prod_{k=1}^n (1 - P_k) \quad (4)$$

где  $P_k$  – значение индекса в каждом тестируемом локусе.

В табл. 1 приведены рассчитанные значения индексов исключения родителей для популяции племенного голштинизированного черно-пестрого скота Свердловской области. Для расчетов использована база данных, включающая генотипы 24544 голов скота, расшифровка которых была проведена в лаборатории молекулярно-генетической экспертизы Уральского НИИСХ – филиала ФГБНУ УрФАНИЦ УрО РАН в период с 2017 по 2021 г. Методика генотипирования и описание популяции приведены нами ранее [20].

Индекс  $P_3$  для панели, включающей 12 рекомендованных ISAG локусов, составляет 0.999999705. Это значение позволяет сформировать ожидания относительно результатов 10 миллионов сравнений генотипов, случайно взятых из популяции отцов, матерей и потомка. Результаты 9 999 997 подобных сравнений покажут, что хотя бы в одном локусе триады отец – мать – потомок для потомка нельзя будет подобрать аллели, наследуемые от предполагаемого отца и предполагаемой матери. В трех оставшихся тестах в каждом локусе потомка будет обнаружен аллель, сходный с аллелем предполагаемого отца и предполагаемой матери (ложноположительное сходство).

Значение индекса  $P_2$  для панели, включающей 12 рекомендованных ISAG локусов, составляет 0.9946. Это значение показывает, что при сравнении генотипов, случайно взятых из популяции матери и потомка, в 9946 тестах из 10000 будет получен результат, в котором хотя бы в одном локусе не будет обнаружен общий аллель, а в 54 оставшихся тестах общий аллель будет обнаружен во всех

**Таблица 1.** Индексы исключения родителей в популяции голштинизированного черно-пестрого скота Свердловской области по данным изменчивости 24544 голов. Значения индексов приведены для каждого локуса, для панели из 12 локусов, рекомендованных ISAG, а также панели из 15 локусов, входящих в набор СOrDIS Cattle

Локус	Индексы исключения родителей		
	P1	P2	P3
BM1824*	0.47	0.30	0.65
BM2113*	0.55	0.37	0.74
ETH3*	0.44	0.27	0.62
ETH10*	0.58	0.40	0.77
ETH225*	0.44	0.27	0.63
INRA23*	0.53	0.35	0.71
SPS115*	0.34	0.18	0.53
TGLA53*	0.71	0.55	0.88
TGLA122*	0.68	0.52	0.86
TGLA126*	0.33	0.19	0.50
TGLA227*	0.66	0.48	0.84
BM1818*	0.35	0.21	0.52
Csrm60	0.42	0.25	0.61
Cssm66	0.52	0.34	0.72
Ilsts006	0.40	0.24	0.57
12 локусов	0.999868	0.9946	0.999999705
15 локусов	0.999978	0.9979	0.999999986

Примечание. \* – локус включен в базовую панель ISAG.

локусах. Это же касается пары предполагаемый отец – потомок.

Высокие значения индексов *P3*, полученные для панели, включающей как 12, так и 15 локусов, на наш взгляд, позволяют проводить не только подтверждение достоверности происхождения, но и восстановление пары родителей (подбор родителей на основании данных о генотипах скота в стаде). Уровни значений индексов *P2* позволяют проводить подтверждение достоверности происхождения в популяции, но недостаточны для восстановления родителя. Отметим, что индексы исключения родителей у локальных пород скота во многих случаях ниже, чем у голштинской породы либо голштинизированного скота [21].

#### Контроль ошибки второго рода (ложноотрицательные заключения)

Причинами отсутствия общих аллелей у родителя и потомка могут быть *de novo* мутации, главным образом изменение в числе мотивов микросателлита, мутации в последовательности ДНК, комплементарной последовательности праймера, ошибки генотипирования, путаница при отборе проб и составлении описей.

**Мутационный процесс.** Частота мутаций в динуклеотидных микросателлитных локусах у человека составляет  $7.88 \times 10^{-5}$  на поколение или одно событие на 12690 проведенных локус-тестов. Наиболее часто мутационное событие заключается в изменении числа мотивов микросателлита на один, при этом четкой линейной связи числа повторов с числом мутационных событий для динуклеотидных повторов не отмечено [22]. Если принять сходный уровень скорости мутационного процесса у человека и крупного рогатого скота, то при использовании панели, включающей 12 микросателлитных локусов, мы вправе ожидать около одного мутационного события примерно на 1000 ( $1 \times 10^{-3}$ ) проведенных тестов (точнее на 1058 тестов, 12690 событий / 12 локусов в тесте = 1058). Таким образом, при выставлении требований к 100%-ному совпадению аллелей во всех 12 локусах крупного рогатого скота ложноотрицательные результаты будут получены в каждом тысячном сравнении родителя и потомка либо в каждом 500-м тесте при сравнении двух родителей и потомка. Отметим, что в одном из миллиона проведенных тестов ( $1 \times 10^{-3} \times 1 \times 10^{-3} = 1 \times 10^{-6}$ ) вследствие мутационного процесса будет наблюдаться несовпадение аллелей в двух локусах.

Мутационный процесс может привести к изменениям в последовательности ДНК, комплементарной последовательности праймера. Следствием мутации может стать либо снижение высоты пика, либо его выпадение (так называемые нуль-аллели). В этом случае гетерозиготный генотип будет считан как гомозигота. Маркером нуль-аллелей может быть отклонение числа наблюдаемых генотипов от их числа, рассчитанного с использованием уравнения Харди–Вайнберга, в сторону увеличения доли гомозигот либо, если нуль-аллели встречаются в гомозиготе, отсутствие пиков в данном локусе. В случае наследования нуль-аллеля генотип потомка будет выглядеть как гомозигота по аллелю, унаследованному от другого родителя. На рис. 1 рассмотрены гипотетические примеры ошибок второго рода, связанных с наличием нуль-аллелей. Отметим, что часть этих ошибок может быть связана с использованием наборов реагентов различных производителей. При этом как ISAG, так и ICAR вводят номенклатуру праймеров [4, 8], однако последовательности праймеров в производимых в РФ

	фрагмент 70 (10 мотивов)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacacacacacacaCCTAGTCAGTCAACCTCTCAC <u>GAGCAAGGTGTTTCCAATC</u> ACACTCCTCCACAAAAAGGTTAGATGCgtgtgtgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG
♂	фрагмент 66 (8 мотивов)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacacacacacaCCTAGTCAGTCAACCTCTCAC <u>GAGCAAGGTGTTTCCAATC</u> ACACTCCTCCACAAAAAGGTTA <u>A</u> ATGCgtgtgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG ↑
♀	фрагмент 74 (12 мотивов)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacacacacacacaCCTAGTCAGTCAACCTCTCAC <u>GAGCAAGGTGTTTCCAATC</u> ACACTCCTCCACAAAAAGGTTAGATGCgtgtgtgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG
	фрагмент 58 (4 мотива)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacaCCTAGTCAGTCAACCTCTCAC <u>GAGCAAGGTGTTTCCAATC</u> ACACTCCTCCACAAAAAGGTTAGATGCgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG
F1 <sup>a</sup>	фрагмент 66 (8 мотивов)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacacacacaCCTAGTCAGTCAACCTCTCAC <u>GAGCAAGGTGTTTCCAATC</u> ACACTCCTCCACAAAAAGGTTA <u>A</u> ATGCgtgtgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG ↑
	фрагмент 58 (4 мотива)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacaCCTAGTCAGTCAACCTCTCAC <u>GAGCAAGGTGTTTCCAATC</u> ACACTCCTCCACAAAAAGGTTAGATGCgtgtgtggGGATCAAGGACTCAGTTGGAGAGTG
F1 <sup>b</sup>	фрагмент 69 (8 мотивов)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacacacacaCCTAGTCAGTCAACCTCTCAC <u>TGTGAGCAAGGTGTTTCCA</u> ACACTCCTCCACAAAAAGGTTAAATGCgtgtgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG
	фрагмент 61 (4 мотива)	TGTGAGCAAGGTGTTTCCAATCTACGcacacacaCCTAGTCAGTCAACCTCTCAC <u>TGTGAGCAAGGTGTTTCCA</u> ACACTCCTCCACAAAAAGGTTAGATGCgtgtgtgtGGATCAAGGACTCAGTTGGAGAGTG

**Рис. 1.** Гипотетический пример ошибок второго рода (ложноотрицательное заключение о происхождении потомка от заявленных родителей), связанных с присутствием мутаций в последовательности ДНК, комплементарной последовательности праймера. В первом тесте при определении генотипа быка ( $\sigma$ ), коровы ( $\varphi$ ) и потомка (F1) был использован набор реагентов “a”, в состав которого входит праймер, в месте комплементарного связывания которого с матрицей ДНК у быка произошла мутация (показана стрелкой). В результате мутации фрагмент размером 66 пн не амплифицируется (нуль-аллель). Генотип быка, обозначенный числом мотивов во фрагменте, будет прочитан как “10 / 10” (гомозигота), хотя фактически его генотип “8 / 10” (гетерозигота). Генотип коровы с использованием этого набора реагентов читается без ошибок “4 / 12”. Потомок наследует аллель “8” быка и аллель “4” коровы. При расшифровке генотипа потомка этим же набором реагентов считывается генотип “4 / 4”, так как аллель “8” не амплифицируется из-за мутации в месте комплементарного связывания праймера с матрицей ДНК (вариант F1<sup>a</sup>). Получившиеся данные позволяют сделать заключение об отсутствии в данном локусе общих аллелей у быка и потомка. В другой лаборатории использовали генетические паспорта родителей, а генотип потомка анализировали с использованием другого набора реагентов “b”. В наборе реагентов “b” для анализа данного локуса разработчики включили другие праймеры, комплементарные иным последовательностям матричной ДНК, в результате чего у потомка читаются оба аллеля как “4”, так и “8” (F1<sup>b</sup>). Однако в этой ситуации будет сделан вывод об отсутствии в данном локусе общих аллелей у быка и потомка.

коммерческих наборах реагентов, предназначенных для мультиплексного анализа микросателлитных маркеров крупного рогатого скота, а именно: COrDIS Cattle (ООО “Гордиз” [23]) и Gene Profile Cattle (НПК Синтол [24]), являются коммерческой тайной разработчиков.

*Ошибки расшифровки.* При анализе больших массивов данных ошибки неизбежны. В случае анализа изменчивости микросателлитов они могут быть связаны с малым числом копий выделенной ДНК, присутствием загрязнения в ходе ПЦР, низким качеством электрофореза, плохой калибровкой генетического анализатора и т. п. Оценить

частоту встречаемости этих ошибок достаточно проблематично.

Отдельно следует рассмотреть ошибки, связанные с многоплодной беременностью у коров. Oko-ло 90% дизиготных близнецов крупного рогатого скота имеют в эмбриональный период плацентарные сосудистые связи, в результате чего происходит взаимный обмен первичными кровеобразующими клетками, которые приживляются в кроветворной ткани [1]. Для клеток крови таких животных характерен химеризм, т. е. клетки их крови состоят из генетически разнородных клеток, происходящих от двух зигот. В случае различного пола эмбрионов это приводит к появлению фримартинизма, с последующей выбраковкой телок из племенного генофонда [25]. В случае многоплодной беременности двумя телочками животные могут оставаться в стаде и проходить генетические тесты. При выделении ДНК близнецов из крови фореграмма выглядит как образец, загрязненный одним другим образцом, при этом высота пиков может заметно различаться (рис. 2, а). Фактически подобное “загрязнение” произошло в утробе матери, поэтому все присутствующие у потомков аллели встречаются и у родителей. При выдаче генотипов близнецов есть соблазн включить в генетический паспорт аллели, соответствующие наиболее высоким пикам. В случае если близнец анализируется как потомок, это не кажется плохим решением, так как все аллели потомка соответствуют аллелям родителей. Но когда подобные генетические паспорта предоставляются фермером в качестве генотипов родителей, довольно часто оказывается, что потомки наследуют аллели, высота пиков которых на фореграмме едва выше уровня “шума” (рис. 2). В случае проведения тестов в одной лаборатории такие случаи могут быть “расследованы”, однако при работе с генетическими паспортами, полученными из иных лабораторий, “феномен близнецов” может приводить к выдаче ложноотрицательных заключений.

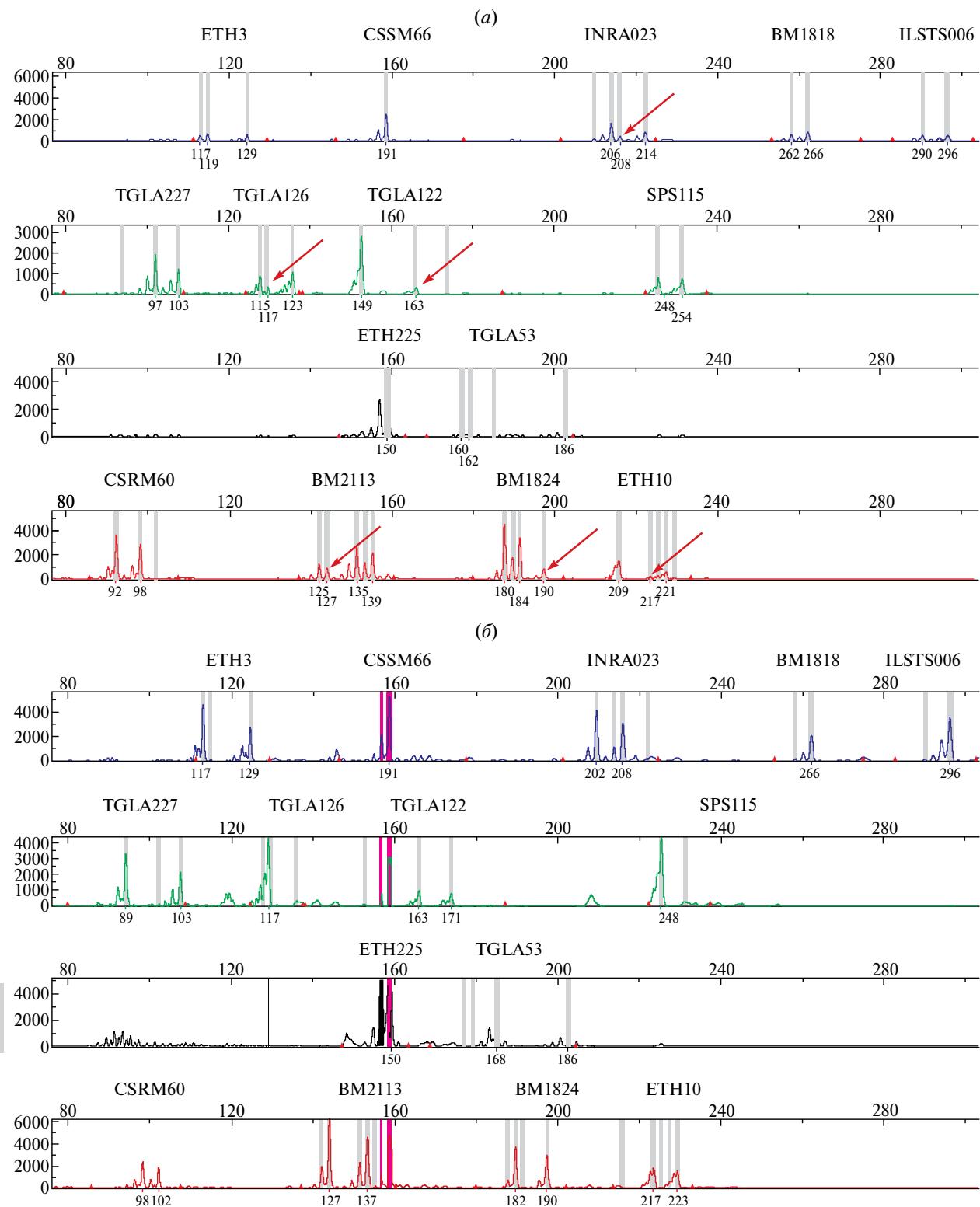
Согласно нашим прямым оценкам, на территории Свердловской области в период с 2000 по 2020 г. появление двух живых телок было отмечено чуть более чем в 1% отелов племенного голштинизированного черно-пестрого скота [26]. Таким образом, стволовые клетки крови более 2% популяции телок региона могут быть химерами и представлять сложности для расшифровки генотипов.

## ЗАКЛЮЧЕНИЕ

Активное перемещение спермопродукции, эмбрионов и поголовья крупного рогатого скота между странами и регионами одной страны делает целесообразной унификацию генетических маркеров, используемых для проведения тестов на достоверность происхождения животных. Наиболее влиятельной международной организацией,

определенющей тип, число и номенклатуру данных маркеров, является ISAG. На сегодняшний день ISAG рекомендует использовать панель, включающую не менее 12 микросателлитных локусов, а именно: BM1824, INRA23, BM2113, SPS115, ETH10, TGLA122, ETH225, TGLA126, TGLA227, BM1818, ETH3 и TGLA53. Эти же локусы входят в панели маркеров, рекомендованных к использованию ICAR и Коллегией Евразийской экономической комиссии. На территории РФ производят два коммерческих набора реагентов, предназначенных для мультиплексного анализа микросателлитных маркеров крупного рогатого скота, а именно: COrDIS Cattle (ООО “Гордиз” [23]) и Gene Profile Cattle (НПК Синтол [24]). В эти наборы, помимо локусов, рекомендованных ISAG, включены дополнительные маркеры, а именно: Csrm60, Cssm66, Ilsts006 (COrDIS Cattle) и Csrm60, Cssm66, Ilsts006, HAUT27 (Gene Profile Cattle). Целесообразность использования для проведения экспертизы на достоверность происхождения скота всех входящих в набор генетических маркеров, либо только 12 локусов, рекомендованных ISAG, определяет вероятность совершения ошибки первого или второго рода при сопоставлении микросателлитных профилей и выдаче заключений о подтверждении либо исключении предполагаемых родителей.

Для контроля ошибки первого рода ICAR и ISAG рекомендуют использовать индексы исключения родителей, однако нам не удалось найти пороговые значения индексов, превышение которых необходимо для выдачи заключения. Оценка индекса  $P_3$  в популяции голштинизированного черно-пестрого скота Свердловской области при использовании панели, включающей 12 локусов, рекомендованных ISAG, составляет 0.999999705 (табл. 1). На наш взгляд, столь высокое значение индекса не только не требует использования в teste дополнительных генетических маркеров, но также позволяет проводить восстановление родословной потомка. Значение индекса  $P_2$  для этой же популяции составляет 0.9946. Эта величина не столь велика, в связи с чем расширение панели генетических маркеров, используемых при сравнении генотипов одного предполагаемого родителя и потомка, является целесообразным. При расширении панели маркеров стоит учитывать, что большинство генетических паспортов быков, выданных за пределами РФ, не включают в себя информацию об изменчивости “дополнительных” локусов, входящих в наборы COrDIS Cattle и Gene Profile Cattle. В связи с этим актуальным становится вопрос о возможности восстановления генотипов быков по “дополнительным” маркерам на основании информации о генотипах подтвержденных потомков и их матерей подобно тому, как это происходило в эпоху иммуногенетики [1]. Альтернативным вариантом



**Рис. 2.** Результаты фрагментного анализа, проведенного с использованием набора реагентов COrDIS Cattle. *а* – предполагаемая мать (дизиготный близнец с химерными клетками крови); *б* – тестируемый потомок (одноплодная беременность). Для каждого фрагмента потомка можно подобрать фрагмент, присутствующий у предполагаемой матери, однако во многих локусах потомок наследовал аллель, высота фрагмента которого у матери была относительно низкой (такие аллели показаны красными стрелками).

получения генотипов быков является использование ДНК, выделенной из спермы.

Рекомендации по контролю ошибки второго рода для тестов, выполненных с использованием STR-маркеров, ISAG и ICAR не дают. Однако ICAR имеет такие рекомендации для тестов, проводимых с использованием панели SNP-маркеров. При этом предлагаются три исхода тестов: родитель подтвержден, родитель сомнителен, родитель исключен. С учетом высокой скорости мутационного процесса, трудностей с выявлением нуль-аллелей и химеризма стволовых клеток крови у близнецов мы считаем, что контроль ошибки второго рода при использовании панели STR-маркеров необходим.

При установлении пороговых величин для контроля ошибок первого и второго уровня оправдано рассмотреть интересы всех участников и выгодоприобретателей экспертиз на достоверность происхождения крупного рогатого скота:

1. Племенная служба. В интересах племенной службы производить выдачу племенных свидетельств только на тех потомков, которые “бесспорно” произошли от предполагаемых родителей. Это требование соответствует высоким значениям индекса исключения родителей при 100%-ном совпадении аллелей во всех тестируемых локусах.

2. Сельхозпроизводитель может иметь две цели. Во-первых, подтверждение того, что спермопродукция происходит от быков, заявленных поставщиком. Это требование соответствует высоким значениям индекса исключения родителей при 100%-ном совпадении аллелей во всех тестируемых локусах. Во-вторых, получение племенного свидетельства на всех заявленных телят, т. е. контроль ложноотрицательных заключений (ошибки второго рода).

3. Поставщики спермопродукции также заинтересованы в контроле ошибки второго рода (выдаче ложноотрицательных заключений).

4. Испытательные лаборатории, проводящие экспертизы, заинтересованы в достоверности выдаваемых экспертиз. На наш взгляд, для этого должны быть соблюдены следующие условия:

4.1) должен проводиться контроль ошибок как первого, так и второго рода, что определяется как фундаментальными основами популяционной генетики, так и несовершенством генетического метода, используемого при проведении экспертиз;

4.2) необходимо иметь возможность получения доступа к первичным файлам генетической идентификации крупного рогатого скота, проведенной в других лабораториях, что особенно важно для быков. На наш взгляд, в едином реестре данных по генотипам крупного рогатого скота к генотипу быка должен быть приложен файл, полученный на генетическом анализаторе, и положительный

контрольный образец (\*.fsa). Данное требование значительно снижает количество ошибочных идентификаций быков;

4.3) совершенствование протоколов определения генетической идентификации близнецов.

Мы предлагаем следующий протокол сопоставления генетических профилей (экспертизы на достоверность происхождения).

Этап I. Задача этапа – подтверждение достоверности происхождения потомка от каждого из родителей.

1. Производится сравнение аллелей в локусах предполагаемого отца и потомка:

- если во всех локусах присутствуют общие аллели и индекс исключения родителей  $P2 > 0.99$ , то родитель подтверждается;

- если общие аллели отсутствуют в двух локусах и более, то родитель исключается;

- если общие аллели отсутствуют в одном локусе, родителю присваивается статус “сомнительный”, см. пункт 3.

2. Пункт 1 проводится для предполагаемой матери.

3. Предполагаемые родители со статусом “сомнительный” подвергаются расширенному анализу. В случае если генетический паспорт “сомнительного” родителя получен из другой лаборатории и возможность проверить первичный файл отсутствует, то статус “сомнительный” меняется на “подтвержден”. Таким образом, мы трактуем возможные опечатки и ошибки расшифровки в пользу сельхозпроизводителя. Если генотип сомнительного родителя получен в нашей лаборатории, то первичные данные проверяются, после чего статус “подтвержден” может быть присвоен в трех случаях:

- аллели потомка и “сомнительного” родителя отличаются на один мотив микросателлита (предполагаем, что имела место мутация);

- аллели потомка и “сомнительного” родителя находятся в гомозиготе (возможен нуль-аллель);

- “сомнительный” родитель – дизиготный близнец, и при проверке его генотипа отмечены низкие фрагменты, соответствующие фрагментам тестируемого потомка.

4. Если ничего из перечисленного в пункте 3 не отмечено, при этом у “сомнительного” родителя и потомка имеется информация об изменчивости “дополнительных” (не входящих в базовую панель ISAG) локусов (минимум трех), то эти дополнительные локусы добавляются в анализ. Если во всех дополнительных локусах у “сомнительного” родителя и потомка имеются общие аллели, то статус предполагаемого родителя меняется на “подтвержден”. В иных случаях статус меняется на “исключен”.

Этап II. Задача этапа – подтверждение достоверности происхождения потомка от пары родителей (у потомка и предполагаемых родителей в каком-либо локусе может быть один аллель, совпадающий с аллелями обоих родителей, в этом случае будет подтвержден каждый родитель по отдельности, однако пара родителей может быть не подтверждена).

5. Если предполагаемый отец и предполагаемая мать имеют статус “подтвержден”, то проводится сравнение генотипа потомка с генотипами пары родителей:

– если во всех локусах присутствуют общие аллели и индекс исключения родителей  $P3 > 0.9999$ , то родители подтверждаются;

– если общие аллели отсутствуют в двух локусах и более, то родители исключаются;

– если общие аллели отсутствуют в одном локусе, статус родителей – “сомнительные”. При наличии дополнительных локусов они включаются в анализ. При наличии общих аллелей в “дополнительных” локусах родители подтверждаются, при несовпадении аллелей хотя бы в одном дополнительном локусе родители исключаются.

6. Допустимо восстановление аллелей в “дополнительных” локусах быков при наличии в выборке не менее чем 10 его потомков и их матерей, имеющих общие аллели в триаде “отец – мать – потомок” во всех 12 локусах базовой панели ISAG.

Слабым местом предложенного протокола является проведение тестов дизиготных близнецов с использованием ДНК, выделенной из крови. Актуальной задачей является совершенствование протокола, позволяющего эффективно анализировать подобные образцы для проведения экспертиз на достоверность происхождения крупного рогатого скота.

Исследования выполнены в рамках Государственного задания № 0532-2023-0006.

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

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

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Методические рекомендации по исследованию и использованию групп крови в селекции крупного рогатого скота. Дубровицы.: Отдел научно-техн. информации, 1974. 40 с.
2. Лазарева Ф.Ф., Сухова Л.Г. Использование групп крови для подтверждения достоверности

происхождения крупного рогатого скота // Тр. УралНИИСХОЗА. 1983. Т. 35. С. 8–14.

3. ISAG/FAO. Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans. Measurement of Domestic Animal Diversity (MoDAD): Recommended Microsatellite Markers. FAO. Italy, Rome: 2004.
4. FAO. Molecular Genetic Characterization of Animal Genetic Resources. FAO. Italy, Rome: 2011.
5. ISAG Conference 2006. Cattle Molecular Markers and Parentage Testing Workshop. [Электронный ресурс] URL: [https://www.isag.us/Docs/ISAG2006\\_CMMPT.pdf](https://www.isag.us/Docs/ISAG2006_CMMPT.pdf) (дата обращения: 02.04.2024).
6. ISAG Conference 2008. Cattle Molecular Markers and Parentage Testing Workshop. [Электронный ресурс] URL: [https://www.isag.us/Docs/ISAG2008\\_CattleParentage.pdf](https://www.isag.us/Docs/ISAG2008_CattleParentage.pdf) (дата обращения: 02.04.2024).
7. ISAG 2021. Cattle Molecular Markers and Parentage Testing. [Электронный ресурс] URL: [https://www.isag.us/Docs/Workshop\\_report\\_CMMPT\\_2021.pdf](https://www.isag.us/Docs/Workshop_report_CMMPT_2021.pdf) (дата обращения: 02.04.2024).
8. ICAR guidelines. Section 4 – DNA technology. Version: February 2022. [Электронный ресурс] URL: <https://www.icar.org/Guidelines/04-DNA-Technology.pdf> (дата обращения 02.04.2024).
9. Решение Коллегии Евразийской экономической комиссии от 2 июня 2020 г. № 74 “Об утверждении Положения о проведении молекулярной генетической экспертизы племенной продукции государств – членов Евразийского экономического союза” [Электронный ресурс] URL: <https://www.garant.ru/products/ipo/prime/doc/74125607> (дата обращения 02.04.2024).
10. Приказ Министерства сельского хозяйства Российской Федерации от 02.06.2022 № 336 “Об утверждении требований к видам племенных хозяйств” [Электронный ресурс] URL: <http://publication.pravo.gov.ru/Document/View/0001202208300022> (дата обращения 02.04.2024).
11. Приказ Министерства здравоохранения и социального развития РФ от 12 мая 2010 г. № 346н “Об утверждении Порядка организации и производства судебно-медицинских экспертиз в государственных судебно-экспертных учреждениях Российской Федерации” [Электронный ресурс] URL: <https://base.garant.ru/12177987/> (дата обращения 02.04.2024).
12. Приказ Министерства здравоохранения Российской Федерации от 25.09.2023 № 491н “Об утверждении Порядка проведения судебно-медицинской экспертизы” [Электронный ресурс] URL: <http://publication.pravo.gov.ru/>

- docs-ment/0001202310250009 (дата обращения 02.04.2024).
13. Методические указания Минздрава РФ № 98/253 от 19.01.1999 года “Использование индивидуализирующих систем на основе полиморфизма длины амплифицированных фрагментов (ПДАФ) ДНК в судебно-медицинской экспертизе идентификации личности и установления родства” [Электронный ресурс] URL: <https://docs.cntd.ru/document/556354310> (дата обращения 02.04.2024).
  14. Животовский Л.А. Критические замечания на “Методические указания” П.Л. Иванова “Использование индивидуализирующих систем на основе полиморфизма длины амплифицированных фрагментов (ПДАФ) ДНК в судебно-медицинской экспертизе идентификации личности и установления родства” // Сиб. мед. журн.”. 2001. № 2. С. 85–86.
  15. Пименов М.Г., Кульгин А.Ю., Кондрашов С.А. Научные и практические аспекты криминалистического ДНК-анализа: Учебное пособие. М.: ГУ ЭКЦ МВД России, 2001. 144 с.
  16. ISAG. Exclusion probability in parentage tests [Электронный ресурс] URL: [https://www.isag.us/Docs/consignment-forms/Exclusion\\_probability.pdf](https://www.isag.us/Docs/consignment-forms/Exclusion_probability.pdf) (дата обращения 02.04.2024).
  17. Jamieson A., Taylor St.C.S. Comparisons of three probability formulae for parentage exclusion // Animal Genetics. 1997. V. 28. P. 397–400.  
doi: 10.1111/j.1365-2052.1997.00186.x
  18. Хедрик Ф. Генетика популяций. М.: Техносфера, 2003. 592 с. (Hedrick P.W. Genetics of population. Jones and Barlet Publ., Inc., 1999.)
  19. Peakall R., Smouse P.E. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research // Mol. Ecol. Notes. 2006. V. 6. № 1. P. 288–295.  
<https://doi.org/10.1111/j.1471-8286.2005.01155.x>
  20. Модоров М.В., Ткаченко И.В., Грин А.А. и др. Генетическая структура популяции голштинизированного черно-пестрого скота на территории Урала // Генетика. 2021. Т. 57. № 4. С. 437–444.  
<https://doi.org/10.1134/S1022795421040104>
  21. Van de Goor L.H.P., Koskinen M.T., van Haeringen W.A. Population studies of 16 bovine STR loci for forensic purposes // Int. J. Legal Med. 2011. V. 125. P. 111–119.  
doi: 10.1007/s00414-009-0353-8
  22. Steely C.J., Scot Watkins W., Baird L., Jorde L.B. The mutational dynamics of short tandem repeats in large, multigenerational families // Genome Biology. 2022. V. 23. Article number: 253.  
<https://doi.org/10.1186/s13059-022-02818-4>
  23. CoRDIS Cattle. Набор реагентов для мультиплексного анализа 15-ти микросателлитных маркеров крупного рогатого скота [Электронный ресурс] URL: <https://gordiz.ru/wp-content/uploads/2023/12/instrukcziya-cordis-cattle-231205.pdf> (дата обращения 02.04.2024).
  24. Набор “Gene Profile Cattle” для генетической паспортизации и определения родства крупного рогатого скота [Электронный ресурс] URL: <https://www.syntol.ru/catalog/reagents-dlya-geneticheskikh-analizatorov/geneprofile-cattle.html> (дата обращения 02.04.2024).
  25. Kozubská-Sobocińska A., Smołucha G., Danielak-Czech B. Early diagnostics of freemartinism in Polish Holstein-Friesian female calves // Animals. 2019. V. 9. № 11. Article number: 971.  
doi: 10.3390/ani9110971
  26. Модоров М.В., Файрушина К.Р., Клещева А.А. и др. Воспроизводство племенного голштинизированного черно-пестрого скота в Свердловской области // Достижения науки и техники АПК. 2022. Т. 36. № 8. С. 67–71.  
doi: 10.53859/02352451\_2022\_36\_8\_67

## Markers Nomenclature and Type I and II Errors Control in Cattle Parentage Exclusion Expertise with Microsatellite Markers

M. V. Modorov<sup>1</sup>, \* , I. V. Tkachenko<sup>1</sup>, A. A. Kleshcheva<sup>1</sup>, M. Yu. Sevost'yanov<sup>1</sup>

<sup>1</sup>*Ural Federal Agrarian Scientific Research Center, Ural Branch, Russian Academy of Sciences,  
Yekaterinburg, 620142 Russia  
\*e-mail: mmodorov@gmail.com*

Erroneous pedigree records reduce the quality of breeding work with cattle; therefore, parentage exclusion expertise has become an integral part of breeding work. For many years on the territory of the Russian Federation, the expertise was carried out using immunogenetic markers, but the improvement of technologies and tightening of regulatory requirements has led to the process of replacement of immunogenetics by microsatellite markers. At present there is no domestic protocol clearly regulating the procedure of cattle parentage exclusion expertise using microsatellite loci, which makes laboratory work more difficult. In particular, the requirements for the number and nomenclature of genetic markers which are used in the genetic expertise are only available for cattle, embryos and semen products being transported across the Eurasian Economic Union. There are no regulations for errors I (false positives) and II (false negatives) types control, which must be taken into account when forming expert judgements. In this paper we will review the approaches to addressing these issues proposed by the International Society for Animal Genetics (ISAG), the International Committee for Animal Recording (ICAR), the Collegium of the Eurasian Economic Commission, as well as domestic regulatory documents governing the production of forensic medical tests related to parentage exclusion expertise. Based on the results of the review, a nomenclature of microsatellite markers and a parentage exclusion protocol will be proposed, in which errors I and II type control are carried out. Special attention will be paid to the description of the sources of II type error and the necessity of its control.

**Keywords:** *Bos taururs*, microsatellites, selection, ISAG, STR, parentage exclusion.

---

## ОБЗОРНЫЕ И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

---

УДК 571.27

# ИНДУКЦИЯ ТРЕНИРОВАННОГО ИММУНИТЕТА С ПОМОЩЬЮ БЦЖ: НЕДАВНИЕ ДАННЫЕ И МНЕНИЯ

© 2024 г. И. В. Алексеенко<sup>1, 2</sup>, Л. Г. Кондратьева<sup>1, 2, \*</sup>, И. П. Чернов<sup>2</sup>, Е. Д. Свердлов<sup>1, \*\*</sup>

<sup>1</sup>Научно-исследовательский центр “Курчатовский институт”, Москва, 123182 Россия

<sup>2</sup>Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова

Российской академии наук, Москва, 117997 Россия

\*e-mail: liakondratyeva@yandex.ru

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

Поступила в редакцию: 28.05.2024 г.

После доработки: 02.07.2024 г.

Принята к публикации: 01.08.2024 г.

Разработанная в начале XX в. для защиты от туберкулеза бацилла Кальметта–Герена (БЦЖ, *Bacillus Calmette–Guerin*) – одна из наиболее широко используемых вакцин в мире. Оказалось, что эта вакцина, вдобавок к защите от туберкулеза, снижает также риск различных других инфекций. В процессе борьбы с коронавирусной инфекцией COVID-19 выяснилось, что вакцинированные БЦЖ медицинские работники имели более низкие показатели выявления вируса SARS-CoV-2 по сравнению с не вакцинированными людьми. Это происходит, по-видимому, благодаря формированию тренированного иммунитета (ТРИМ) и неспецифического защитного эффекта (НЗЭ, non-specific effects), которые, скорее всего, объясняются эпигенетическим и метаболическим перепрограммированием клеток врожденного иммунитета, называемым тренированным иммунитетом. НЗЭ открывает перспективы разработки будущих терапевтических средств, которые будут защищать от непредвиденных пандемий. Кроме того, индукторы ТРИМ открывают новые перспективы в повышении эффективности стандартных вакцин путем их включения в составы вакцин для усиления как специфических, так и неспецифических иммунных ответов.

**Ключевые слова:** вакцины, БЦЖ, тренированный иммунитет, врожденная иммунная система, эпигенетические модификации.

**DOI:** 10.31857/S0016675824120022 **EDN:** WAVTHU

## ГОРЬКИЙ ОПЫТ ПАНДЕМИИ COVID-19. ВРЕМЯ МЕЖДУ ПОЯВЛЕНИЕМ SARS-COV-2 И ЭФФЕКТИВНОЙ ВАКЦИНОЙ

Несмотря на принятие многими странами решительных мер, пандемию COVID-19 не удалось сдержать, и осенью – зимой 2020 г. многие страны Северного полушария пережили вторую волну инфекции. Клинический спектр COVID-19 неоднороден: около 40% инфицированных людей не проявляли или почти не проявляли симптомов, тогда как у меньшинства инфицированных развивалась тяжелая пневмония с дыхательной недостаточностью, вплоть до летального исхода [1]. Полагают [2, 3], что инфекция SARS-CoV-2 станет эндемичной, и в дальнейшем вспышки будут происходить, например, в холодное время года. В последнее время, кроме того, сообщалось об эффективности ряда вакцин против COVID-19 (см., например, недавний метаанализ [4]). Тем не менее за

время, прошедшее между появлением вируса и эффективной вакцинацией во всем мире, COVID-19 причинил страдания и смерть миллионам людей, а также привел к катастрофическим экономическим последствиям. Вакцинация остается самой реалистичной надеждой обуздать распространение вируса.

Важнейшим уроком, извлеченным из истории с COVID-19, является понимание, что при появлении новой пандемии потребуется несколько месяцев, прежде чем можно будет найти эффективный способ вакцинации, произвести достаточное количество доз и распространить среди населения развитых стран, и гораздо больше времени, чтобы вакцинировать все население мира. Время между появлением вируса и эффективной всемирной вакцинацией будет периодом страданий и смертей миллионов людей, а также катастрофических экономических последствий [5]. Поэтому остро стоит

вопрос о неспецифической к конкретному патогену профилактической защите. В процессе пандемии COVID-19 были предложены альтернативные подходы к профилактике заболевания. И, естественно, возникла идея использовать известные свойства некоторых старых вакцин вызывать защиту от инфекций, отличающихся от заболевания, для защиты от которого они были созданы [6]. Вакцины, содержащие живые аттенуированные микроорганизмы, в частности, бацилла Кальметта — Герена (**БЦЖ**, *Bacillus Calmette–Guerin*), вакцины против кори, паротита и краснухи (**КПК**, *Measles, mumps and rubella*, MMR) и оральная вакцина против полiovируса (**ОПВ**, oral polio vaccine, OVP) по ряду наблюдений способны снижать смертность среди детей, помимо защиты от целевых заболеваний [7, 8]. Этот эффект получил название неспецифического защитного эффекта (**НЗЭ**, non-specific effects) [9].

### ЗАЩИТА, ИНДУЦИРОВАННАЯ БЦЖ

Эпидемиологические обследования показали, что вакцинация БЦЖ значительно снижает младенческую смертность, что не может быть объяснено только снижением заболеваемости туберкулезом [9]. БЦЖ-индуцированная защита, по-видимому, обусловлена эффектами, в частности, против инфекций дыхательных путей [10], что также позже было подтверждено в рандомизированных исследованиях у взрослых [11, 12]. Эти клинические испытания были дополнены экспериментальными исследованиями, в которых изучались механизмы, с помощью которых БЦЖ вызывает эти защитные эффекты. Были получены данные, свидетельствующие о том, что в эту защиту вовлечен врожденный иммунитет и соответствующие клетки [13, 14]. Оказалось, что функциональные изменения моноцитов связаны с транскрипционным, эпигенетическим и метаболическим перепрограммированием предшественников миелоидных клеток у лиц, привитых БЦЖ [15]. Стало ясным, что долгосрочные изменения в фенотипе клеток врожденного иммунитета после вакцинации БЦЖ представляют собой то, что получило название “тренированный иммунитет” (**ТРИМ**). Имеются данные о том, что ТРИМ является (по крайней мере частично) механизмом, с помощью которого вакцинация БЦЖ вызывает защиту против вирусных инфекций [16]. В результате была выдвинута гипотеза о том, что вакцинация БЦЖ может использоваться в качестве профилактической меры против инфекции SARS-CoV-2.

Несколько эпидемиологических исследований давали основание считать, что в странах с программой вакцинации БЦЖ наблюдается более низкое число случаев заражения COVID-19 и более низкий уровень смертности [17–19]. Однако демографические, этнические и генетические различия

между населением в разных странах, а также ряд других особенностей [20], различающих разные популяции, вызывают систематическую ошибку в таких анализы.

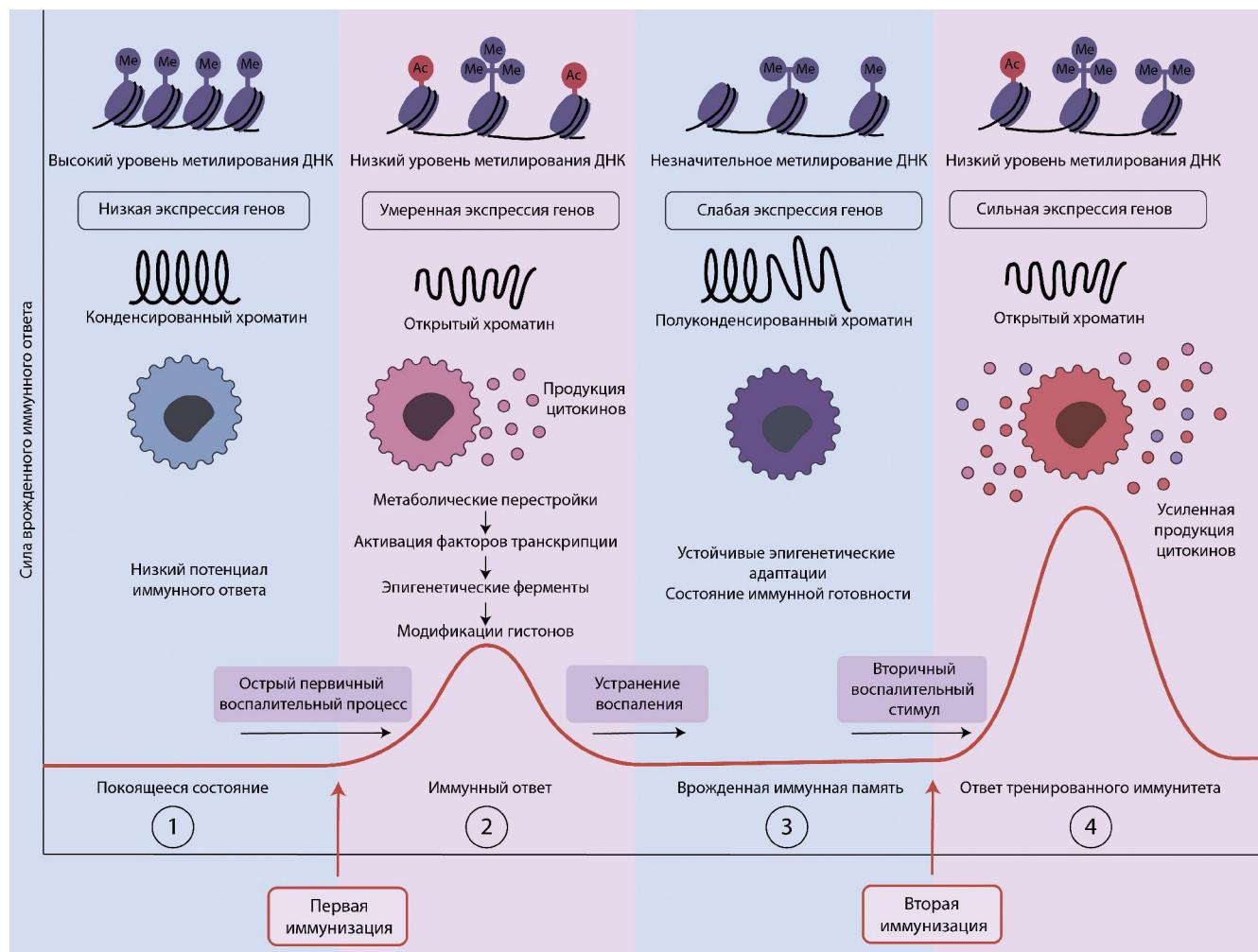
Гипотеза о профилактике COVID-19 с помощью БЦЖ получила поддержку благодаря исследованию [17]. Была исследована заболеваемость, вызванная инфекциями SARS-CoV-2 в большой когорте медицинских работников из системы здравоохранения Лос-Анджелеса. Почти треть добровольцев, участвовавших в исследовании, ранее прошли вакцинацию БЦЖ, которая сопровождалась заметно меньшим количеством (30–40%) диагнозов COVID-19, положительных результатов ПЦР-тестов с обратной транскрипцией COVID-19 и серопозитивных тестов на анти-SARS-CoV-2. Было показано, что вакцинация БЦЖ связана с меньшей заболеваемостью COVID-19, несмотря на то, что вакцинированная группа была немного старше и имела больше сопутствующих заболеваний. Симптомы, связанные с COVID-19, также были заметно менее тяжелыми в группе, вакцинированной БЦЖ, чем в невакцинированной. Интересно, что внутренний контроль показал, что такие положительные эффекты не были вызваны другими вакцинами, такими как вакцины против гриппа, менингококка или пневмококка, поэтому эффект, по-видимому, относится к иммунизации БЦЖ. Следует, однако, сказать, что число людей, которые не были вакцинированы против гриппа, было слишком малым, чтобы делать однозначные выводы [17]. Недавнее обследование лиц, привитых БЦЖ в течение последних двух лет [21], и одно клиническое исследование медицинского персонала, получившего вакцинацию БЦЖ во время первой фазы пандемии [22], согласуются с представлением о защитном действии БЦЖ против инфекции SARS-CoV-2.

Появившиеся в конечном счете вакцины на основе РНК от BioNTech-Pfizer и Moderna и российская аденоовирусная вакцина “Гам-КовидВак” обеспечивают более 90% защиты от COVID-19, в том числе у пожилых людей [4, 23, 24]. Третья вакцина, основанная на аденоовирусной платформе от Oxford-AstraZeneca, так же, по-видимому, обеспечивает защиту не менее 60% [4, 23]. Однако период перед их появлением является предостережением и аргументом в пользу необходимости продолжения исследований БЦЖ и других вакцин, вызывающих неспецифический защитный эффект. БЦЖ и, возможно, другие живые аттенуированные вакцины могли бы стать промежуточной вакцинацией в период между появлением нового патогена и разработкой конкретной вакцины, чтобы уменьшить страдания населения как от самой болезни, так и от ее экономических последствий.

## РЕГУЛЯТОРНЫЕ МЕХАНИЗМЫ, УПРАВЛЯЮЩИЕ ТРЕНИРОВАННЫМ ИММУНИТЕТОМ

Человека от инфекций защищают две ветви иммунной системы: врожденный и адаптивный иммунитет (для недавнего обзора см. [25]). Первый обеспечивает постоянную защиту от патогенов, в то время как второй запоминает характеристики индивидуальных патогенов, с которыми человек уже сталкивался. Традиционные вакцины нацелены на индукцию адаптивной иммунной системы антигенами определенных патогенов и требуют специальной вакцины для каждого нового патогена. Врожденная иммунная система, в свою

очередь, играет ключевую роль в первичной обороне организма, пока активируется специфический (адаптивный) иммунный ответ. Она представляет собой набор механизмов, которые реагируют быстро и неспецифически на угрозы. Долгосрочные изменения в фенотипе клеток врожденного иммунитета могут привести к формированию ТРИМ. При первоначальном воздействии патогена или вакцины, индуцирующих ТРИМ (рис. 1), клетки врожденного иммунитета подвергаются метаболической перестройке и эпигенетическому пере програмированию, приводя к усиленному и ускоренному ответу на вторичную гетерологическую стимуляцию.



**Рис. 1.** Тренированный иммунитет опосредован метаболической и эпигенетической перестройкой в клетках врожденного иммунитета, что приводит к увеличению транскрипции генов и защите хозяина от гетерологических патогенов. После первого контакта с живым или живым ослабленным патогеном врожденный иммунный ответ приходит в “тренированное” состояние, готовое к усилению после второй инфекции. Наивная, нестимулированная клетка (1) демонстрирует сильно конденсированный хроматин. Ее первичная стимуляция приводит к эпигенетическим модификациям хроматина (таким как ацетилирование гистонов (Ac) или метилирование (Me)), которые способствуют разворачиванию хроматина и облегчают транскрипцию генов (2). Эти эпигенетические изменения могут сохраняться или могут быть устранины лишь частично после прекращения действия раздражителя (3) и, таким образом, могут генерировать устойчивую и запоминаемую врожденную иммунную память, реализуемую в ответ на вторичный воспалительный стимул (4) (по мотивам [5, 26, 27]).

Эпигенетическое перепрограммирование БЦЖ включает модификации гистонов, в частности приобретение активирующих меток H3K27ac и H3K4me3, снижение уровня метилирования гистонов (а именно репрессивной метки H3K9me3 на промоторах генов провоспалительных цитокинов, таких как IL6, TNFA и IL1B). Ацетилирование гистонов приводит к раскрытию хроматина. После инициирования в костном мозге клетки сохраняют метки ацетилирования на протяжении всей дифференцировки в лимфоидную или миелоидную линию, способствуя миелопоэзу [28].

После первого контакта с БЦЖ клетки претерпевают долгосрочные метаболические и эпигенетические изменения, приводящие к усилению иммунных реакций на последующие не связанные с ними стимулы. Изменения в уровнях метаболитов определяют активность ферментов модификации гистонов. Важные изменения происходят в системах гликолиза, окислительного фосфорилирования (OXPHOS), цикла TCA и липидного обмена. Индукция тренированного иммунитета также определяется на уровне метилирования ДНК в клетках врожденного иммунитета. Показано, что у вакцинированных БЦЖ лиц, у которых успешно развился ТРИМ, наблюдалась значительная потеря метилирования ДНК в промоторах генов, связанных с иммунитетом, по сравнению с невосприимчивыми [29].

Метаболические и эпигенетические процессы, участвующие в возникновении ТРИМ, тесно взаимосвязаны. Метаболическая перестройка поддерживает потребности в энергии и промежуточных метаболитах для эпигенетических и геномных изменений и напрямую модулирует активность эпигенетических ферментов.

ТРИМ, индуцированный БЦЖ, непосредственно влияет и на адаптивный иммунный ответ: тренированные моноциты продукции значительно больше провоспалительных медиаторов, таких как TNF, IL-1 $\beta$  и IL-6, а также проявляют повышенную экспрессию МНС-II, костимулирующих молекул (CD80/CD86) и повышают уровень аэробного гликолиза [30, 31]. Вследствие этого ТРИМ изменяет пластичность Т-клеток и может смещать баланс между Th1 и Th2 в сторону формирования цитотоксического Th1-ответа и между Treg и Th17 в сторону активации воспалительных реакций через Th17. Усиленная продукция IFN- $\gamma$  Th1-клетками и IL-17 Th17-клетками может взаимно влиять на функцию клеток врожденного иммунитета. Было показано, что выработка IFN- $\gamma$  индуцирует ТРИМ в тканевых резидентных макрофагах [31]. В исследовании [32] было обнаружено, что применение вакцины БЦЖ предотвращает антителозависимое усиление инфекции после вакцинации против респираторного синцитиального вируса (RSV, respiratory syncytial virus) за счет формируемого ею

ТРИМ и индукции изменений в Т-клетках. Авторы полагают, что БЦЖ, вводимая вскоре после рождения с последующей вакцинацией от RSV, может индуцировать тренированные макрофаги, тканевые резидентные Т-клетки памяти и специфические цитотоксические Т-лимфоциты. При этом использование только вакцины от RSV способствует развитию толерантности макрофагов, подавляя их эффекторные функции при вторичной стимуляции патогеном, а комбинация с БЦЖ предотвращает ее за счет формирования тренированных макрофагов, способных быстро реагировать на последующий стимул [32]. В другой работе [33] *in vitro* была показана обратимость иммунотолерантности, возникающей вследствие вакцинации от брюшного тифа (TVF, typhoid fever vaccine) с помощью механизмов ТРИМ, вызванных БЦЖ. Авторы показали, что БЦЖ не модулирует адаптивный иммунитет при TVF, но при этом частично предотвращает ингибирование врожденных иммунных реакций, вызванных TVF [33].

Использование БЦЖ в качестве предварительного бустера перед вакцинацией бесклеточной противококлюшной вакциной на выборке из здоровых добровольцев в рандомизированном клиническом исследовании [34] показало, что ТРИМ, индуцированный БЦЖ, усиливает продукцию противококлюшных специфических антител, противококлюшные-специфические реакции Th1-клеток и общую реакцию В-клеток памяти через 2 недели после иммунизации. Это наблюдение положительно коррелировало с повышением уровня продукции биомаркеров ТРИМ, таких как IL-1 $\beta$  и IL-6. Таким образом, БЦЖ вызывает устойчивые изменения в иммунной системе, связанные с неспецифическим ответом на инфекции, как на уровне врожденного тренированного иммунитета, так и на уровне гетерологичных реакций адаптивного иммунитета [35].

Недавнее фундаментальное исследование, проведенное большим коллективом авторов [36], дает следующую сводку разнообразных защитных эффектов, наблюдавшихся при вакцинации БЦЖ:

1. При применении в виде однократной дозы БЦЖ усиливает анти микробактериальную реакцию организма, защищающую от туберкулеза, хотя и с различной эффективностью в зависимости от человека и географического региона.
2. БЦЖ может обеспечить состояние повышенной реактивности врожденной иммунной системы и может увеличить общую выживаемость детей в развивающихся странах, хотя этот результат не был последовательно воспроизведен.
3. БЦЖ может обеспечить частичную защиту от респираторных инфекций: SARS-CoV-2, гриппа, RSV (респираторно-синцитиальный вирус),

*Staphylococcus aureus*, *Escherichia coli* и *Streptococcus pneumoniae*, малярии и некоторых вирусов [37–39]. Это направление представляется особенно важным, поскольку именно респираторные инфекции являются одной из наиболее вероятных причин будущих пандемий [40, 41]. Известно, в частности, что эпидемические пики совместно циркулирующих респираторных вирусов в зимне-весенние сезоны могут быть синхронными и асинхронными [42]. Наличие временных закономерностей в эпидемиях, вызванных некоторыми респираторными вирусами, позволяет предположить, что они могут негативно взаимодействовать друг с другом. Это может быть результатом влияния врожденной иммунной памяти. Во время пандемии COVID-19 было снижено распространение большинства респираторных вирусов, а после отмены противоковидных санитарных мер циркуляция возобновилась в необычные периоды [42]. Повышение врожденного иммунитета может иметь важное значение для разработки новых и эффективных противопаразитарных вакцин [43].

Помимо защитных свойств от инфекционных процессов, БЦЖ показывает положительные результаты при применении для терапии некоторых онкологических заболеваний, таких как рак мочевого пузыря [44], рак желудка, простаты, легких [45, 46]. В последние четыре десятилетия БЦЖ активно используется для лечения немышечного инвазивного рака мочевого пузыря. За этот период применения данного метода лечения было установлено, что он не только эффективен, но и имеет преимущества перед внутрипузырной химиотерапией. В отличие от нее, данный метод существенно снижает не только вероятность рецидива, но и риск его прогрессирования, что подтверждает его значительную эффективность в долгосрочной перспективе [47]. В своем применении для лечения рака мочевого пузыря БЦЖ является наиболее успешным на сегодняшний день примером микробиологического лечения рака [48].

Хотя многочисленные обзоры по врожденной иммунной памяти подробно описывают механизм тренированного иммунитета, их индукторы и роль эффекторных иммунных клеток, в обзоре [36] впервые рассматриваются новые экспериментальные модели тренированного иммунитета и последние достижения, которые позволяют удовлетворить повышающийся спрос на широкий спектр защитных средств от респираторных патогенов. В этом контексте в обзоре подчеркивается интеграция обученных индукторов иммунитета с антигенами для придания свойств врожденной и адаптивной иммунной памяти для создания вакцин следующего поколения. Этот подход к гибридной вакцинации может быть использован для удовлетворения потребности в разработке быстрых и эффективных стратегий иммунизации и обладает огромным

потенциалом для изменения парадигмы профилактики против возникающих пандемий [37].

Чтобы исследовать эту проблему детально, упомянутая группа авторов [36] вакцинировала БЦЖ 323 здоровых добровольца и проанализировала влияние на иммунную систему. Основываясь на увеличении неспецифической продукции цитокинов клетками иммунной системы при их активации микробными стимулами на 90 день после вакцинации БЦЖ, испытуемых разделили на две группы: 213 человек были определены как способные к обучению врожденного иммунитета (responders, “респондеры”), а 78 – как неспособные (nonresponders, “нереспондеры”). Авторы обнаружили, что индукция тренированного иммунитета была наиболее эффективной у людей с дремлющим (dormant) врожденным иммунитетом на момент иммунизации, что отражалось в характерном состоянии эпигенетических клеток, предсказывающем реакцию на вакцину. Оказалось, что у “нереспондеров” уже была более высокая исходная иммунная готовность, чем у “респондеров”.

Клетки нативной иммунной системы у тех людей, которые реагируют на вакцинацию БЦЖ, переключаются от пассивного, покоящегося (dormant) состояния к активному состоянию [36, 49]. У тех же, которые не реагируют, эти клетки и до вакцинации находятся в хотя бы частично активированном состоянии. Испытуемые с конденсированным до вакцинации хроматином в генах, участвующих во врожденном иммунитете, демонстрировали сильный тренированный иммунитет, индуцированный БЦЖ, и приобретали открытый хроматин в этих генах, в то время как “нереспондеры” несли такой открытый хроматин до вакцинации БЦЖ без дальнейшего усиления иммунных реакций после вакцинации. Это открытие объясняет, как эпигенетика позволяет иммунным клеткам переключаться между различными уровнями иммунной готовности для эффективной защиты организма от патогенов, минимизируя при этом возможность возникновения излишних и патологических иммунных ответов.

Конечно, нужно иметь в виду, что индивидуумы крайне вариабельны, и нельзя точно предсказать, какой эффект может оказать индукция врожденной иммунной системы на данного конкретного человека. Этой изменчивости способствовали как генетические факторы, так и факторы окружающей среды, но наиболее интересные различия наблюдались в эпигенетических состояниях иммунных клеток. Эпигенетические состояния клеток, реализуемые посредством изменений в доступности хроматина, которые облегчают или затрудняют активацию генов, отражают регуляторную пластичность клетки и ее способность быстро реагировать на изменения в окружающей среде, что делает их

вероятными кандидатами на регулирование тренированного иммунитета.

Результаты авторов [36], по их мнению, показывают, что люди могут быть сильными “респондерами” на врожденный иммунный ответ, но слабыми на адаптивный ответ на вакцинацию БЦЖ, и наоборот. Другими словами, врожденная и адаптивная иммунная память представляется в значительной степени независимыми биологическими процессами, каждый из которых управляет отдельными регуляторными механизмами. Эти фенотипические эффекты связаны с эпигенетическим состоянием клеток (рис. 1), связанным с доступностью хроматина и определяющим активацию генов, приводящую в конечном счете к быстрой реакции клеток на изменения в окружающей среде.

Эти результаты дают не только новое представление о возможном новом поколении универсальных вакцин, но и могут составить основу будущих новых терапевтических средств, которые целенаправленно пробуждают дремлющую иммунную систему [50, 51]. В частности, можно думать о реактивации подавленной иммунной системы у пациентов, больных раком. Несколько фармацевтических компаний уже ищут способы вызвать тренированный иммунитет, не полагаясь на вакцину БЦЖ [8, 52].

### ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Включение модуляторов ТРИМ в современные вакциные технологии имеет значительные последствия для разработки вакцин. Используя врожденную иммунную память в составах вакцин следующего поколения, мы потенциально можем повысить эффективность вакцин против различных инфекционных заболеваний, что приведет к улучшению контроля заболеваний и их профилактики.

Эффективность живых вакцин может быть значительно усиlena с помощью генетических технологий. Такие попытки активно предпринимаются [53–57]. Будущие исследовательские усилия должны быть сосредоточены на оптимизации интеграции индукторов ТРИМ в составы вакцин, понимании активируемых путей и механизмов их действия, а также оценке их долгосрочного воздействия на иммунные реакции. Было показано, что ТРИМ стимулируется в костном мозге и периферических тканях. Однако различия в индукции ТРИМ на разных уровнях, таких как роль тренированных периферических, тканевых резидентов или клеток-предшественников, требуют детально-го исследования.

Важно иметь в виду, что помимо полезных эффектов, ТРИМ может приводить к гиперактивации иммунных клеток. Для изучения побочных эффектов, связанных с индукторами ТРИМ, необходима

щательная доклиническая и клиническая оценка, чтобы избежать потенциальной иммунопатологии или вмешательства в специфическую иммунную память. На данном этапе совместные усилия иммунологов, вакцинологов и клиницистов необходимы для разработки и проведения хорошо контролируемых клинических испытаний для оценки эффективности и безопасности вакцин на основе ТРИМ. Кроме того, изучение комбинаторных подходов, таких как стратегии, сочетающие индукторы ТРИМ со специфическими антигенами, могло бы еще больше повысить эффективность вакцины.

Работа выполнена за счет гранта Российского научного фонда № 22-14-00308, <https://rscf.ru/project/22-14-00308/>.

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. Huang C., Wang Y., Li X. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China // Lancet. 2020. V. 395. № 10223. P. 497–506. doi: 10.1016/S0140-6736(20)30183-5
2. Shaman J., Galanti M. Will SARS-CoV-2 become endemic? // Science. 2020. V. 370. № 6516. P. 527–529. doi:10.1126/science.abe5960
3. Otto S.P., MacPherson A., Colijn C. Endemic does not mean constant as SARS-CoV-2 continues to evolve // Evolution. 2024. V. 10. doi: 10.1093/evolut/qpa041
4. Soheili M., Khateri S., Moradpour F. et al. The efficacy and effectiveness of COVID-19 vaccines around the world: A mini-review and meta-analysis // Ann. Clin. Microbiology and Antimicrobials. 2023. V. 22. № 1. P. 42. doi: 10.1186/s12941-023-00594-y
5. Netea M.G., van der Meer J.W., van Crevel R. BCG vaccination in health care providers and the protection against COVID-19 // J. Clin. Invest. 2021. V. 131. № 2. doi: 10.1172/JCI145545
6. Aaby P., Benn C.S., Flanagan K.L. et al. The non-specific and sex-differential effects of vaccines // Nat. Rev. Immunol. 2020. V. 20. № 8. P. 464–470. doi: 10.1038/s41577-020-0338-x
7. Higgins J.P., Soares-Weiser K., Lopez-Lopez J.A. et al. Association of BCG, DTP, and measles containing vaccines with childhood mortality: Systematic review // BMJ. 2016. V. 355. doi: 10.1136/bmj.i5170

8. Baydemir I., Dulfer E.A., Netea M.G., Dominguez-Andres J. Trained immunity-inducing vaccines: Harnessing innate memory for vaccine design and delivery // Clin. Immunol. 2024. V. 261. doi: 10.1016/j.clim.2024.109930
9. Shann F. The non-specific effects of vaccines // Arch. Dis. Child. 2010. V. 95. № 9. P. 662–667. doi: 10.1136/adc.2009.157537
10. Stensballe L.G., Nante E., Jensen I.P. et al. Acute lower respiratory tract infections and respiratory syncytial virus in infants in Guinea-Bissau: A beneficial effect of BCG vaccination for girls community based case-control study // Vaccine. 2005. V. 23. № 10. P. 1251–1257. doi: 10.1016/j.vaccine.2004.09.006
11. Nemes E., Geldenhuys H., Rozot V. et al. Prevention of *M. tuberculosis* infection with H4:IC31 vaccine or BCG revaccination // N. Engl. J. Med. 2018. V. 379. № 2. P. 138–149. doi: 10.1056/NEJMoa1714021
12. Giamarellos-Bourboulis E.J., Tsilika M., Moorlag S. et al. Activate: Randomized clinical trial of BCG vaccination against infection in the elderly // Cell. 2020. V. 183. № 2. P. 315–323. doi: 10.1016/j.cell.2020.08.051
13. Spencer J.C., Ganguly R., Waldman R.H. Nonspecific protection of mice against influenza virus infection by local or systemic immunization with Bacille Calmette-Guerin // J. Infect. Dis. 1977. V. 136. № 2. P. 171–175. doi: 10.1093/infdis/136.2.171
14. Kleinnijenhuis J., Quintin J., Preijers F. et al. Bacille Calmette-Guerin induces NOD2-dependent non-specific protection from reinfection via epigenetic reprogramming of monocytes // Proc. Natl Acad. Sci. USA. 2012. V. 109. № 43. P. 17537–17542. doi: 10.1073/pnas.1202870109
15. Cirovic B., de Bree L.C.J., Groh L. et al. BCG vaccination in humans elicits trained immunity via the hematopoietic progenitor compartment // Cell Host Microbe. 2020. V. 28. № 2. P. 322–334. doi: 10.1016/j.chom.2020.05.014
16. Arts R.J.W., Moorlag S., Novakovic B. et al. BCG vaccination protects against experimental viral infection in humans through the induction of cytokines associated with trained immunity // Cell Host Microbe. 2018. V. 23. № 1. P. 89–100. doi: 10.1016/j.chom.2017.12.010
17. Rivas M.N., Ebinger J.E., Wu M. et al. BCG vaccination history associates with decreased SARS-CoV-2 seroprevalence across a diverse cohort of health care workers // J. Clin. Invest. 2021. V. 131. № 2. doi: 10.1172/JCI145157
18. Escobar L.E., Molina-Cruz A., Barillas-Mury C. BCG vaccine protection from severe coronavirus disease 2019 (COVID-19) // Proc. Natl Acad. Sci. USA. 2020. V. 117. № 30. P. 17720–17726. doi: 10.1073/pnas.2008410117
19. Berg M.K., Yu Q., Salvador C.E. et al. Mandated Bacillus Calmette-Guerin (BCG) vaccination predicts flattened curves for the spread of COVID-19 // Sci. Adv. 2020. V. 6. № 32. doi: 10.1126/sciadv.abc1463
20. Hamiel U., Kozer E., Youngster I. SARS-CoV-2 rates in BCG-vaccinated and unvaccinated young adults // JAMA. 2020. V. 323. № 22. P. 2340–2341. doi: 10.1001/jama.2020.8189
21. Moorlag S., van Deuren R.C., van Werkhoven C.H. et al. Safety and COVID-19 symptoms in individuals recently vaccinated with BCG: A retrospective cohort study // Cell. Rep. Med. 2020. V. 1. № 5. doi: 10.1016/j.xcrm.2020.100073
22. Amirlak L., Haddad R., Hardy J.D. et al. Effectiveness of booster BCG vaccination in preventing Covid-19 infection // Hum. Vaccin. Immunother. 2021. V. 17. № 11. P. 3913–3915. doi: 10.1080/21645515.2021.1956228
23. Pascolo S. Vaccines against COVID-19: Priority to mRNA-based formulations // Cells. 2021. V. 10. № 10. doi: 10.3390/cells10102716
24. Sukhikh G.T., Priputnevich T.V., Ogarkova D.A. et al. Sputnik light and sputnik v vaccination is effective at protecting medical personnel from COVID-19 during the period of delta variant dominance // Vaccines. 2022. V. 10. № 11. doi: 10.3390/vaccines10111804
25. Алексеенко И.В., Василов Р.Г., Кондратьева Л.Г. и др. Клеточные и эпигенетические аспекты программирования тренированного иммунитета и перспективы создания универсальных вакцин в преддверии учащихся пандемий // Генетика. 2023. Т. 59. № 9. С. 981–1001. doi: 10.31857/S0016675823090023 (Alekseenko I., Vasilev R., Kondratyeva L. et al. The Cellular and Epigenetic Aspects of Trained Immunity and Prospects for Creation of Universal Vaccines on the Eve of More Frequent Pandemics // Russian Journal of Genetics. 2023. V. 59. № 9. P. 851–868.)
26. Munoz-Wolf N., Lavelle E.C. Promotion of trained innate immunity by nanoparticles // Semin. Immunol. 2021. V. 56. doi: 10.1016/j.smim.2021.101542
27. Hajishengallis G., Netea M.G., Chavakis T. Innate immune memory, trained immunity and nomenclature clarification // Nat. Immunol. 2023. V. 24. № 9. P. 1393–1394. doi: 10.1038/s41590-023-01595-x
28. Kaufmann E., Sanz J., Dunn J.L. et al. BCG educates hematopoietic stem cells to generate protective innate immunity against tuberculosis // Cell. 2018. V. 172. № 1–2. P. 176–190. doi: 10.1016/j.cell.2017.12.031
29. Verma D., Parasa V.R., Raffetseder J. et al. Anti-microbial activity correlates with altered DNA methylation pattern in immune cells from BCG-vaccinated

- subjects // Sci. Rep. 2017. V. 7. № 1. P. 12305.  
doi: 10.1038/s41598-017-12110-2
30. Singh S., Saavedra-Avila N.A., Tiwari S., Porcelli S.A. A century of BCG vaccination: Immune mechanisms, animal models, non-traditional routes and implications for COVID-19 // Front. Immunol. 2022. V. 13. doi: 10.3389/fimmu.2022.959656
  31. Murphy D.M., Mills K.H.G., Basdeo S.A. The effects of trained innate immunity on T cell responses; clinical implications and knowledge gaps for future research // Front. Immunol. 2021. V. 12. doi: 10.3389/fimmu.2021.706583
  32. Wang Y., Ge F., Wang J. et al. Mycobacterium bovis BCG given at birth followed by inactivated respiratory syncytial virus vaccine prevents vaccine-enhanced disease by promoting trained macrophages and resident memory T cells // J. Virol. 2023. V. 97. № 3. doi: 10.1128/jvi.01764-22
  33. Blok B.A., Arts R.J.W., van Crevel R. et al. Differential effects of BCG vaccine on immune responses induced by vi polysaccharide typhoid fever vaccination: an explorative randomized trial // Eur. J. Clin. Microbiol. Infect. Dis. 2020. V. 39. № 6. P. 1177–1184. doi: 10.1007/s10096-020-03813-y
  34. Gillard J., Blok B.A., Garza D.R. et al. BCG-induced trained immunity enhances acellular pertussis vaccination responses in an explorative randomized clinical trial // NPJ Vaccines. 2022. V. 7. № 1. P. 21. doi: 10.1038/s41541-022-00438-4
  35. Kleinnijenhuis J., Quintin J., Preijers F. et al. Long-lasting effects of BCG vaccination on both heterologous Th1/Th17 responses and innate trained immunity // J. Innate. Immun. 2014. V. 6. № 2. P. 152–158. doi: 10.1159/000355628
  36. Moorlag S.J., Folkman L., Ter Horst R. et al. Multi-omics analysis of innate and adaptive responses to BCG vaccination reveals epigenetic cell states that predict trained immunity // Immunity. 2024. V. 57. № 1. P. 171–187. doi: 10.1016/j.jimmuni.2023.12.005
  37. Joseph J. Trained immunity as a prospective tool against emerging respiratory pathogens // Vaccines. 2022. V. 10. № 11. doi: 10.3390/vaccines10111932
  38. Chen J., Gao L., Wu X. et al. BCG-induced trained immunity: History, mechanisms and potential applications // J. Transl. Med. 2023. V. 21. № 1. P. 106. doi: 10.1186/s12967-023-03944-8
  39. O'Neill L.A.J., Netea M.G. BCG-induced trained immunity: Can it offer protection against COVID-19? // Nat. Rev. Immunol. 2020. V. 20. № 6. P. 335–337. doi: 10.1038/s41577-020-0337-y
  40. Neumann G., Kawaoka Y. Which virus will cause the next pandemic? // Viruses. 2023. V. 15. № 1. doi: 10.3390/v15010199
  41. He Y., Liu W.J., Jia N. et al. Viral respiratory infections in a rapidly changing climate: The need to prepare for the next pandemic // EBioMedicine. 2023. V. 93. doi: 10.1016/j.ebiom.2023.104593
  42. Piret J., Boivin G. The impact of trained immunity in respiratory viral infections // Rev. Med. Virol. 2024. V. 34. № 1. doi: 10.1002/rmv.2510
  43. Zhu J., Liu J., Yan C. et al. Trained immunity: a cutting edge approach for designing novel vaccines against parasitic diseases? // Front. Immunol. 2023. V. 14. doi: 10.3389/fimmu.2023.1252554
  44. Гладков О., Зуков Р., Мамвеев В. et al. Практические рекомендации по лекарственному лечению рака мочевого пузыря // Злокачественные опухоли. 2022. V. 12. № 3S2-1. P. 589–606. doi: 10.18027/2224-5057-2022-12-3s2-589-606
  45. Kaur G., Singh S., Nanda S. et al. Fiction and facts about bcg imparting trained immunity against COVID-19 // Vaccines. 2022. V. 10. № 7. doi: 10.3390/vaccines10071006
  46. Chen J., Gao L., Wu X. et al. BCG-induced trained immunity: history, mechanisms and potential applications // J. of Transplational Medicine. 2023. V. 21. № 1. P. 106. doi: 10.1186/s12967-023-03944-8
  47. Коняшикина С.Ю., Рева С., Петров С. Лабораторные предикторы ответа на внутривенную БЦЖ-терапию у пациентов с немышечно-инвазивным раком мочевого пузыря // Вестник урологии. 2022. Т. 10. № 4. Р. 155–164. doi: 0.21886/2308-6424-2022-10-4-155-164
  48. Jiang S., Redelman-Sidi G. BCG in bladder cancer immunotherapy // Cancers (Basel). 2022. V. 14. № 13. doi: 10.3390/cancers14133073
  49. Pshennikova E., Voronina A. Dormancy: There and back again // Mol. Biology. 2022. V. 56. № 5. P. 735–755. doi: 10.1134/S0026893322050119
  50. De Araujo A., Mambelli F., Sanches R.O. et al. Current understanding of bacillus calmette-guerin-mediated trained immunity and its perspectives for controlling intracellular infections // Pathogens. 2023. V. 12. № 12. doi: 10.3390/pathogens12121386
  51. Ziogas A., Bruno M., van der Meel R. et al. Trained immunity: Target for prophylaxis and therapy // Cell Host Microbe. 2023. V. 31. № 11. P. 1776–1791. doi: 10.1016/j.chom.2023.10.015
  52. Dominguez-Andres J., Dos Santos J.C., Bekkering S. et al. Trained immunity: Adaptation within innate immune mechanisms // Physiol. Rev. 2023. V. 103. № 1. P. 313–346. doi: 10.1152/physrev.00031.2021
  53. Chapman R., Bourn W.R., Shephard E. et al. The use of directed evolution to create a stable and

- immunogenic recombinant BCG expressing a modified HIV-1 Gag antigen // PLoS One. 2014. V. 9. № 7. doi: 10.1371/journal.pone.0103314
54. *Kim B.J., Kim B.R., Kook Y.H., Kim B.J.* Development of a live recombinant BCG expressing Human Immunodeficiency Virus Type 1 (HIV-1) Gag using a pMyong2 vector system: potential use as a novel HIV-1 vaccine // Front. Immunol. 2018. V. 9. doi: 10.3389/fimmu.2018.00643
55. *Kilpelainen A., Saubi N., Guitart N. et al.* Recombinant BCG expressing HTI prime and recombinant ChAdOx1 boost is safe and elicits HIV-1-specific T-Cell responses in BALB/c mice // Vaccines (Basel). 2019. V. 7. № 3. P. 78. doi: 10.3390/vaccines7030078
56. *Gupta P.K.* New disease old vaccine: Is recombinant BCG vaccine an answer for COVID-19? // Cell Immunol. 2020. V. 356. doi: 10.1016/j.cellimm.2020.104187
57. *Peng X., Zhou Y., Zhang B. et al.* Mucosal recombinant BCG vaccine induces lung-resident memory macrophages and enhances trained immunity via mTORC2/HK1-mediated metabolic rewiring // J. Biol. Chem. 2024. V. 300. № 1. doi: 10.1016/j.jbc.2023.105518

## Induction of Trained Immunity by BCG: Recent Data and Opinions

**I. V. Alekseenko<sup>1, 2</sup>, L. G. Kondratyeva<sup>1, 2, \*</sup>, I. P. Chernov<sup>2</sup>, E. D. Sverdlov<sup>1, \*\*</sup>**

<sup>1</sup>National Research Center “Kurchatov Institute”, Moscow, 123182 Russia

<sup>2</sup>Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, 117997 Russia

\*e-mail: liakondratyeva@yandex.ru,

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

Bacillus Calmette-Guerin (BCG), developed in the early 20th century to protect against tuberculosis, is one of the most widely used vaccines in the world. In addition to protecting against tuberculosis, this vaccine has also been shown to reduce the risk of various other infections. During the COVID-19 pandemic, it was found that BCG-vaccinated health care workers had lower detection rates of the SARS-CoV-2 virus compared to unvaccinated individuals. This appears to be due to the generation of trained immunity (TRIM) and non-specific protective effects (NSE), which are likely attributed to the epigenetic and metabolic reprogramming of innate immune cells, known as trained immunity. NSE further holds promise for the development of future therapeutics that can protect against unpredictable pandemics. Additionally, TRIM inducers provide new perspectives on improving the efficacy of standard vaccines by incorporating them into vaccine formulations to enhance both specific and nonspecific immune responses.

**Keywords:** vaccines, BCG, trained immunity, innate immune system, epigenetic modifications.

---

## ГЕНЕТИКА МИКРООРГАНИЗМОВ

---

УДК 575.223:582.282.23

# РЕПАРАТИВНАЯ ДНК-ПОЛИМЕРАЗА *eta* ( $\text{Pol}\eta$ ) ИГРАЕТ КЛЮЧЕВУЮ РОЛЬ В МУТАГЕНЕЗЕ ПРИ НИЗКИХ ДОЗАХ УФ-ОБЛУЧЕНИЯ У ДРОЖЖЕЙ *Saccharomyces cerevisiae*

© 2024 г. Е. А. Алексеева<sup>1, 2, \*</sup>, Т. А. Евстюхина<sup>1, 2</sup>, И. И. Скобелева<sup>1</sup>,  
В. Т. Пешехонов<sup>1, 2</sup>, В. Г. Королев<sup>1, 2</sup>

<sup>1</sup>Петербургский институт ядерной физики им. Б.П. Константинова Национального исследовательского центра  
“Курчатовский институт”, Ленинградская область, Гатчина, 188300 Россия

<sup>2</sup>Курчатовский геномный центр – ПИЯФ, Ленинградская область, Гатчина, 188300 Россия

\*e-mail: alekseeva\_ea@pnpi.nrcki.ru

Поступила в редакцию: 20.05.2024 г.

После доработки: 12.07.2024 г.

Принята к публикации: 16.07.2024 г.

В нормальных условиях и при воздействии малых доз мутагенов ключевым способом борьбы с повреждениями ДНК у бактериальных и эукариотических клеток является система толерантности к повреждениям ДНК (ТПД). Две различные ветви ТПД, ошибочная (translesion synthesis – TLS) и безошибочная (рекомбинационная), функционируют во всех эукариотических организмах. TLS вовлекает белковый комплекс полимеразы дзета ( $\text{Pol}\zeta$ ) (кодируемый генами *REV1*, *REV3* и *REV7*) и полимеразу эта ( $\text{Pol}\eta$ ) (кодируемую геном *RAD30*); эти полимеразы эволюционно консервативны от дрожжей до человека. Мы показали, что при низких уровнях доз УФ-излучения решающую роль в мутагенезе играет высокая активность  $\text{Pol}\eta$ , которая обходит подавляющее число оказавшихся на пути репликативной машины повреждений ДНК. Как в случае облучения высокими, так и низкими дозами УФ-лучей большую роль в индуцированном мутагенезе играет чекпойнт. Однако функции генов, участвующих в регуляции УФ-индуцированного мутагенеза при низких и высоких дозах, часто различаются. Инактивация гена *RAD30* при низких дозах приводит к повышенному уровню индуцированного мутагенеза, а при высоких дозах мутагенез практически не отличается от уровня штамма дикого типа. Делекции генов *HIM1*, *HSM3* и *HIFI*, наоборот, снижают высокий уровень мутагенеза, характерный для высоких доз, до уровня штамма дикого типа при низких дозах. Эти различия, по-видимому, характеризуют локальный и глобальный чекпойнт.

**Ключевые слова:**  $\text{Pol}\eta$ , *S. cerevisiae*, чекпойнт, низкие дозы, УФ-индуцированный мутагенез.

**DOI:** 10.31857/S0016675824120035 **EDN:** WAUMSD

Клетка в ответ на повреждения ДНК использует сеть сигнальных переносчиков, возникающих при прохождении клеточного цикла (чекпойнт) и в течение репарационного процесса [1]. Согласно с этими сигналами чекпойнт формирует ответ клетки на повреждения ДНК [2]. Хотя хорошо известно, что чекпойнт имеет решающее значение для поддержания целостности генома, но как клетка балансирует между остановкой клеточного цикла и клеточной пролиферацией в условиях постоянных эндогенных и экзогенных источников повреждения ДНК остается загадкой. Например, для устранения небольших количеств повреждений ДНК (таких как спонтанные повреждения оснований) используются эксцизионная репарация и процесс толерантности к повреждениям ДНК. При этом

существенной задержки клеточного цикла не происходит. Однако при ответе на более высокие уровни повреждений ДНК эти процессы становятся тесно интегрированными с задержкой клеточного цикла (глобальный чекпойнт) [1]. Для генотоксических агентов существует порог дозы, ниже которого активация чекпойнта минимальна, несмотря на низкую активность путей репарации ДНК. Этот порог может варьировать в зависимости от повреждающего агента, организма и типа клеток.

Доминирующую роль в толерантности к повреждениям ДНК играет безошибочная ветвь пострепликативной репарации (ПРР), которую часто называют рекомбинационной, так как оба типа репарации имеют общую стадию образования D-петли. Наибольший прогресс в изучении связи между

мутагенезом, репарацией, динамикой хроматина и клеточным циклом достигнут на модельном одноклеточном эукариотическом организме – почковавшихся дрожжах *Saccharomyces cerevisiae*. Эксперименты с дрожжами показали, что безошибочные механизмы являются основными путями ПРР при любых репликативных стрессах [2–6]. Тем не менее в литературе есть примеры, когда ошибочный путь синтеза через повреждение (TLS) при малых количествах повреждений ДНК может быть также эффективным [1].

В дрожжах во время нормальной репликации ДНК канонический путь глобального чекпойнта, регулируемого киназами Mec1 и Rad53, не функционирует. Перед вступлением в S-фазу клетки не накапливают больших количеств дНТФ, а скорее входят в нее с субоптимальными уровнями дНТФ [7]. Невысокий уровень дезоксинуклеотидов при входе в S-фазу быстро истощается. При этом ДНК-полимеразы, вероятно, не способны эффективно использовать дНТФ в таких низких концентрациях и должны останавливаться и активировать путь, контролируемый Mec1 – Rad53. Активация синтеза дНТФ происходит в ответ на накопление многочисленных однонитевых брешей за остановленными вилками репликации. Временная активация Mec1-зависимого пути приводит к деградации репрессора Sml1 и активации экспрессии генов, кодирующих субъединицы комплекса рибонуклеотидредуктазы (RNR), и, как следствие, синтезу необходимых количеств дНТФ.

**Таблица 1.** Штаммы дрожжей, использованные в работе

Штамм	Генотип
LMG-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1</i>
4-EAA-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad30Δ</i>
TAE-152	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad30Δ hsm3Δ</i>
TAE-141	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad30Δ rad53+HA-F</i>
TAE-142	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad30Δ sml1Δ</i>
TAE-140	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad30Δ pms1Δ</i>
5-EAA-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad30Δ him1Δ</i>
6-DVF-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 sml1Δ</i>
1-EAA-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 him1Δ</i>
10-DVF-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad53+HA-F</i>
9-DVF-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 rad53+HA-F him1Δ</i>
CAY-3	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 hif1Δ</i>
5-LMG-3031	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 hsm3Δ</i>
TAE-143	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 hif1Δ rad30Δ</i>
2-IVF-312	<i>MAT<math>\alpha</math> ade2Δ-248 ura3-160, 188 leu2-3, 112 trp1 pms1Δ</i>

Изучение молекулярного механизма биологического действия сверхмалых количеств повреждений ДНК является актуальной задачей. Из литературы известно, что только мутанты с нарушением путей пострепликативной репарации проявляли какую-либо чувствительность к хроническому воздействию низких доз УФ-излучения [4, 8]. В настоящей работе мы показали, что при низких уровнях доз УФ-излучения решающую роль в мутагенезе играет высокая активность Pol $\eta$ , которая обходит подавляющее число повреждений ДНК, оказавшихся на пути репликативной машины. При этом частота УФ-индуцированного мутагенеза у мутанта *rad30Δ* зависит от уровня активации комплекса RNR. В отсутствие Pol $\eta$  ее роль в пострепликативной репарации выполняет склонная к ошибкам Pol $\zeta$ , которая создает ошибки в сайте повреждения и на неповрежденной нити в процессе репаративного синтеза ДНК.

## МАТЕРИАЛЫ И МЕТОДЫ

### Штаммы

Делекции генов получали путем ПЦР-опосредованного разрушения генов [9] с использованием специфических маркеров. Генотипы штаммов дрожжей, использованных в настоящем исследовании, представлены в табл. 1.

### Получение штаммов

ПЦР-генерированный *KanMX* модуль был амплифицирован из плазиды pFLA6A-*kanMX*6, полученной от В.Ф. Латыпова, с использованием праймеров: *RAD30* DelL: 5'-ACTTGGAAAGGAGTT GATTAGCTGGTCCCCAGTAAAGCATACGA GTCCTCCTTAGCATGTATGCCCGCTCGTACG CTGCAGGTG-3';

*RAD30* DelR: 5'-CTTGTAAGGATGATAAGATG TTTTGGAAAGATGTAACCTGTTCTGAGGT GTGGCAGTATGTTGTGGCATAGGCCACTAGTG GATC-3'.

10-DVF-3031, 6-DVF-3031, CAY-3 и 2-IVF-312 трансформировали фрагментом ДНК, содержащим маркер *KanMX*, с фланкирующими нуклеотидными последовательностями, гомологичными флангам гена *RAD30*. Трансформанты с делецией гена *RAD30* отбирали по их устойчивости к канамицину (G418) с использованием полной среды, содержащей G418 в концентрации 200 мкг/мл. Разрушение гена *RAD30* было подтверждено с помощью ПЦР. Таким образом получены двойные мутанты: TAE-141, TAE-142, TAE-143 и TAE-140 (табл. 1).

### Питательные среды

Состав полной среды YPD для выращивания культур и учета выживаемости приведен в [9]. В качестве селективной среды для учета частоты УФ-индуцированных мутаций устойчивости к канаванину использовали минимальную среду, обогащенную полным набором аминокислот (за исключением аргинина) и азотистых оснований [10]. В зависимости от использованных штаммов концентрации канаванина достигали 100 мг/л.

### Чувствительность к УФ-излучению

Тесты на выживаемость клеток проводили путем выращивания в течение ночи культуры соответствующего штамма в жидкой YPD при 30 °C. Клетки промывали и ресуспендировали в воде с плотностью  $1 \times 10^7$  клеток/мл. Супензию клеток облучали УФ-лампой БУВ-30 (диапазон УФ-С) с мощностью дозы на верхнем уровне 1.4 Дж / м<sup>2</sup> × с. Аликвоты отбирали после различных доз облучения, разбавляли и помещали на чашки с YPD для определения количества выживших.

### Мутационный тест

Мутационные тесты проводили путем выращивания в течение ночи клеток соответствующего штамма в жидкой YPD при 30 °C. Клетки промывали и ресуспендировали в воде с плотностью  $1 \times 10^7$  клеток/мл. Клетки облучали УФ-лампой БУВ-30. Аликвоты отбирали после различных доз облучения, разбавляли и помещали на чашки с YPD для определения количества выживших.

Для определения частоты мутаций неразведенные аликвоты высевали на среду с канаванином.

### ПЦР в реальном времени

Для проведения ПЦР в реальном времени была использована система детекции RT-PCR CFX96 (Bio-Rad, Великобритания). Реакции проводили в объеме 25 мкл, состоящем из 10 мкл 2.5-кратной реакционной смеси для ОТ-ПЦР в присутствии красителя SYBR Green I и эталонного красителя Rox (Syntol, Россия), 13.8 мкл воды, 1 мкл кДНК и 0.1 (2 мМ) соответствующих праймеров (праймеры для гена *RNR3*: For 5'-ACACCTTCATGGTTATAAG-3' и Rev 5'-CGACGATTTCACAACATAA-3'; для гена *ACT1*: For 5'-GAAGGTCAAGATCATTGC-3 и Rev 5'-GTTGGAAGGTAGTCAAAG-3').

Условия амплификации ПЦР были следующими: 1 цикл по 5 мин при 95°C, затем 39 циклов по 15 с при 95°C и 20 с при 52°C. Анализ кривой плавления показал 5-секундное увеличение на 1°C с 55 до 95°C.

Контрольные реакции с праймером и без матрицы реакционных смесей были включены. Два биологических и три технических повтора были выполнены для каждого образца. Результаты были обработаны с использованием программы CFX Manager.

### Статистический анализ

Экспериментальные данные представлены как среднее по крайней мере по трем повторностям со стандартной ошибкой среднего (SEM), а статистические различия определялись с помощью *t*-критерия Стьюдента. Значимость определяли на уровне  $p < 0,05$  [11].

## РЕЗУЛЬТАТЫ

Как отмечалось выше, путь TLS может эффективно использоваться при низких уровнях повреждений ДНК. Это подтверждается данными работы [12], из которой видно, что мутация *rad30Δ* десятикратно повышает репаративный спонтанный мутагенез. Одним из основных участников TLS-пути является Polη, кодируемая геном *RAD30*. На неповрежденной матрице Polη создает мутации со скоростью  $10^{-2} - 10^{-3}$ , в то же время она способна обходить УФ-индуцированные cis-syn циклобутановые димеры с высокой точностью [13–15]. Уровень УФ-индуцированного мутагенеза у мутанта *rad30Δ* зависит в значительной степени от дозы облучения. При высоких уровнях УФ-доз (выше 40 Дж/м<sup>2</sup>) выживаемость мутанта *rad30Δ* практически не отличалась от выживаемости штамма дикого типа, частота индуцированного мутагенеза также не отличалась или даже была ниже, чем у клеток

дикого типа [12]. При снижении дозы уровень мутагенеза начинает превышать уровень дикого типа [16].

#### УФ-индуцированный мутагенез у мутанта *rad30Δ* при низких дозах облучения

Показано, что УФ-облучение клеток дрожжей, находящихся в различных стадиях клеточного цикла, влияет на количество молекул Pol $\eta$  в большей степени, когда клетки облучались в начале S-фазы, чем в G<sub>1</sub>- или в G<sub>2</sub>-фазе. В S-фазе отмечены два противоположных эффекта. С одной стороны, с увеличением дозы УФ в начале этой фазы значительно задерживалось движение по циклу, как результат, происходило накопление Pol $\eta$ . С другой стороны, степень накопления полимеразы была обратно пропорциональна дозе УФ-лучей, так что при дозе 10 Дж/m<sup>2</sup> концентрация Pol $\eta$  была значительно больше, чем при 80 Дж/m<sup>2</sup> [17]. Возникает вопрос, зачем клетки после УФ-облучения накапливают большое количество полимеразы Pol $\eta$  в самом начале S-фазы? Для разрешения этого вопроса мы изучили УФ-индуцированный мутагенез у мутанта *rad30Δ* при низких дозах облучения в сравнении с клетками дикого типа. Мутант *rad30Δ* показал высокую чувствительность к мутагенному действию низких доз УФ-излучения. При дозе 7 Дж/m<sup>2</sup> частота мутагенеза у мутанта *rad30Δ* была примерно в 3 раза выше, чем у клеток штамма дикого типа (рис. 1). Таким образом, можно предположить, что высокая концентрация Pol $\eta$  в начале S-фазы необходима для эффективного и точного обхода относительно небольшого числа повреждений ДНК, которые остаются после репарационного процесса и попадают в вилки репликации, что может обеспечить нормальную скорость репликативного процесса.

В клетках мутанта *rad30Δ TLS* происходит при участии Pol $\zeta$  [16]. Отсюда следует, что причиной *rad30Δ*-опосредованного УФ-индуцированного мутагенеза является замена Pol $\delta$  на ошибочную Pol $\zeta$  при синтезе ДНК в репаративных брешах. Эта полимераза обладает высокой процессивностью и способна заполнить репаративную брешь. При этом она будет создавать мутации в сайте повреждения и из-за отсутствия у нее редакторской функции делать ошибки при репаративном синтезе неповрежденной ДНК. Последние являются субстратом для репарации ошибочно спаренных оснований (мисматч-репарация). Чтобы проверить это предположение, мы удалили ген *PMS1*, который является одним из ключевых участников мисматч-репарации, у мутанта *rad30Δ*. Одиночный мутант *rad30Δ* при низких уровнях доз УФ-излучения показал значимую чувствительность, в то время как одиночный *pms1Δ* не отличался от штамма дикого типа (рис. 1). Двойной мутант *pms1Δ rad30Δ* не отличался по этому показателю от штамма дикого типа и

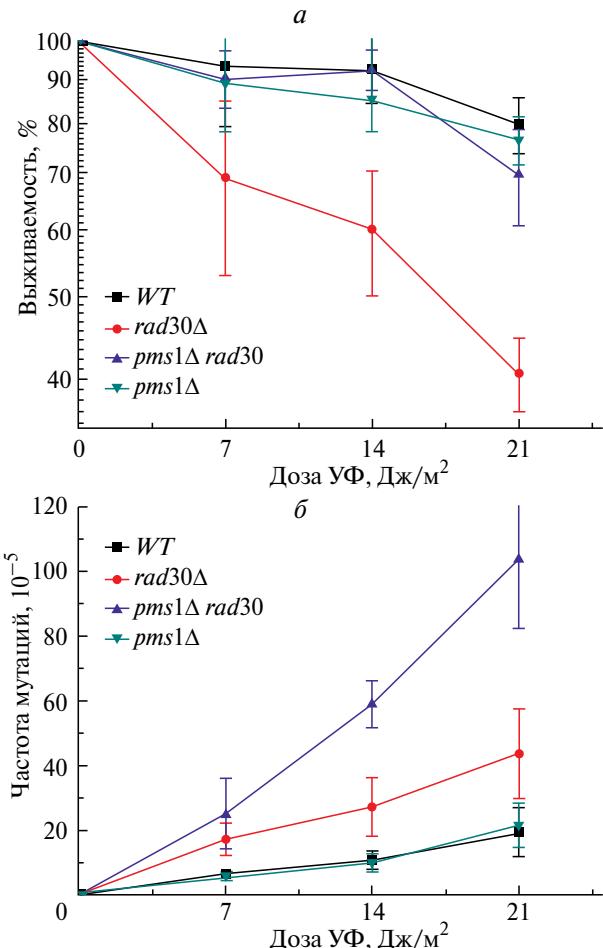
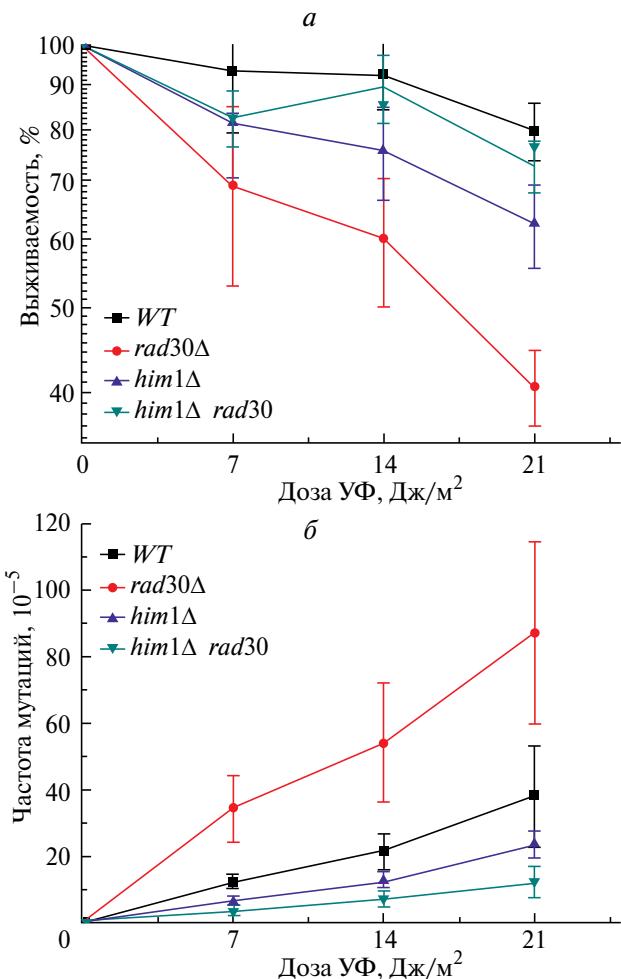


Рис. 1. Чувствительность к УФ-свету и частота УФ-индуцированного мутагенеза в локусе *CAN1* у штамма дикого типа и мутантных штаммов *rad30Δ*, *pms1Δ* и *pms1Δ rad30Δ* при обработке дозами 7, 14 и 21 Дж/m<sup>2</sup> ультрафиолетового излучения. На графиках показаны стандартные ошибки среднего значения ( $\pm$  SEM), полученные в результате пяти независимых экспериментов.

одиночного *pms1Δ*. Этот результат показывает, что УФ-чувствительность одиночного мутанта *rad30Δ* связана с работой мисматч-репарации. В двойном мутанте *pms1Δ rad30Δ* уровень индуцированного мутагенеза при низких дозах УФ был в 2 раза выше, чем у одиночного мутанта *rad30Δ* (рис. 1). На основании полученных результатов можно сделать вывод, что в отсутствие Pol $\eta$  в клетках дрожжей ее роль в пострепликативной репарации выполняет ошибочная Pol $\zeta$ , которая создает мутации в сайте повреждения и на неповрежденной нити в процессе репаративного синтеза ДНК. Репаративный синтез может происходить, когда Pol $\zeta$  обходит повреждение и продолжает синтез до заполнения бреши. При этом процесс обхода повреждения будет осуществлен без участия рекомбинационной ветви пострепликативной репарации. Второй путь обхода



**Рис. 2.** Чувствительность к УФ-свету и частота УФ-индуцированного мутагенеза в локусе *CAN1* у штамма дикого типа и мутантных штаммов *rad30Δ*, *him1Δ* и *him1Δ rad30Δ* при обработке дозами 7, 14 и 21 Дж/м<sup>2</sup> ультрафиолетового излучения. На графиках показаны стандартные ошибки среднего значения ( $\pm$  SEM), полученные в результате пяти независимых экспериментов.

повреждения по механизму безошибочной ветви пострепликативной репарации может осуществляться при участии Pol $\zeta$  для заполнения участка бреши, образовавшегося при преждевременной терминации D-петли [11].

#### УФ-индуцированный мутагенез у мутанта *him1Δ* при низких дозах облучения

Ранее мы показывали, что гены, контролирующие репаративную сборку хроматина, задействованы в контроле локального чекпойнта через подавление экспрессии генов комплекса RNR [11, 18]. В связи с этим мы провели эпистатический анализ взаимодействия мутации *rad30Δ* с мутациями в генах, контролирующих репаративную сборку хроматина. Предполагается, что белок Him1 в клетках

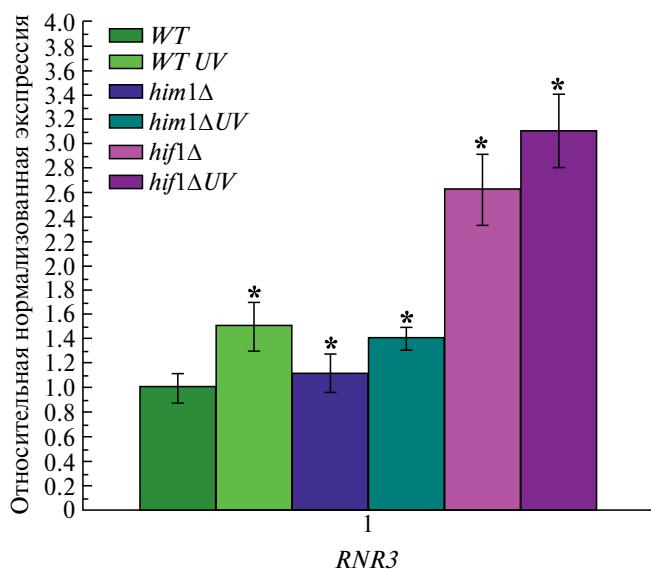
дрожжей выполняет функцию гистонового шаперона при репаративной сборке хроматина [18]. Мутация *him1Δ* имеет фенотип, подобный мутации *hsm3Δ*, но супрессирует УФ-индуцированную активность рибонуклеотидредуктазного комплекса RNR более эффективно, чем *hsm3Δ* [19]. Одиночный мутант *him1Δ* показал слабую чувствительность к летальному действию УФ по сравнению со штаммом дикого типа (рис. 2), как и при высоких дозах.

Однако уровень индуцированного мутагенеза при всех использованных дозах у мутанта *him1Δ* был ниже, чем у штамма дикого типа, что кардинально отличается от этого показателя при высоких дозах УФ-излучения [19]. Этот результат можно объяснить тем, что у одиночного мутанта *him1Δ* после работы Pol $\zeta$  остается мало повреждений ДНК, что снижает до минимума активацию киназы Rad53. При этом будет понижена роль рекомбинационной ветви репарации и, как следствие, понижен *him1Δ*-специфический мутагенез. У двойного мутанта *him1Δ rad30Δ* при низких дозах УФ мутация *him1Δ* подавляла чувствительность мутанта *rad30Δ* так, что двойной мутант не отличался по выживаемости от штамма дикого типа (рис. 2). В отношении УФ-индуцированного мутагенеза двойной мутант характеризовался слабо выраженным понижением уровня частоты мутагенеза, по сравнению с одиночным мутантом *him1Δ* и полной супрессией *rad30Δ*-специфического мутагенеза (рис. 2). Понижение частоты мутагенеза в двойном мутанте по сравнению с одиночным мутантом происходит вследствие увеличения количества выживших клеток при одинаковой дозе у двойного мутанта. Для подтверждения этого предположения мы провели эксперимент по измерению экспрессии гена *RNR3* при низкой дозе УФ-лучей в мутанте *him1Δ* и штамме дикого типа. Из рис. 3 видно, что уровень экспрессии в обоих штаммах практически не отличался, что подтверждает наше предположение.

Суммируя полученные результаты, можно заключить, что мутация *him1Δ* супрессирует УФ-чувствительность и резко подавляет УФ-индуцированный мутагенез мутанта *rad30Δ*. Таким образом, частота УФ-индуцированного мутагенеза у мутанта *rad30Δ* зависит от уровня активации комплекса RNR.

#### УФ-индуцированный мутагенез у мутантов по генам комплекса NuB4 при низких дозах УФ-излучения

Мутации *hsm3Δ* и *hif1Δ* инактивируют обе субъединицы ацетилазного комплекса NuB4, выполняющего функцию гистонового шаперона при репаративной сборке хроматина, и частично супрессируют УФ-индуцированную активность RNR. Белки, кодируемые этими генами, как и



**Рис. 3.** Относительная нормализованная экспрессия гена *RNR3* в штамме дикого типа и мутантных штаммах *him1Δ* и *hif1Δ* до и после облучения их ультрафиолетовым светом (после УФ-облучения клетки выдерживали в течение четырех часов при 30 °C в термостате для индукции), доза УФ-излучения составляла 14 Дж/м<sup>2</sup>; \*  $p < 0,05$ , *t*-критерий Стьюдента.

белок Him1, функционируют в контроле рекомбинационной ветви постстрепликативной репарации [12, 18]. На рис. 4 приведены кривые зависимости выживаемости от дозы УФ-лучей для одиночных *hsm3Δ*, *rad30Δ* и двойного *hsm3Δ rad30Δ* мутантов.

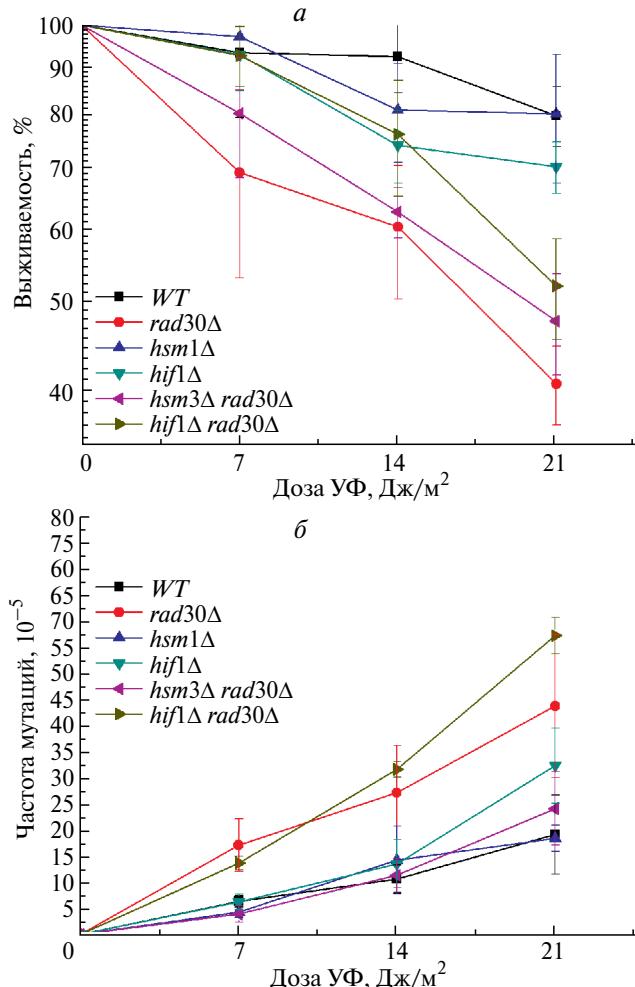
На рис. 4 можно видеть, что мутация *hsm3Δ*, возможно, имеет слабый супрессорный эффект на выживаемость мутанта *rad30Δ*. Мутация *hsm3Δ* при низких дозах УФ не влияет на мутагенез и полностью подавляет *rad30Δ*-специфический мутагенез. Как и в случае с мутантом *him1Δ*, в одиночном мутанте *hsm3Δ* после работы Polη будет снижена активация киназы Rad53, но в меньшей степени, чем у мутанта *him1Δ*. При этом мутация *hsm3Δ* супрессирует *rad30Δ*-специфический мутагенез до уровня штамма дикого типа.

Одиночный мутант *hif1Δ* при низких дозах УФ оказался немного чувствительнее штамма дикого типа и более устойчивым, чем мутант *rad30Δ* (рис. 4). У двойного мутанта мутация *hif1Δ* эпистатировала к мутации *rad30Δ*. Одиночный мутант *hif1Δ* показал уровень УФ-индуцированного мутагенеза, сравнимый с уровнем дикого типа. Уровень УФ-мутагенеза в штамме *hif1Δ rad30Δ* значительно не отличается от одиночного мутанта *rad30Δ* и значительно превышает уровень мутагенеза в штамме дикого типа. Это свидетельствует о том, что мутация *hif1Δ* в меньшей степени, чем *hsm3Δ*, супрессирует активность киназы Rad53, что подтверждается данными измерения экспрессии гена *RNR3* в одиночном мутанте *hif1Δ* (рис. 3). В то же

время при высоких дозах УФ-облучения, когда глобальный чекпойнт полностью активирован, уровень УФ-индуцированного мутагенеза у мутантов *him1Δ*, *hsm3Δ* и *hif1Δ* более выражено превышает уровень штамма дикого типа [12]. Эти результаты подтверждают высказанное выше заключение, что частота УФ-индуцированного мутагенеза в мутанте *rad30Δ* зависит от уровня активации комплекса RNR.

#### УФ-индуцированный мутагенез у мутантов по генам, отвечающим за путь dНТФ в клетке, при низких дозах УФ-излучения

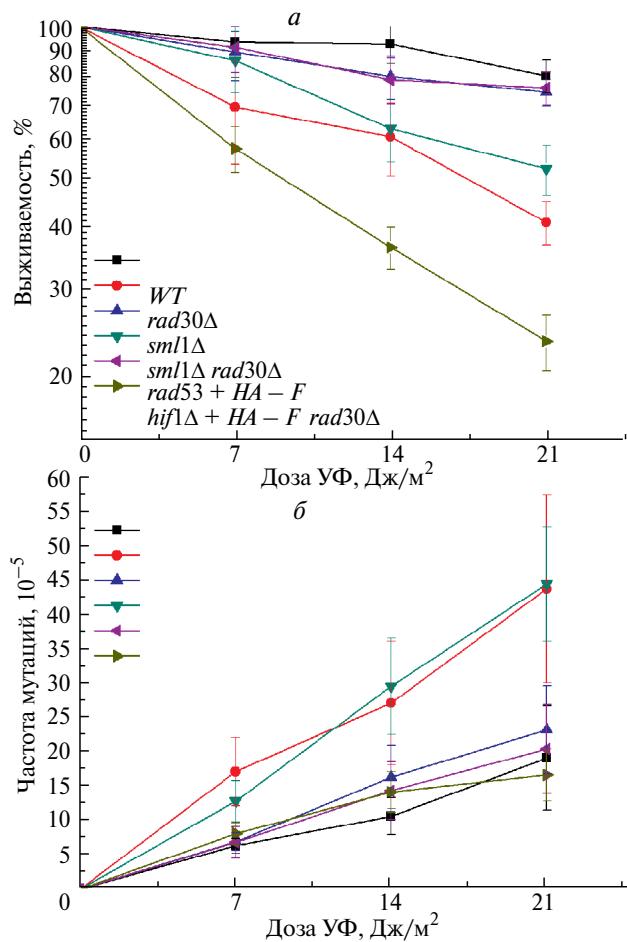
Для проверки этого заключения мы разрушили у мутанта *rad30Δ* ген *SML1*, который кодирует



**Рис. 4.** Чувствительность к УФ-свету и частота УФ-индуцированного мутагенеза в локусе *CAN1* у штамма дикого типа и мутантных штаммов *rad30Δ*, *hsm3Δ*, *hif1Δ*, *hsm3Δ rad30Δ* и *hif1Δ rad30Δ* при обработке дозами 7, 14 и 21 Дж/м<sup>2</sup> ультрафиолетового излучения. На графиках показаны стандартные ошибки среднего значения ( $\pm$  SEM), полученные в результате пяти независимых экспериментов.

супрессор каталитической субъединицы Rnr1 комплекса RNR [20, 21]. Мутация *sml1Δ* увеличивает в клетке пул дНТФ в 2–2.5 раза. Одиночный мутант *sml1Δ* не отличался от штамма дикого типа ни по УФ-чувствительности, ни по уровню УФ-индуцированного мутагенеза (рис. 5). В двойном мутанте мутация *sml1Δ* супрессировала УФ-чувствительность и *rad30Δ*-специфический мутагенез до уровня штамма дикого типа.

Мутация *rad53 + HA-F* блокирует увеличение экспрессии гена *RNR3*, кодирующего субъединицу комплекса RNR при нормальных условиях роста и после облучения высокими дозами УФ-лучей [12]. Мы использовали эту мутацию для подтверждения высказанного выше утверждения о зависимости частоты УФ-индуцированного мутагенеза



**Рис. 5.** Чувствительность к УФ-свету и частота УФ-индуцированного мутагенеза в локусе *CAN1* у штамма дикого типа и мутантных штаммов *rad30Δ*, *sml1Δ*, *rad53 + HA-F*, *sml1Δ rad30Δ* и *rad53 + HA-F rad30Δ* при обработке дозами 7, 14 и 21 Дж/м² ультрафиолетового излучения. На графиках показаны стандартные ошибки среднего значения ( $\pm$  SEM), полученные в результате пяти независимых экспериментов.

у мутанта *rad30Δ* от степени активации комплекса RNR. Был сконструирован двойной мутант *rad30Δ rad53 + HA-F* и изучена эффективность низких доз УФ в индукции мутагенеза у двойного мутанта. На рис. 5 представлен полученный результат, из которого следует, что мутация *rad53 + HA-F* эпистатирует к мутации *rad30Δ*, снижая уровень мутагенеза у двойного мутанта до уровня одиночного мутанта *rad53 + HA-F*. Этот результат согласуется с высказанным выше выводом о том, что частота УФ-индуцированного мутагенеза у мутанта *rad30Δ* зависит от уровня активации экспрессии генов комплекса RNR, которая, в свою очередь, зависит от активности киназы Rad53.

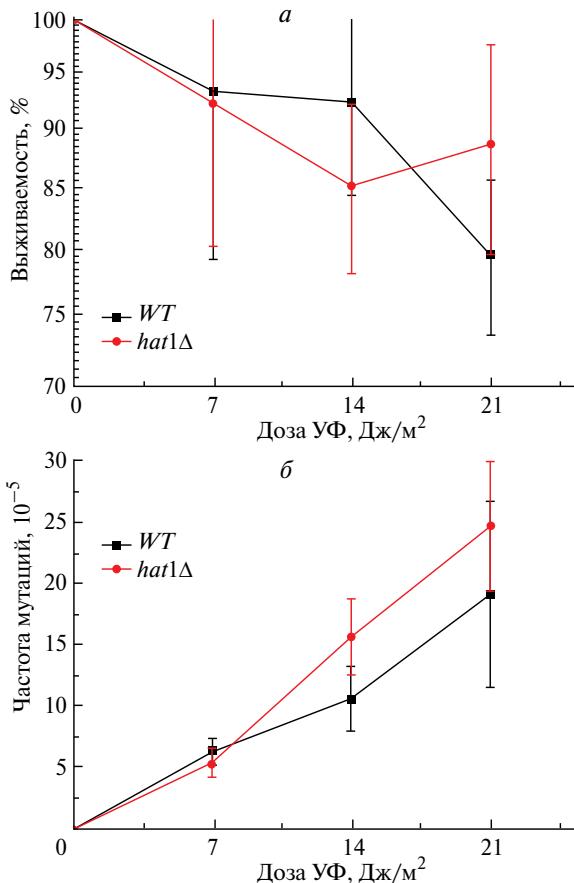
Ген *HAT1* кодирует каталитическую субъединицу комплекса NuB4. Мутация *hat1Δ* инактивирует функцию NuB4, но в отличие от вспомогательных субъединиц практически не влияет на частоту УФ-индуцированного мутагенеза при высоких дозах облучения и приводит к гиперактивации генов комплекса RNR даже в нормальных условиях роста [11].

При низких дозах УФ-облучения зависимость частоты мутаций от дозы была такая же, как при высоких уровнях доз (рис. 6). Этот результат неудивителен, так как гиперактивация генов комплекса RNR ограничивает уровень УФ-индуцированного мутагенеза. Суммируя полученные данные, можно заключить, что при низких и высоких дозах УФ-лучей дозовая зависимость мутагенеза у изученных мутантов по генам, контролирующим репаративную сборку хроматина, контролируется уровнем экспрессии генов комплекса RNR.

## ОБСУЖДЕНИЕ

Считается, что Polη является первым из ферментов, осуществляющих TLS. Она быстро привлекается к остановленным репликативным вилкам и отвечает за безошибочный обход 8-оксогуаниновых повреждений ДНК и УФ-индуцированных пиримидиновых димеров [13, 22]. Мотив PIP1 Polη связывается с убиквитин-конъюгирующим комплексом Rad6–Rad18, необходимым для TLS, что способствует локализации Polη в очагах повреждения ДНК *in vivo* [23]. Предполагается, что Polη уже находится в комплексе, когда происходит реакция переноса убиквитина [24]. Эти наблюдения подтверждают модель, в которой Polη может быть задействована в остановленной вилке репликации, когда комплекс Rad6–Rad18–Polη связывается с PCNA [25].

Мутация *rad30Δ* приводит к тому, что после УФ-облучения низкими дозами клеток этого мутанта, остановленных на повреждениях вилок репликации остается кратно больше по сравнению с клетками дикого типа. Вследствие этого происходит резкое увеличение количества однонитевой



**Рис. 6.** Чувствительность к УФ-свету и частота УФ-индуцированного мутагенеза в локусе *CAN1* у штамма дикого типа и мутантного штамма *hat1Δ* при обработке дозами 7, 14 и 21 Дж/м<sup>2</sup> ультрафиолетового излучения. На графиках показаны стандартные ошибки среднего значения ( $\pm$  SEM), полученные в результате пяти независимых экспериментов.

ДНК в репарационных брешах. Однонитевая ДНК, покрытая белком Rpa, служит сигналом для активации киназы Rad53. В результате у мутанта *rad30Δ* произойдет избыточная активация киназы и увеличение концентрации дНТФ по сравнению с клетками дикого типа при одинаковой дозе УФ-лучей. Это увеличение и создает предпосылки для *rad30Δ*-специфического мутагенеза. Согласно результатам наших предыдущих работ [11, 18], условия для специфического повышенного мутагенеза возникают, когда активность киназы Rad53 снижается с максимальных значений или повышается с минимальных до промежуточного уровня.

Мутанты дрожжей по генам, кодирующим белки, участвующие в репаративной сборке хроматина, показали высокие уровни спонтанного репаративного мутагенеза, что может свидетельствовать о том, что эти гены могут играть позитивную роль в регуляции локального чекпойнта [12, 18]. Это предположение было подтверждено в экспериментах

по изучению УФ-индукционного мутагенеза у двойных мутантов *him1Δ rad2Δ*, *hif1Δ rad2Δ*, *hsm3Δ rad2Δ* и *hat1Δ rad2Δ*. Мутанты дрожжей с заблокированной эксцизионной репарацией нуклеотидов очень чувствительны к низким дозам УФ-излучения, которые не индуцируют глобальный чекпойнт. Для всех двойных мутантов мы наблюдали синергическое увеличение частоты индуцированных мутаций, хотя у последнего мутанта это увеличение было в пределах доверительного интервала только при минимальной дозе УФ [11, 12, 26]. Из этих экспериментов следует, что все использованные мутации ответственны за высокий уровень мутагенеза при низких дозах УФ-лучей.

Ранее мы показали, что индуцированный мутагенез у мутанта *rad30Δ* при высоких дозах УФ не отличается от штамма дикого типа, а при очень высоких дозах даже ниже, чем в диком типе [19]. Эти данные противоречили результатам, полученным в работе S. G. Kozmin и соавт. [16], в которой мутант *rad30Δ* показал уровень УФ-индукционного мутагенеза выше, чем у клеток дикого типа. В настоящей работе мы разрешили это противоречие. Оказалось, что при низких дозах УФ уровень мутагенеза у штамма *rad30Δ* действительно значительно выше по сравнению со штаммом дикого типа (рис. 1). Авторы работы [16] объясняют высокий уровень мутагенеза при низких дозах УФ заменой Pol $\eta$  на Pol $\zeta$ , которая делает частые ошибки на повреждениях, которые в норме безошибочно обходятся Pol $\eta$ . Однако эксперименты с блокированием мисматч-репарации у мутанта *rad30Δ* ставят под сомнение это утверждение. После обхода повреждения с ошибочным либо безошибочным встраиванием нуклеотидов образуется структура, которая является субстратом для нуклеотид-эксцизионной, но не мисматч-репарации. В то же время мы наблюдаем уровень УФ-индукционного мутагенеза в двойном мутанте *rad30Δ pms1Δ* в 2 раза выше, чем у одиночного мутанта *rad30Δ* (рис. 1). Этот результат можно объяснить, если субстраты для мисматч-репарации возникают в результате ошибок Pol $\zeta$  в ходе репарационного синтеза на неповрежденной матрице. Это предположение вполне допустимо, так как Pol $\zeta$  лишена редакторской функции, что снижает точность синтеза ДНК на два порядка по сравнению с репликативными полимеразами. Pol $\zeta$  может осуществить репаративный синтез либо сразу после обхода повреждения, либо после преждевременной терминации D-петли в течение рекомбинационного обхода повреждения. Чтобы решить, какой из этих вариантов предпочтительней реализуется в клетке, мы осуществили ряд экспериментов. Если предпочтителен первый вариант, то уровень УФ-мутагенеза не должен заметно зависеть от уровня дНТФ в клетке, в то время как реализация второго варианта требует строго определенного повышения этого уровня

[12]. Мы разрушили ген *SML1* у мутанта *rad30Δ*, что привело к кратному повышению концентрации дНТФ и снижению уровня УФ-мутагенеза до уровня штамма дикого типа (рис. 5). Этот результат свидетельствует в пользу второго варианта. В следующих экспериментах мы, наоборот, ограничили индукцию дНТФ, введя в клетки мутанта *rad30Δ* мутацию *rad53+HA-F* [12]. У мутанта *rad53+HA-F rad30Δ* уровень УФ-индуцированного мутагенеза также снизился до уровня штамма дикого типа (рис. 5). На основании полученных данных мы заключили, что *rad30Δ*-специфический мутагенез реализуется в течение рекомбинационной ветви пострепликативной репарации в результате заполнения репарационной бреши Pol $\zeta$  после преждевременной терминации *D*-петли, которое требует гипоактивации экспрессии генов комплекса RNR.

Комплекс RNR состоит из четырех субъединиц Rnr1, Rnr2, Rnr3 и Rnr4, которые кодируются генами *RNR1*, *RNR2*, *RNR3* и *RNR4* соответственно. Из этих генов нас больше всего интересовал *RNR3*, так как его экспрессия увеличивается в ответ на повреждения ДНК, в отличие от других генов данного комплекса, чья экспрессия не изменяется в ответ на повреждения ДНК. Поэтому мы рассматривали экспрессию гена *RNR3* у наших мутантных штаммов как до, так и после облучения их УФ-светом.

Для подтверждения этого заключения мы разрушили у мутанта *rad30Δ* гены *HSM3* и *HIF1*. Ранее мы показали, что делеции генов *HSM3* и *HIF1* приводят к частичному подавлению УФ-индуцированной экспрессии гена *RNR3*, кодирующего одну из катализитических субъединиц комплекса RNR, следствием чего было появление *hsm3Δ*- и *hif1Δ*-зависимого УФ-индуцированного мутагенеза. У двойного мутанта *hsm3Δ hif1Δ* происходило дальнейшее подавление экспрессии *RNR3* и исчезновение повышенного УФ-индуцированного мутагенеза, характерного для одиночных мутантов [11]. Как можно видеть из рис. 4, УФ-индуцированный мутагенез у мутантов *hsm3Δ rad30Δ* и *hif1Δ rad30Δ* был ниже по сравнению с одиночным мутантом *rad30Δ*. Этот результат означает, что *rad30Δ*-специфический мутагенез действительно зависит от уровня экспрессии генов комплекса RNR. К такому же результату приводит разрушение гена *HIM1* у мутанта *rad30Δ* (рис. 2).

При высоких дозах УФ-лучей уровень мутагенеза в одиночных мутантах *him1Δ*, *hsm3Δ* и *hif1Δ* кратно превышал уровень мутагенеза, характерного для штамма дикого типа. Однако при низких дозах УФ мутагенез у одиночных мутантов *hsm3Δ* и *hif1Δ* не имел значимых отличий от штамма дикого типа, а у мутанта *him1Δ* он был даже ниже этого уровня (рис. 2, 4). Это различие можно объяснить тем, что при низких дозах УФ-лучей активная работа нуклеотид-эксцизионной репарации оставляет нерепарированными относительно небольшое

число повреждений, которые успешно обходятся Pol $\eta$ , в результате активация киназы Rad53 низкая, что приводит к низкой экспрессии генов комплекса RNR, а мутации *him1Δ*, *hsm3Δ* и *hif1Δ* дополнитель но снижают уровень этой экспрессии.

В отличие от клеток дикого типа у мутанта *rad30Δ* остается значительно больше нерепарированных повреждений, которые будут увеличивать активность киназы Rad53. Эти повреждения будут субстратами пострепликативной репарации в G<sub>2</sub>-фазе клеточного цикла.

Наши результаты показывают, что весь *rad30Δ*-специфический мутагенез происходит на рекомбинационной ветви репарации, так как он зависит от промежуточной активации киназы Rad53. Можно предположить, что Pol $\delta$  вставляет напротив повреждения случайный нуклеотид и далее происходит перенос вновь синтезированной нити в сестринскую двунитевую ДНК с образованием *D*-петли. На конце внедрившейся нити образуется мисматч, который является субстратом исключительно Pol $\zeta$ , которая продолжит репаративный синтез в *D*-петле и после ее преждевременной терминации на материнской молекуле в бреши. Мисматч-репарация исправит ошибочно спаренные основания, возникшие из-за ошибок Pol $\zeta$  на неповрежденной матрице, но не затронет ошибку над повреждением.

Суммируя полученные данные, можно сделать вывод, что при низких уровнях доз УФ-излучения решающую роль играет высокая активность Pol $\eta$ , которая обходит подавляющее число оказавшихся на пути репликативной машины повреждений ДНК. Как в случае облучения высокими, так и низкими дозами УФ-лучей большую роль в индуцированном мутагенезе играет чекпойнт. Однако роли генов, участвующих в регуляции УФ-индуцированного мутагенеза, при низких и высоких дозах часто различаются. Например, инактивация гена *RAD30* при низких дозах приводит к повышенному уровню индуцированного мутагенеза, а при высоких дозах практически не отличается от уровня штамма дикого типа. Инактивация генов *HIM1*, *HSM3* и *HIF1*, наоборот, снижает высокий уровень мутагенеза, характерный для высоких доз, до уровня штамма дикого типа при низких дозах. Эти различия характеризуют локальный и глобальный чекпойнт.

Работа выполнена при финансовой поддержке гранта РНФ № 23-24-00119.

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. Lazzaro F., Giannattasio M., Puddu F. et al. Checkpoint mechanisms at the intersection between DNA damage and repair // *DNA Repair*. 2009. V. 8. P. 1055–1067. doi: 10.1016/j.dnarep.2009.04.022
2. Prakash L. Characterization of postreplication repair in *Saccharomyces cerevisiae* and effects of *rad6*, *rad18*, *rev3* and *rad52* mutations // *Mol. Gen. Genet.* 1981. V. 184. P. 471–478. doi: 10.1007/BF00352525
3. Baynton K., Bresson-Roy A., Fuchs R.P.P. Analysis of damage tolerance pathways in *Saccharomyces cerevisiae*: A requirement for Rev3 DNA polymerase in translesion synthesis // *Mol. Cell. Biol.* 1998. V. 18. P. 960–966. doi: 10.1128/MCB.18.2.960
4. Hishida T., Kubota Y., Carr A.M., Iwasaki H. RAD6–RAD18–RAD5-pathway-dependent tolerance to chronic low-dose ultraviolet light // *Nature*. 2009. V. 457. P. 612–615. doi: 10.1038/nature07580
5. Huang D., Piening B.D., Paulovich A.G. The preference for error-free postreplication repair in *Saccharomyces cerevisiae* exposed to low-dose methyl methanesulfonate is cell cycle dependent // *Mol. Cell. Biol.* 2013. V. 33. P. 1515–1527. doi: 10.1128/MCB01392-12
6. Zhang H., Lawrence C.W. The error-free component of the RAD6/RAD18 DNA damage tolerance pathway of budding yeast employs sister-strand recombination // *Proc. Natl Acad. Sci. USA*. 2005. V. 102. P. 15954–15959. doi: 10.1073/pnas.0504586102
7. Forey R., Poveda A., Sharma S. et al. Mec1 is activated at the onset of normal S phase by low-dNTP pools impending DNA replication // *Mol. Cell*. 2020. V. 78. P. 396–410. doi: 10.1016/molcel.2020.02.021
8. Tercero J.A., Longhese M.P., Diffley J.F. A central role for DNA replication forks in checkpoint activation and response // *Mol. Cell*. 2003. V. 11. P. 1323–1336. doi: 10.1016/s1097-2765(03)00169-2
9. Захаров И.А., Кожин С.А., Кожина Т.А. и др. Сборник методик по генетике дрожжей-сахаромицетов. Л.: Наука, 1984, изд. 2-е. 144 с.
10. Fedorova I.V., Kovaltzova S.V., Gracheva L.M. et al. Requirement of *HSM3* gene for spontaneous mutagenesis in *Saccharomyces cerevisiae* // *Mutat. Res.* 2004. V. 554. P. 65–75. doi: 10.1016/j.mrfmmm.2004.03.003
11. Evstyukhina T.A., Alekseeva E.A., Fedorov D.V. et al. Genetic analysis of the Hsm3 protein function in yeast *Saccharomyces cerevisiae* NuB4 complex // *Genes*. 2021. V. 12. doi: 10.3390/genes12071083
12. Evstyukhina T.A., Alekseeva E.A., Peshekhonov V.T. et al. The role of chromatin assembly factors in induced mutagenesis at low levels of DNA damage // *Genes*. 2023. V. 14. doi: 10.3390/genes14061242
13. Johnson R.E., Prakash S., Prakash L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Polη // *Science*. 1999. V. 283. P. 1001–1004. doi: 10.1126/science.283.5404.1001
14. Prakash S., Johnson R.E., Prakash L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function // *Annu. Rev. Biochem.* 2005. V. 74. P. 317–353. doi: 10.1146/annurev.biochem.74.082803.133250
15. Wang Z., Xiao W. Distinct requirement for budding yeast Rev1 and Polη in translesion DNA synthesis across different types of DNA damage // *Curr. Genet.* 2020. V. 5. P. 1019–1028. doi: 10.1007/s00294-020-01092-w
16. Kozmin S.G., Pavlov Y.I., Kunkel T.A., Sage E. Roles of *Saccharomyces cerevisiae* DNA polymerases Polη and Polξ in response to irradiation by simulated sunlight // *Nucl. Acids Res.* 2003. 2003. V. 31. P. 4541–4552. doi: 10.1093/nar/gkg489
17. Sobolevska A., Halas A., Plachta M. et al. Regulation of the abundance of Y-family polymerases in the cell cycle of budding yeast in response to DNA damage // *Curr. Genet.* 2020. V. 66. P. 749–763. doi: 10.1007/s00294-020-01061-3
18. Alekseeva E.A., Evstyukhina T.A., Peshekhonov V.T. et al. The Role of the RPD3 complex of *Saccharomyces cerevisiae* yeast in the activation of UV-induced expression of RNR complex genes // *J. Biomed. Res. Environ. Sc.* 2024. V. 5. P. 360–372. doi: 10.37871/jbres1902
19. Alekseeva E.A., Evstyukhina T.A., Peshekhonov V.T., Korolev V.G. Participation of the *HIM1* gene of yeast *Saccharomyces cerevisiae* in the error-free branch of post-replicative repair and role Polη in *him1*-dependent mutagenesis // *Curr. Genet.* 2021. V. 67. P. 141–151. doi: 10.1007/s00294-020-01115-6
20. Zhao X., Chabes A., Domkin V. et al. The ribonucleotide reductase inhibitor Sml1 is a new target of the Mec1/Rad53 kinase cascade during growth and in response to DNA damage // *EMBO J.* 2001. V. 20. P. 3544–3553. doi: 10.1093/emboj/20.13.3544
21. Chabes A., Georgieva B., Domkin V. et al. Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase // *Cell*. 2003. V. 112. P. 391–401. doi: 10.1016/s0092-8674(03)00075-8
22. Haracska L., Johnson R.E., Prakash L., Prakash S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta //

- Nature genet. 2000. V. 25. P. 458–461.  
doi: 10.1038/78169
23. Watanabe K., Tateishi S., Kawasaji M. et al. Rad18 guides poleta to replication stalling sites through physical interaction and PCNA monoubiquitination // EMBO J. 2004. V. 23. P. 3886–3896.  
doi: 10.1038/sj.emboj.7600383.
24. Durando M., Tateishi S., Vaziri C. A non-catalytic role of DNA polymerase eta in recruiting Rad18 and promoting PCNA monoubiquitination at stalled replication forks // Nucl. Acids Res. 2013. V. 41. P. 3079–3093.  
doi: 10.1093/nar/gkt016
25. Ripley B.M., Gildenberg M.S., Washington M.T. Control of DNA damage bypass by ubiquitylation of PCNA // Genes. 2020. V. 11. P. 138.  
doi.org/10.3390/genes11020138
26. Fedorova I.V., Gracheva L.M., Kovaltzova S.V. et al. The yeast *HSM3* gene acts in one of the mismatch repair pathways // Genetics. 1998. V. 148. P. 963–973.  
doi: 10.1093/genetics/148.3.963

## The Reparative DNA Polymerase eta (Polη) Plays A Key Role in Mutagenesis at Low Doses of UV Radiation in Yeast *Saccharomyces cerevisiae*

E. A. Alekseeva<sup>1, 2, \*</sup>, T. A. Evstyukhina<sup>1, 2</sup>, I. I. Skobeleva<sup>1</sup>, V. T. Peshekhanov<sup>1, 2</sup>, V. G. Korolev<sup>1, 2</sup>

<sup>1</sup>Konstantinov Petersburg Nuclear Physics Institute of the National Research Centre “Kurchatov Institute”, Leningrad oblast, Gatchina, 188300 Russia

<sup>2</sup>Kurchatov Genome Center — Petersburg Nuclear Physics Institute, Leningrad oblast, Gatchina, 188300 Russia  
\*e-mail: alekseeva\_ea@pnpi.nrcki.ru

Under normal conditions and when exposed to low doses of mutagens, DNA damage tolerance systems (DDT) are a key way to combat DNA damage in bacterial and eukaryotic cells. Two different DDT pathways, erroneous (translesion synthesis – TLS) and error-free (recombination), operate in all eukaryotic organisms. TLS involves the protein complex polymerase zeta (Polζ) (encoded by the genes *REV1*, *REV3* and *REV7*), and polymerase eta (Polη) (encoded by the *RAD30* gene); these polymerases are conserved from yeast to humans. We have shown that at low levels of UV radiation doses, a decisive role in mutagenesis is played by the high activity of Poly, which bypasses the overwhelming number of DNA damage that got in the way of the replication machine. Both in the case of irradiation with high and low doses of UV rays, checkpoint plays an important role in induced mutagenesis. However, the roles of genes involved in the regulation of UV-induced mutagenesis at low and high doses often differ. Inactivation of the *RAD30* gene at low doses leads to an increased level of induced mutagenesis, and at high doses it practically does not differ from the level of the wild-type strain. Deletions of the *HIM1*, *HSM3* and *HIF1* genes, on the contrary, reduce the high level of mutagenesis characteristic of high doses to the level of the wild-type strain at low doses. These differences characterize the local and global checkpoint.

**Keywords:** Polη, *S. cerevisiae*, checkpoint, low doses, UV-induced mutagenesis.

## ТРАНСКРИПЦИОННАЯ АКТИВНОСТЬ *CCA1* У РАСТЕНИЙ *Arabidopsis thaliana* СЕВЕРНОЙ ПОПУЛЯЦИИ В УСЛОВИЯХ ИЗМЕНЕННОГО СВЕТОВОГО РЕЖИМА

© 2024 г. М. В. Зарецкая<sup>1,\*</sup>, О. М. Федоренко<sup>1,\*\*</sup>

<sup>1</sup>Институт биологии Карельского научного центра Российской академии наук, Петрозаводск, 185910 Россия

\*e-mail: genmt@mail.ru

\*\*e-mail: fedorenko\_om@mail.ru

Поступила в редакцию 20.05.2024 г.

После доработки 13.06.2024 г.

Принята к публикации 28.06.2024 г.

Проанализирована динамика транскрипционной активности одного из ключевых генов циркадной сети *CCA1* в условиях естественного светового фотопериода длинного дня (16L : 8D) и при инвертированном световом режиме (8D : 16L) у растений *A. thaliana* северной природной популяции (Карелия). Показано, что в условиях инвертированного смещения светового режима происходит резкий подъем экспрессии этого гена со сдвигом фазы в циркадном ритме на 2 часа. Уровень транскрипционной активности *CCA1* оказался почти в два раза выше по сравнению с естественными световыми условиями. При этом эндогенный ритм гена сохранялся, но с меньшей амплитудой. С возрастом у 30-дневных растений, выросших в инвертированных условиях, произошла потеря эндогенного циркадного ритма *CCA1*. Полученные результаты позволяют заключить, что, вероятно, циркадные ритмы *A. thaliana* северных природных популяций выполняют важную роль в адаптации к изменению световых условий и что один из ключевых генов часов *CCA1* играет в этом процессе существенную роль.

**Ключевые слова:** *Arabidopsis thaliana*, северные популяции, циркадные ритмы, транскрипционная активность *CCA1*, фотопериод.

**DOI:** 10.31857/S0016675824120046 **EDN:** WARNCQ

Циркадные ритмы являются одним из важнейших механизмов, участвующих в адаптации растений наряду со временем цветения и сроками прорастания семян. У большинства живых организмов имеются внутренние биологические часы, генерирующие суточные (циркадные) ритмы биологических процессов с периодом около 24 часов. Эта внутренняя система циркадных ритмов полагается на интегрированные внешние сигналы окружающей среды (продолжительность дня и колебание температуры) для контроля важных физиологических процессов, определяющих развитие, и, таким образом, способствует общей приспособленности организма. Несмотря на связь с внешними стимулами, циркадные ритмы имеют эндогенное происхождение. Они не следуют пассивно циклическим изменениям светового режима, а являются результатом слаженной работы сложной системы, состоящей из центральных и периферических осцилляторов [1, 2].

После синхронизации с внешней средой осцилляторы регулируют ритмическое накопление значительной доли транскриптов, белков и

метаболитов. Такая циркадная система способствует синхронизации различных биологических функций с периодически изменяющимися условиями окружающей среды (циклами света и темноты). Это результат эволюционной адаптации к жизни в определенных геофизических условиях на вращающейся планете, где длина дня зависит от географической широты и времени года [3]. Таким образом, циркадные часы дают преимущества в приспособленности, позволяя организмам предвидеть изменения окружающей среды и адаптироваться к ним [4, 5].

Однако до сих пор нет четкого понимания принципов построения молекулярной сети, лежащей в основе системы циркадных ритмов. Известно, что это своего рода внутренний осциллятор, который в большинстве клеток живых организмов представлен комплексом белков часовых генов (*clock-genes proteins*) [2, 6, 7]. У *A. thaliana* ключевые гены циркадной сети – *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* и *LATE ELONGATED HYPOCOTYL (LHY)*, взаимосвязаны через гены нисходящего пути с генами цветения в условиях длинного дня [8, 9, 10].

Так, GIGANTEA (GI) – компонент циркадных часов, активирует фотопериодическое цветение, т. е. синхронизирует циркадные ритмы со световыми сигналами окружающей среды [2, 11]. Он также регулирует экспрессию флоральных активаторов – *CONSTANS* (*CO*) и *FLOWERING LOCUS T* (*FT*) [12]. К тому же показано, что генная сеть, контролирующая циркадные ритмы растений, частично перекрываетя сетью, регулирующей цветение [2, 13]. Это может свидетельствовать о том, что циркадные ритмы являются важным механизмом в адаптации растений, способным влиять на время начала цветения в зависимости от геофизических световых условий.

Не так давно было обнаружено, что значительные изменения в режиме свет – темнота вызывают стрессовые реакции, снижая эффективность фотосинтеза. Этот ответ, называемый циркадным стрессом, приводит к снижению экспрессии *CCA1* и *LHY* при продолжительном освещении (32 ч) и длительном периоде темноты (16 ч) [14]. Однако, как было показано, циркадные часы способны регулировать реакцию организма на абиотический стресс. Эта регуляция и подгонка к суточному ритму осуществляется с помощью эпигенетических механизмов. К ним относят модификации гистонов и ремоделирование хроматина, которые регулируют доступность ДНК путем открытия или уплотнения октамера гистонов [5].

В связи с этим с целью выявления способности растений *A. thaliana* карельских популяций адаптироваться к изменению светового режима и определения, насколько быстро это происходит, мы проанализировали в сравнительном аспекте транскрипционную активность одного из ключевых генов циркадных часов – *CCA1* – при естественном световом режиме длинного дня (16L : 8D) и при инвертированном смещении его (8D : 16L). Инвертированное смещение светового режима подразумевает предоставление света в вечернее и ночное время суток при сохранении продолжительности фотопериода.

## МАТЕРИАЛЫ И МЕТОДЫ

В работе использовали растения *A. thaliana*, выращенные из собранных в природе семян северной природной популяции, находящейся в Карелии. Это популяция Шуйская, названная в соответствии с близлежащим населенным пунктом (61°94' с. ш.; 34°25' в. д.). Анализ экспрессии гена *CCA1* проводили на молодых листьях растений. В качестве контроля использована раннецветущая линия – *Ler* (*Landsberg erecta*).

### Выращивание растений в лабораторных условиях

Выращивание растений в лабораторных условиях проводили по общепринятым методикам

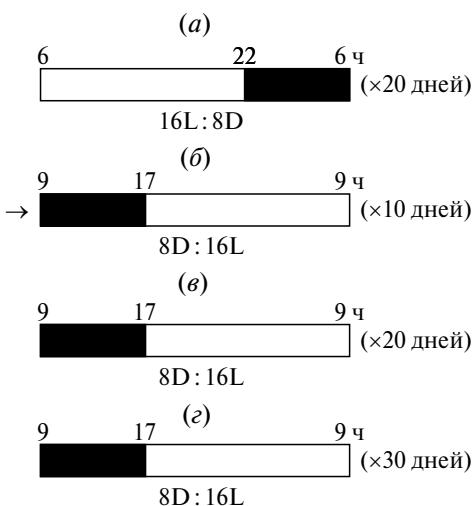
культивирования *A. thaliana* [15]. Семена растений из популяции Шуйская высевали в чашки Петри и проращивали на простой среде по Гихнеру – Велеминскому, которая готовилась на основе 8%-ного агар-агара с добавлением растворов макро- и микроэлементов. Растения выращивали в условиях естественного светового фотопериода длинного дня (16L : 8D) и при инвертированном смещении его. В первом случае свет включался в 6 ч, а выключался в 22 ч; во втором – свет предоставлялся в вечернее и ночное время (с 17 до 9 ч), а темный период – в дневное время (с 9 до 17 ч). При этом длина дня оставалась прежней (8D : 16L). Условия роста на свету: освещение – 10000 лк; температура 22 °C, в темноте – 22 °C.

Для анализа уровня транскриптов *CCA1* использовали следующие группы растений, выращенные в неодинаковых условиях (рис. 1):

*a* – 20-дневные растения карельской популяции и *Ler*, выросшие при естественном световом режиме в условиях длинного дня (16L : 8D); 20-дневные растения (Ш20) анализировали;

*b* – растения карельской популяции из группы “*A*” (Ш20) добрачивали десять дней при инвертированном световом режиме (8D : 16L); в результате анализировали 30-дневные растения (Ш30);

*в* и *г* – семена растений из карельской популяции, посеянные в чашки Петри, сразу помещали в инвертированные световые условия (8D : 16L);



**Рис. 1.** Графическое представление условий выращивания растений при длинном световом дне: (а) – семена посеяны при естественном световом фотопериоде (16L : 8D); свет включался с 6 до 22 ч; 20-дневные растения (Ш20) анализировали. (б) – растения из группы “*a*” добрачивали 10 дней при инвертированном световом режиме (8D : 16L); свет включался с 17 до 9 ч; 30-дневные растения (Ш30) анализировали. (в, г) – семена посеяны в инвертированных световых условиях (8D : 16L); 20-дневные (Ш20) и 30-дневные (Ш30) растения анализировали.

20-дневные растения (C20) и 30-дневные (C30) анализировали.

#### Анализ уровня транскриптов *CCA1*

Выделение суммарной РНК из листьев растений осуществлялось с использованием набора ExtractRNA (Евроген, Россия) по протоколу производителя. Качество и количество РНК определяли на спектрофотометре Smart Spec (Bio-Rad, США). Первую цепь кДНК синтезировали с помощью набора для обратной транскрипции MMLV RT kit (Евроген). Содержание мРНК оценивали методом ПЦР в режиме реального времени с интеркалирующим красителем SYBR Green на приборе Light Cycler 96 (Roche, Германия) с набором для ПЦР-РВ (Евроген). Для определения уровня экспрессии РНК каждую ПЦР проводили на двух независимых образцах кДНК. Последовательности праймеров для анализа экспрессии *CCA1* – f: 5'-GTGATGATGTTGAGGCAGATG -3', r: 5'-GAGAGCTTGAAGGCAATTG-3'.

Анализ относительного содержания транскриптов проводился с помощью метода  $2^{-\Delta\Delta Ct}$  [16], основанного на нормализации данных по экспрессии относительно двух референсных генов. Рассчитывалась разница значений  $C_t$  ( $\Delta C_t$ ) между целевым и референсными генами, затем сравнивались значения  $\Delta C_t$  контрольного и опытного образцов. В качестве референсных использованы гены *18sRNA* и *ACTIN8*, которые характеризуются конститтивной экспрессией. Последовательности праймеров референсных генов: *18sRNA* – f: 5'-TG-CCCGTTGCTCTGATGA-3', r: 5'-GGATGTGG-TAGCCGTTCT-3'; *ACTIN8* – f: 5'-GCAGACCG-TATGAGCAAAGAG-3', r: 5'-TGAGGGAAAGCAAG-GATAGAAC-3'. О специфичности фрагментов судили по кривым плавления.

#### Статистическая обработка данных

Экспериментальные данные обрабатывали с использованием статистических программ Microsoft Excel и Statgraphics 2.1 (ANOVA). Исследования выполнены на научном оборудовании Центра коллективного пользования Федерального исследовательского центра “Карельский научный центр Российской академии наук”.

### РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

#### Анализ транскрипционной активности гена *CCA1*

*A. thaliana* относится к длиннодневным растениям, т. е. ему требуется продолжительность светового дня не менее 12 ч (14–16 ч) [17]. В связи с этим для изучения транскрипционной активности *CCA1* растения выращивали в условиях длинного дня: 16 ч день и 8 ч ночь. В условиях естественного

светового фотопериода (16L : 8D) анализ проводили на 20-дневных растениях *A. thaliana* популяции Шуйская и линии *Ler*. Свет включался в 6 ч. Результаты представлены на графике (рис. 2). Динамика экспрессии *CCA1* в обеих группах растений оказалась сходной: фаза пришла на одно и то же время – 7 ч, т. е. максимальная амплитуда экспрессии гена достигалась уже через час после светового стимула и была примерно одинакова в обеих группах.

Изучение динамики транскрипционной активности гена *CCA1* растений *A. thaliana* в условиях инвертированного светового режима (8D : 16L) проводилось на растениях популяции Шуйская. Анализировали три группы растений: первая – это 30-тидневные растения, которые только последние десять дней росли в инвертированных условиях; вторая и третья группы – это 20-ти и 30-тидневные растения, выросшие в инвертированных световых условиях из семян. Результаты представлены на двух диаграммах: на одной – экспрессия *CCA1* анализировалась, когда свет предоставлялся в вечернее время с 17 ч (рис. 3), а на другой – свет продолжался в утренние часы с 5 до 9 ч, что соответствует началу светового периода в естественных условиях (рис. 4).

Результаты показали, что у растений, выращиваемых в условиях естественного светового режима 20 дней и перенесенных в условия инвертированного светового режима на 10 дней (Ш30), происходит резкий скачок экспрессии *CCA1* со сдвигом фазы на 2 часа (рис. 3, a) относительно группы растений, выросших в условиях 16L : 8D (Ш20) (рис. 2). Максимальная амплитуда экспрессии *CCA1* пришла на 20 ч, т. е. через 3 ч

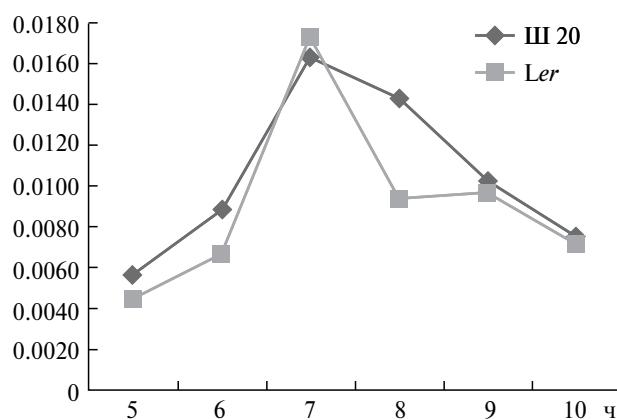
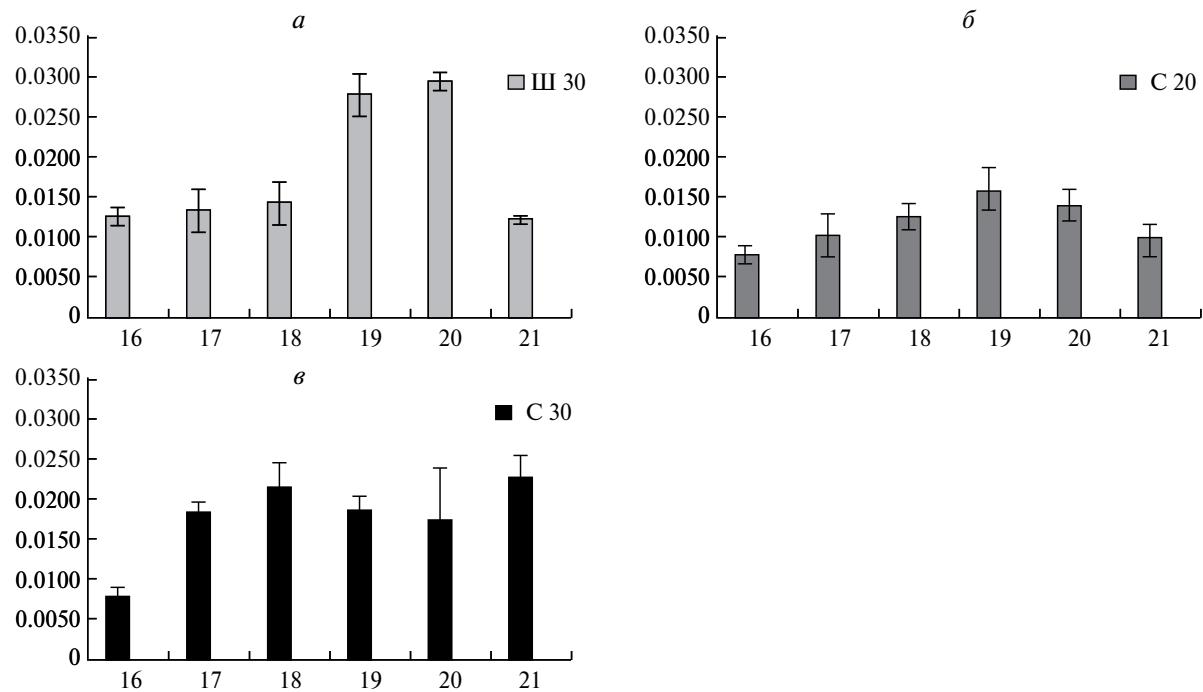
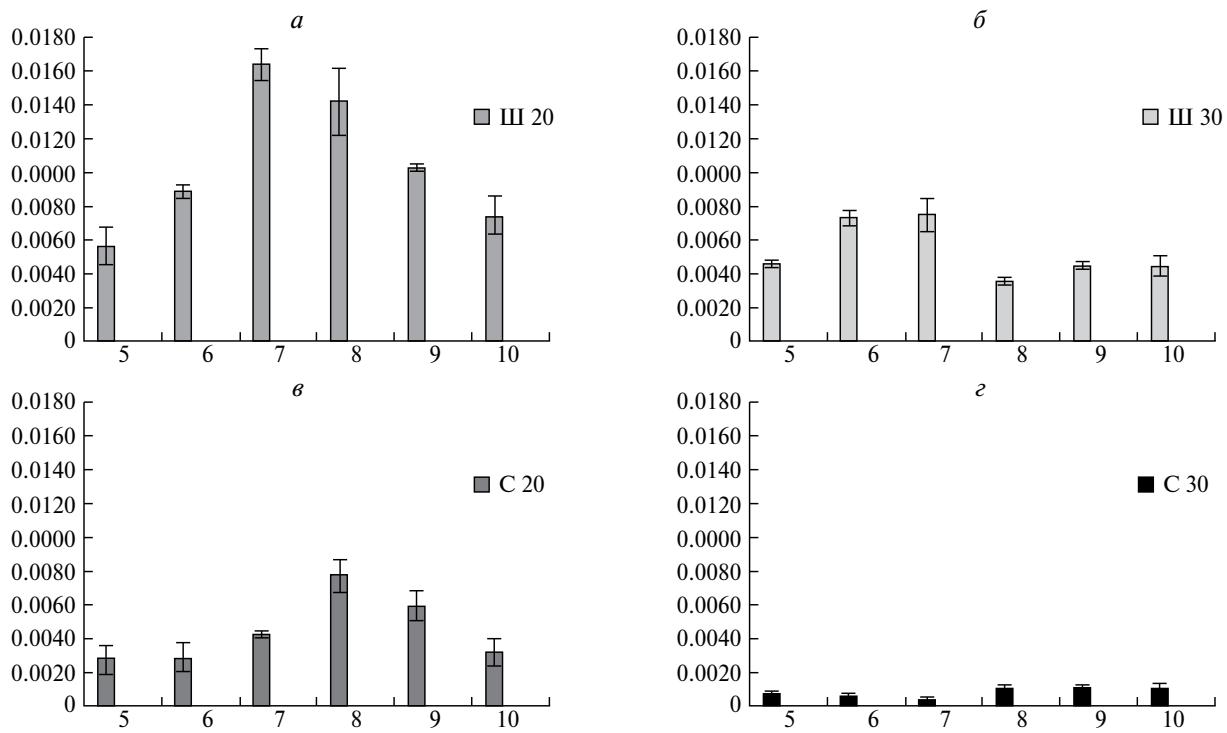


Рис. 2. Динамика транскрипционной активности *CCA1* у 20-дневных растений *A. thaliana* популяции Шуйская (Ш20) и *Ler* при выращивании их в условиях естественного светового фотопериода 16L : 8D (свет включался в 6 ч). Примечание. Здесь и на рисунках 3 и 4: по оси X – время суток в часах; по оси Y – относительный уровень транскриптов *CCA1*.



**Рис. 3.** Динамика транскрипционной активности *CCA1* у растений *A. thaliana* популяции Шуйская при выращивании их в условиях инвертированного светового режима 8D : 16L (свет включался в 17 ч). *a* – 30-дневные растения только последние 10 дней росли в инвертированных световых условиях (Ш30); *б* и *в* – 20-дневные (С20) и 30-дневные (С30) растения, постоянно выращиваемые при инвертированном световом режиме.



**Рис. 4.** Динамика транскрипционной активности *CCA1* у растений *A. thaliana* популяции Шуйская при выращивании их в условиях инвертированного светового режима 8D : 16L (свет продолжается с 17 ч до 9 ч утра). *а* – растения росли в условиях естественного светового фотопериода (16L : 8D) (Ш20) (данные приведены из рис. 1 для сравнения); *б* – 30-дневные растения только последние 10 дней росли в инвертированных световых условиях (Ш30); *в* – 20-дневные (С20) и 30-дневные (С30) растения, постоянно выращиваемые при инвертированном световом режиме.

после светового стимула, а у растений в условиях естественного светового режима — через один час (6–7 ч). Таким образом, в данных условиях происходит задержка транскрипционной активности гена, однако амплитуда экспрессии *CCA1* указанных растений оказалась почти в два раза выше по сравнению с растениями, выросшими в условиях естественного светового режима. По-видимому, это связано с более активным откликом на свет после 10-дневного роста в инвертированных световых условиях. Сдвиг фазы на 2 ч отражает потребность в дополнительном времени для адаптивной перестройки молекулярно-генетических процессов, участвующих в контроле циркадных ритмов. Когда продолжительность освещения достигла утренних часов суток, после 6 часов происходит увеличение транскрипционной активности *CCA1* с максимальной амплитудой в 7 ч, но примерно в два раза меньшей (рис. 4, б), чем у растений, выросших в условиях естественного светового периода (рис. 4, а). Это служит доказательством сохранения эндогенного циркадного ритма *CCA1*.

Для определения реакции растений, сразу после прорастания оказавшихся в инвертированных световых условиях, был предпринят следующий эксперимент: посевные семена сразу поместили в условия 8D : 16L на 20 и 30 дней (рис. 3). Стояла задача выяснить, как изменяется циркадный фенотип (который визуализируется с помощью количественных показателей — фазы и амплитуды) и как на это влияет возраст растений, т. е. длительность пребывания в инвертированных условиях.

Результаты анализа транскрипционной активности *CCA1* в вечернее время показали, что у растений C20 (рис. 3, б) фаза (максимальная экспрессия) достигалась через два часа после светового стимула в 17 ч (т. е. в 19 ч), что немного быстрее, по сравнению с 30-дневными растениями (Ш30) в таких же условиях (рис. 3, а), но амплитуда почти точно соответствовала амплитуде 20-дневных растений Ш20 (рис. 4, а), выросших в условиях естественного светового режима. Таким образом, за 20 дней роста в условиях инвертированного светового режима растения смогли к нему адаптироваться. Растения, выращиваемые в таких же условиях 30 дней (C30), продемонстрировали активный отклик на свет: фаза экспрессии *CCA1* пришла на первый час после светового стимула (на 18 ч) и амплитуда примерно в 1.5 раза выше (рис. 3, в) по сравнению с растениями C20 (рис. 3, б). Далее, в утренние часы у растений C20 в инвертированных световых условиях фаза транскрипционной активности *CCA1* была достигнута через 2 ч (в 8 ч) (рис. 4, в). По сравнению с растениями, выросшими в естественных световых условиях (Ш20; рис. 4, а), происходит задержка фазы на 1 ч, но с меньшей (в два раза) амплитудой. Такая же амплитуда у растений Ш30 наблюдается в подобных условиях

в утренние часы (рис. 4, б), что свидетельствует о сохранении эндогенного ритма экспрессии *CCA1* у растений C20 в инвертированных световых условиях. 30-дневные растения (C30) в аналогичных условиях показали в утренние часы очень низкую экспрессию *CCA1* (рис. 4, г). Такая низкая экспрессия указывает на потерю эндогенного ритма гена. Возможно, это связано с возрастом растений и/или является результатом адаптации к необычным световым условиям.

Таким образом, настоящее исследование показало, что искусственно измененные световые условия влияют на циркадные ритмы растений *A. thaliana* северной природной популяции. Были зафиксированы следующие реакции на инвертированный световой режим:

- резкий подъем транскрипционной активности одного из ключевых генов циркадной сети *CCA1* на световой стимул;
- сдвиг фазы экспрессии этого гена (при различных условиях выращивания растений), указывающий на необходимость определенного периода времени для молекулярно-генетических процессов адаптации;
- сохранение эндогенного ритма экспрессии *CCA1* у растений в течение 10-ти дней (Ш30) и 20-ти дней (C20) роста в инвертированных световых условиях и потеря его у 30-тидневных растений (C30) в этих же условиях.

Последний из указанных результатов в какой-то мере согласуется с мнением A. Venkat и S. Muneer [2], которые считают, что при снятии внешних сигналов дня и ночи эндогенный ритм экспрессии генов часов становится свободно текущим и через некоторое время перестает совпадать с суточными изменениями.

Исследований, подобных нашему, когда светлый и темный период меняются местами в суточном ритме, в литературе не встречается. В отличие от работ других авторов, в которых использовано существенное изменение длительности светлого и темного периодов [9, 14], в настоящем исследовании естественный фотопериод длинного дня сохранен. Так, H. Sugiyama с соавт. [9] изучали влияние ультракоротких фотопериодов (3L : 3D; 2L : 2D; 1L : 1D) на растения *A. thaliana*, а S. Nitschke с соавт. [14] подвергали этот вид продолжительному световому воздействию (32 ч) с последующим длительным периодом темноты (16 ч). Оказалось, что такое изменение фотопериода в обоих случаях приводит к циркадному стрессу и влияет на физиологические показатели растений и экспрессию генов. В работе [9] показано, что значительное сокращение фотопериода приводит к снижению экспрессии флоригена FT и задержке цветения растений. В исследовании S. Nitschke с соавт. [14] циркадный стресс вызвал подавление интенсивности

фотосинтеза, гибель клеток листьев и снижение транскрипционной активности *CCA1* и *LHY*. Также выращивание длиннодневного растения *A. thaliana* на коротком дне приводит к задержке цветения [18], а на постоянном свету применяется как стрессовый фактор [19]. В нашем исследовании природный фотопериод длинного дня, свойственный растениям *A. thaliana*, сохранялся, и это привело к усилению транскрипционной активности *CCA1* в ответ на световой стимул в необычное время.

Таким образом, результаты настоящего исследования показали, что радикальное изменение светового режима при сохранении длины дня не приводит к стрессовым реакциям, таким как снижение экспрессии *CCA1*. Наоборот, наблюдается процесс адаптации к новым световым условиям. Полученные результаты позволяют заключить, что вероятно, циркадные ритмы *A. thaliana* северных природных популяций выполняют важную роль в адаптации к изменению светового режима, и что один из ключевых генов часов *CCA1* играет в этом процессе существенную роль.

Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (№ темы FMEN-2022-0009).

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

- Zhu Z., Quint M., Anwer M.U. *Arabidopsis EARLY FLOWERING 3* controls temperature responsiveness of the circadian clock independently of the evening complex // J. Exp. Bot. 2022. V. 73. № 3. P. 1049–1061.  
<https://doi.org/10.1093/jxb/erab473>
- Venkat A., Muneer S. Role of circadian rhythms in major plant metabolic and signaling pathways // Front. in Plant Sci. 2022. V. 13.  
<https://doi.org/10.3389/fpls.2022.836244>
- Ueda H.R. Systems biology flowering in the plant clock field // Mol. Syst. Biol. 2006. V. 2(60).  
<https://doi.org/10.1038/msb4100105>
- Yamashino T., Ito S., Niwa Y. et al. Involvement of *Arabidopsis* clock-associated pseudo-response regulators in diurnal oscillations of gene expression in the presence of environmental time cues // Plant Cell Physiol. 2008. V. 49(12). P. 1839–1850.  
<https://doi.org/10.1093/pcp/pcn165>
- Ronald J., Davis S.J. Making the clock tick: The transcriptional landscape of the plant circadian clock // F1000Res. 2017. V. 6(951).  
<https://doi.org/10.12688/f1000research.11319.1>
- Flis A., Fernández A.P., Zielinski T. et al. Defining the robust behaviour of the plant clock gene circuit with absolute RNA timeseries and open infrastructure // Open Biol. 2015. V. 5(10).  
<https://doi.org/10.1098/rsob.150042>
- Linde A.M., Eklund D.M., Kubota A. et al. Early evolution of the land plant circadian clock // New Phytol. 2017. V. 216. P. 576–590.  
<https://doi.org/10.1111/nph.14487>
- Suárez-López P., Wheatley K., Robson F. et al. *G. CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis* // Nature. 2001. V. 410(6832). P. 1116–1120.  
<https://doi.org/10.1038/35074138>
- Sugiyama H., Natsui Y., Hara M. et al. Late flowering phenotype under ultra-short photoperiod (USP) in *Arabidopsis thaliana* // Plant Biotechnology. 2014. V. 31(1). P. 29–34.  
<https://doi.org/10.5511/plantbiotechnology.13.1104a>
- Rees H., Joynson R., Brown J.K.M. et al. Naturally occurring circadian rhythm variation associated with clock gene loci in Swedish *Arabidopsis* accessions // Plant Cell Environ. 2021. V. 44. P. 807–820.  
<https://doi.org/10.1111/pce.13941>
- Anwer M.U., Davis A., Davis S.J., Quint M. Photoperiod sensing of the circadian clock is controlled by *EARLY FLOWERING 3* and *GIGANTEA* // Plant J. 2020. V. 101(6). P. 1397–1410.  
<https://doi.org/10.1111/tpj.14604>
- Sawa M., Kay S.A. *GIGANTEA* directly activates *FLOWERING LOCUS T* in *Arabidopsis thaliana* // Proc. Natl Acad. Sci. USA. 2011. V. 108. P. 11698–11703.  
<https://doi.org/10.1073/pnas.1106771108>
- Salathia N., Davis S.J., Lynn J.R. et al. *FLOWERING LOCUS C*-dependent and -independent regulation of the circadian clock by the autonomous and vernalization pathways // BMS Plant Biol. 2006. V. 6. № 10.  
<https://doi.org/10.1186/1471-2229-6-10>
- Nitschke S., Cortleven A., Iven T., Feussner I. et al. Circadian stress regimes affect the circadian clock and cause jasmonic acid-dependent cell death in ctoxin-deficient *Arabidopsis* plants // Plant Cell. 2016. V. 28(7). P. 1616–1639.  
<https://doi.org/10.1105/tpc.16.00016>
- Иванов В.И., Касьяnenко А.Г., Санина А.В. и др. Краткая характеристика *A. thaliana* и некоторые сведения о его культивировании, технике скрещиваний и учете изменчивости // Генетика. 1966. Т. 8. № 1. С. 115–120.
- Livak K.J., Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2– $\Delta\Delta Ct$  method // Methods. 2001. V. 25. P. 402–408.  
<https://doi.org/10.1006/meth.2001.1262>

17. Квятко К.В., Мюллер А. Новый объект для генетических исследований – *Arabidopsis thaliana* (L.) Heynh. // Исследования по генетике. Т. 1. Л.: Изд-во ЛГУ. 1961. С. 79.
18. Fujiwara S., Oda A., Yoshida R. et al. Circadian clock proteins LHY and CCA1 regulate SVP protein accumulation to control flowering in *Arabidopsis* // Plant Cell. 2008. V. 20(11). P. 2960–2971.  
<https://doi.org/10.1105/tpc.108.061531>
19. Millar J., Carrington J.T., Tee W. et al. Changing planetary rotation rescues the biological clock mutant *lhy cca1* of *Arabidopsis thaliana* // bioRxiv. 2015.  
<https://doi.org/10.1101/034629>

## Transcriptional Activity of *CCA1* in Northern Population *Arabidopsis thaliana* Plants under Altered Light Conditions

M. V. Zaretskaya<sup>1,\*</sup>, O. M. Fedorenko<sup>1,\*\*</sup>

<sup>1</sup> Institute of Biology of Karelian Research Centre Russian Academy of Sciences,  
Petrozavodsk, 185910 Russia  
 \*e-mail: genmg@mail.ru  
 \*\*e-mail: fedorenko\_om@mail.ru

The dynamics of the transcriptional activity of one of the key genes of the circadian network, *CCA1*, was analyzed under conditions of a natural light photoperiod of a long day (16L : 8D) and under an inverted light regime (8D : 16L) in *A. thaliana* plants of the northern natural population (Karelia). It has been shown that under conditions of an inverted shift in the light regime, there is a sharp increase in the expression of this gene with a phase shift in the circadian rhythm by 2 hours. The level of *CCA1* transcriptional activity was almost two times higher compared to the natural light conditions. At the same time, the endogenous rhythm of the gene was preserved, but with a smaller amplitude. With age, 30-day-old plants grown under inverted conditions experienced a loss of endogenous *CCA1* circadian rhythm. The results obtained allow us to conclude that the circadian rhythms of *A. thaliana*, northern natural populations, probably play an important role in adaptation to changing light conditions, and that one of the key clock genes, *CCA1*, plays a significant role in this process.

**Keywords:** *Arabidopsis thaliana*, northern natural populations, circadian rhythms, dynamics of *CCA1* transcriptional activity, photoperiod.

## ГЕНЫ ПАРТЕНОГЕНЕЗА КУКУРУЗЫ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ МУТАЦИЙ

© 2024 г. Е. М. Моисеева<sup>1</sup>, В. В. Фадеев<sup>1</sup>, Ю. В. Фадеева<sup>1</sup>, С. И. Мазилов<sup>1</sup>,  
А. Ю. Колесова<sup>2</sup>, М. И. Чумаков<sup>1,\*</sup>

<sup>1</sup>Институт биохимии и физиологии растений и микроорганизмов, Федеральный исследовательский центр  
“Саратовский научный центр Российской академии наук”, Саратов, 410049 Россия

<sup>2</sup>Федеральный аграрный научный центр Юго-Востока, Саратов, 410010 Россия

\*e-mail: chumakov\_m@ibppr.ru

Поступила в редакцию 30.05.2024 г.

После доработки 31.07.2024 г.

Принята к публикации 06.08.2024 г.

Приведен анализ полиморфизма нуклеотидных последовательностей генов, предположительно связанных с партеногенетическим развитием зародыша и эндосперма у кукурузы. С помощью секвенирования и последующего множественного выравнивания транскриптов целевых генов (*Hdt104*, *Chr106*, *Fie1* и *Fie2*) исследуемых в работе партеногенетических линий AT-1, AT-3 и AT-4 и референсной линии кукурузы B73 определено наличие однонуклеотидных замен, делеций и вставок, построены филогенетические деревья по изучаемым генам.

**Ключевые слова:** партеногенез, кукуруза, мутации, полиморфизм генов.

**DOI:** 10.31857/S0016675824120057 **EDN:** WANBDA

Половое размножение является доминирующим, но не единственным способом размножения у покрытосеменных растений. Бесполое размножение семенами (апомиксис) существует у многих видов растений, включая злаки [1], и является запасным вариантом размножения для растений в отсутствие опылителя или при расхождении сроков созревания яйцеклетки, центральной клетки и спермииев в неблагоприятных погодных условиях. При апомиксисе эмбрион и эндосперм начинают развиваться спонтанно, без опыления.

Апомиксис у растений возникает в результате переключения с полового размножения на бесполое при запуске каскада генов (программы) в условиях отсутствия опыления [2–5]. За последние годы получен ряд данных, касающихся сравнительной экспрессии генов метилирования ДНК у современных сортов кукурузы с исключительно половым размножением и апомиктических гибридов кукурузы и трипсакума [6–8].

Апомиксис, как правило, не наблюдается у культурных форм растений, поскольку традиционная селекция культурных растений на протяжении тысячелетий проводилась человеком путем гибридизации и отбора (т. е. путем полового размножения). Например, дикий предок кукурузы трипсакум (*Tripsacum dactyloides* L.) и гибриды кукурузы и трипсакума способны к апомиктическому способу

размножения [1, 2, 9]. А кукуруза (*Zea mays* L.) в ходе селекции человеком в течение нескольких тысяч лет приобрела ряд важных агрономических свойств, но утеряла некоторые признаки дикого предка, в частности способность к бесполому размножению [10].

Предполагается, что апомиктические формы растений могут репродуцироваться непосредственно производителями сельскохозяйственной продукции без помощи семеноводческих хозяйств, поскольку не требуется получения гибридов, что значительно снижает затраты на получение, хранение, поддержание, транспортировку семенного материала, сокращает зависимость от зарубежных производителей семян.

Путем скрещиваний и селекции во второй половине XX и начале XXI в. в Саратове (Россия) были получены уникальные линии, у которых зафиксированы элементы апомиксиса: независимое от опыления начало развития эмбриона и эндосперма (автономный матроклиний партеногенез). В частности, саратовскими селекционерами около 40 лет назад была получена линия кукурузы AT-1 [11] и ее производные AT-3 [12], ATT [13], ATTМ [14], у которых автономные (без опыления) эмбрио- и эндоспермогенез наблюдались с повышенной (6–50% и более) частотой.

Линия АТ-1 выделена в самоопыленном потомстве гибрида, полученного после скрещивания линий Stock 6 и Коричневый тестер [11]. Для линии АТ-1 характерен наследуемый тип партеногенеза с независимым от опыления эндоспермогенезом, полизибрионией, образованием гаплоидов в потомстве. Этот признак контролируется генами, расположенными в ядре растительной клетки. Но линия АТ-1 плохо вызревала в условиях Юго-Востока России. У более скороспелой линии АТ-3 партеногенетические зародыши развивались из неопыленных яйцеклеток через 7–10 дней после появления пестичных рыльцев [12]. У тетраплоидной линии АТТ (исходной для диплоидной линии АТ-4) было обнаружено 0,6% зародышевых мешков с развивающимися зародышами (от двухклеточных до глобулярных (93 клетки)) [13]. На данный момент известен ряд генов, которые связывают с независимым от опыления началом развития эмбриона и эндосперма.

Кандидатами для изучения генов, связанных с независимым от опыления развитием зародыша и эндосперма у кукурузы, являются гены с различным уровнем экспрессии при половом и апомиктическом развитии [8, 15–18].

Установлено, что подавление экспрессии генов, кодирующих хроматин-модифицирующие белки (**ХМБ**) (*Chr106*, *Hdt104*, *Hon101*), и генов, кодирующих ДНК-метилтрансферазы (*Dmt102*, *Dmt103*, *Dmt105*), коррелирует с признаками апомиксиша у гибридов кукурузы и трипсакума [8, 19]. Уровни экспрессии ХМБ у партеногенетических линий (АТ-3 и АТ-4) и обычных линий кукурузы (ГПЛ-1) существенно не различались, за исключением генов *Hon101* и *Hdt104* [20]. Уровни экспрессии *Hon101* и *Hdt104* были в 3 и 1,4 раза ниже соответственно в зародышевых мешках линии АТ-3 по сравнению с таковыми у линии ГПЛ-1. Кроме того, уровень экспрессии гена *Hon101* был в 2 раза ниже в эндосперме линии АТ-3, чем в эндосперме линии ГПЛ-1 [20]. Гены метилирования ДНК (*Dmt103*, *Dmt105*) и гены, кодирующие ферменты, модифицирующие хроматин (*Chr106*, *Hdt104*, *Hon101*), имели значительно более высокий уровень экспрессии в партеногенетических зародышах линии АТ-4 по сравнению с неопыленными зародышевыми мешками [21]. Изменение экспрессии генов *Hdt104*, *Hon101* может быть связано с переходом к партеногенетическому развитию у линий АТ.

Гены, контролирующие у кукурузы развитие эндосперма (*fertilization independent endosperm*, *Fie*), входят в Polycomb Repressive Complex 2 (PRC2) [22]. PRC2 представляет собой четырехсубъединичный гистон-метилтрансферазный комплекс, который катализирует trimетилирование лизина 27 на гистоне Н3 (H3K27me3), что способствует изменению структуры хроматина и длительной репрессии генов [22]. У *Arabidopsis thaliana* (резуховидки Таля)

идентифицированы три гена, которые контролируют независимое от оплодотворения развитие семян: *FIS1/MEDEA*, *FIS2* и *FIS3/FIE* [23–25]. В отличие от резуховидки Таля, имеющей единственный ген *Fie*, у кукурузы [15] и риса [26] обнаружены два гена *Fie*. Для мутантов *OsFie2* риса показано автономное образование зародышеподобной структуры в отсутствие оплодотворения, а для двойных мутантов *OsFie1/OsFie2* – автономное образование эмбриона и эндосперма с высокой частотой [27]. Аналогичные структуры наблюдались и у партеногенетических линий кукурузы АТ [11, 28].

У кукурузы и риса *Zm\_Fie2* является, вероятно, препрессором развития эндосперма до опыления [15, 16, 18]. Экспрессия *Zm\_Fie1* зарегистрирована исключительно в центральной клетке кукурузы после опыления [15, 16, 18]. У партеногенетической линии кукурузы АТ-4 в 2021 г. нами был впервые зафиксирован необычный характер экспрессии генов *ZmFie1* и *ZmFie2* по отношению к непартеногенетическим линиям, наряду с экспрессией некоторых генов, контролирующих хроматин-модифицирующие белки, что может явиться причиной спонтанного (без опыления) развития зародыша и эндосперма [21]. В дальнейшем предстоит выяснить, какие гены связаны с наследуемым партеногенезом у линий кукурузы саратовской селекции и какие изменения нуклеотидной последовательности генов и функционирования кодируемых белков привели к этому.

## МАТЕРИАЛЫ И МЕТОДЫ

### Растения

Партеногенетическая линия кукурузы АТ-1 была выделена в самоопыленном потомстве гибрида линии-гаплоиднуктора Stock 6 (США) [29] и линии Коричневый тестер (США) [11]. Линия АТ-3 была получена путем скрещивания линии АТ-1 со скороспелой линией саратовской селекции ГПЛ-1 (дигаплоидизированный гаплоид, Саратов) [28]. Партеногенетическая линия кукурузы АТТ была получена путем скрещивания линии АТ-1 с тетраплоидной линией Кр-П-1 (Краснодар) и последующим отбором диплоидной формы гибрида из тетраплоидной партеногенетической линии АТТ [13, 30]. В потомстве линии АТТ среди 850 тетраплоидных растений было обнаружено одно диплоидное растение, отличавшееся от тетраплоидов меньшей высотой и более ранним цветением. В течение нескольких лет проводился отбор растений с наибольшим количеством двоен и гаплоидов, а также с высокой частотой автономных зародышей, в результате чего была получена линия АТ-4. Для всех линий АТ характерны такие элементы апомикса, как независимое от опыления развитие зародыша

**Таблица 1.** Праймеры для секвенирования генов партеногенеза кукурузы

Название гена	Праймер	Последовательность нуклеотидов 5' → 3'
<i>Chr106</i>	seq1-1-F	TCTCGGCCTGTTCCCTTCTAT
	seq1-1-2-R	GAAGCTGGTCTTCGTCTTG
	seq1-2-F	GACGAAGACCAGCTTCTGGA
	seq1-2-R	CGACTTCAACTTCCCACCACT
	seq2-F	AAAAGGCCAAGACAGCAGTG
	seq2-R	TCCCAGTCAAAAGGAGCTTG
	seq3-F	ATTGAGGGAGATGAAGCGCA
	seq3-R	TGCGTGCAAGTAGGGAAGTT
	seq4-F	TGGAACAACAAGCCTTATCC
	seq4-R	GCGTTAGGCTTGCTCTTC
<i>Hdt104</i>	seq6-F	GTGGTGATAGGCAAGGGACA
	seq6-R	AACCGATAAAGCATGGCAGT
	seq1-F	CAATGGAGTTCTGGGGTGAA
	seq1-R	TCTTCTTGCCAACTAACCTCA
	rt-F1	ACCCCTAAACCCAGCACCAG
	seq2-F	TGGTGATGATGATGATTCACTG
	seq2-R	TTTGGGGACTTCTTGTCAAG
	seq3-F	CAACTTCTCATCCTGCAAAGC
	seq3-R	GGCCTTACGGGTACAATAA
<i>Fie1</i>	seq1-F	CCGCCACCATATAGAACAC
	seq1-R	AAATTGAACCCGATGGCATA
	seq2-F	TATGCCATCGGGTTCAATT
	seq2-R	CAGCCCTGCAAAGACTAAG
	seq3-F	CTGCCAGCAAGGATGAATCT
	seq3-R	GGACTGCACTTCCCAGACAT
	seq4-F	CAGATGGCAATAGGCAACAA
	seq4-R	TCCACGGTGCAAATTACAA
<i>Fie2</i>	seq1-F1	GCGGAACCGGAAATCTTG
	seq1-R1	AGATTCATCCTGCTGCAGA
	seq1-F2	GCCATGGCGACTCAATAAT
	seq1-R2	TGGACTGTACTTCCCACACG
	seq1-F3	GCGATAGGCAACCGTGAA
	seq1-R3	TTTCCACCAACGGATTCAAC

до стадии глобулы, начало эндоспермогенеза, по-лиэмбриония, а также псевдогамия [31, 32].

#### *Определение нуклеотидной последовательности транскриптов генов*

РНК выделяли из завязей початков, материал замораживали и хранили при  $-70^{\circ}\text{C}$ . кДНК получали согласно инструкции производителя ревертазы (Евроген, Россия). Участки кДНК исследуемых генов амплифицировали с помощью набора реактивов Tersus Plus PCR kit (Кат. № PK121, Евроген) и специфических праймеров (Евроген) для получения перекрывающихся фрагментов последовательностей. Праймеры подбирали с помощью ресурса Primer-Blast (табл. 1).

ПЦР-продукты очищали при помощи набора Cleanup Mini (Кат. № BC023S, Евроген) и секвенировали в компаниях Евроген (Россия) и Синтол (Россия). Транскрипты генов *Chr106*, *Hdt104*, *Fie1* использованных в статье линий кукурузы были секвенированы и представлены в базе данных GenBank (см. Результаты и обсуждение).

#### *Методы биоинформатики*

Нуклеотидные последовательности кодирующих ХМБ и *Fie* генов партеногенетических линий кукурузы сравнивали между собой, с линиями, не обладающими данными признаками, и референсной линией B73 методом множественного выравнивания с использованием программ BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) и Clustal Omega (<https://www.ebi.ac.uk/jDispatcher/msa/clustalo>). Филогенетический анализ генов был проведен в программе MEGA 11 [33]. Эволюционные расстояния были вычислены с использованием метода максимального составного правдоподобия [34].

#### *Микроскопический анализ зародышевых мешков кукурузы*

Семяпочки кукурузы фиксировали и окрашивали, как описано в работе [20]. Анализ препаратов проводили под микроскопом AxioScope A1 (Karl Zeiss, Германия).

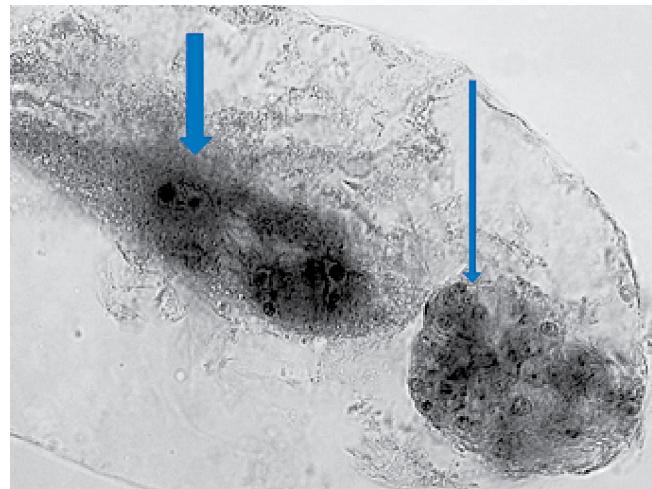
### РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для экспериментальной задержки опыления изоляцию рылец линии кукурузы AT-4 осуществляли пергаментными пакетами на 7–14 дней. У партеногенетической линии кукурузы AT-4 без опыления наблюдаются элементы апомиксиса, такие как независимое от опыления развитие зародыша до стадии глобулы, начало эндоспермогенеза (рис. 1).

Как видно из рис. 1, неопыленный (10 дней после изоляции рылец) зародышевый мешок (3М)

кукурузы линии AT-4 содержит развивающийся партеногенетический (36-клеточный) зародыш и 4-ядерный партеногенетический эндосперм. Эмбриологический анализ зародышевых мешков семи растений линии AT-4 (2021–2022 гг.) в отсутствие опыления (через 7–10 суток после появления рылец) показал, что автономные зародыши у них формируются с частотой 28–52% и содержат от двух до 120 клеток. У трех растений выявлено автономное развитие эндосперма, содержащего от двух до 16 ядер (1–8% зародышевых мешков). В нескольких мегагаметофитах наблюдалось одновременное развитие автономного зародыша и эндосперма (рис. 1). У растений исходной тетрапloidной линии ATT частота автономного эмбриогенеза была значительно ниже, чем у диплоидной линии AT-4, и составляла 0.5–5.9% [13].

Как правило, у линий AT развивающиеся партеногенетические зародыши и эндосперм начинают деградировать на 12–14 день после изоляции рылец. Полностью развившихся без опыления зародыша и эндосперма не удавалось наблюдать в многолетних исследованиях на большом статистическом материале (250–300 тыс. зародышевых мешков) [35]. Одной из возможных причин можно назвать несоответствие полидности эндосперма у спонтанно делящейся центральной клетки и отсутствие соответствующих сигнальных молекул для развивающегося зародыша. Но тем не менее у части 3М (0.5–5.9%) линии ATT происходит деление яйцеклетки и центральной клетки [13]. Примерно



**Рис. 1.** Зародышевый мешок (3М) кукурузы диплоидной линии AT-4, содержащий партеногенетические зародыши и эндосперм. Неопыленный 3М выделен через 10 дней после изоляции пестичных нитей. Полевой сезон 2022 г. Тонкой стрелкой показан партеногенетический (36-клеточный) зародыш; толстой стрелкой показан 4-ядерный партеногенетический эндосперм. Ув.  $\times 400$ , микроскоп – Axio Scope A1 (Karl Zeiss, Германия).

Chr106_B73	1	ATGGCTGGAAGGGATGCCAGTGCCCTGCCGAATGGGGCAAGCCGATGGGCCGCC	60
Chr106_KM	1	.....	60
Chr106_AT-4	1	.....A.....	60
Chr106_AT-3	1	.....	60
Chr106_B73	61	ACCGGCCGCCGCCCTAACGCCAGATTTCTCCCTAAATCCGTTCTGAGGACGAGAAA	120
Chr106_KM	61	.....	120
Chr106_AT-4	61	.....G.....	120
Chr106_AT-3	61	.....	120
Chr106_B73	121	GAATCCAAGAATGGGAACGCCCTCCITAGCCACGGAGCAGTTAACAGCAAGGGACGAGACC	180
Chr106_KM	121	.....	180
Chr106_AT-4	121	.....A.....A.....	180
Chr106_AT-3	121	.....	180
Chr106_B73	181	TCCGACGATTTCCTGCGATGCCATTGATGTCCCCTCCCTGACATGAAATCGACAGTGC	240
Chr106_KM	181	.....	240
Chr106_AT-4	181	.....A.....	240
Chr106_AT-3	181	.....	240
Chr106_B73	241	GATGTTGCCCTCTGCTGTTGAGGCAGCTAACAGACGAAGACCAGCTCTGGAGCCCTGCCAG	300
Chr106_KM	241	.....	300
Chr106_AT-4	241	.....	300
Chr106_AT-3	241	.....	300
Chr106_B73	301	GAGGAGAAATTGATGATTCTGGATGAGCTTATCTATAACCTATTGATCTTGAGGCC	360
Chr106_KM	301	.....	360
Chr106_AT-4	301	.....	360
Chr106_AT-3	301	.....	360
Chr106_B73	361	AAGAAATGGTGATGATCCCTGATCACAGAACGAAATGAAAAAGGAAGAACAGCTCGAG	420
Chr106_KM	361	.....	420
Chr106_AT-4	361	.....	420
Chr106_AT-3	361	.....	420
Chr106_B73	421	GAGGCCCGGATTAAAGCAAGGAAAGAAAGAGAACGCCAGGAAGAGCAGGAGAACTTGCAAAA	480
Chr106_KM	421	....A.....	480
Chr106_AT-4	421	.....	480
Chr106_AT-3	421	.....	480
Chr106_B73	481	CTTGGTTTTGATCTCTGAGTCGGGGTATAAGCAACTTAGATGAGCTACTGACAAAGCACAA	540
Chr106_KM	481	.....C.....	540
Chr106_AT-4	481	.....C.....	540
Chr106_AT-3	481	.....	540
Chr106_B73	541	CTATTTTCAGAGTTCCCTACTTGAGAAATGGATAAAATGCCGATGAAAGTTGTTGAAACCT	600
Chr106_KM	541	.....	600
Chr106_AT-4	541	.....T.....T.....	600
Chr106_AT-3	541	.....T.....T.....	600
Chr106_B73	601	CAAGGTGAAGAGTCACCAAGTAGAGAACAAAGAAAAGACGTGGCCCAAAGAGAAAGCTAA	660
Chr106_KM	601	.....	660
Chr106_AT-4	601	.....	660
Chr106_AT-3	601	.....	660
Chr106_B73	661	GCCACACCAAAATCAATGATAAAAAGGCCAAGACAGCAGCTGGCACCCATCTTACAAGA	720
Chr106_KM	661	.....	720
Chr106_AT-4	661	.....	720
Chr106_AT-3	661	.....	720
Chr106_B73	721	TCTCGTGAAGATGCTCTGCTGATGACTCTACTCTCTCGAGAAAGAAAGCTTCAAAAAAA	780
Chr106_KM	721	.....T.....	780
Chr106_AT-4	721	.....T.....	780
Chr106_AT-3	721	.....	780
Chr106_B73	1441	aaaagaagGGTTCATGTTGTTCAAAACGCTTCACTGCCATTTCGGCCATTCTCTGAGA	1500
Chr106_KM	1441	.....	1500
Chr106_AT-4	1441	.....C.....	1500
Chr106_AT-3	1441	.....	1500
Chr106_B73	1501	CGGATGAAAGAGCATGAGAACAAATCTTCCACGGAAGAAAGAGATTAATCTTATGCT	1560
Chr106_KM	1501	.....	1560
Chr106_AT-4	1501	.....	1560
Chr106_AT-3	1501	.....	1560
Chr106_B73	1561	AACATGACTGAGGTTCAAGGCCATTCCAGGATCACTTAGTTGAGAAGAACCTTTGATGTC	1620
Chr106_KM	1561	.....	1620
Chr106_AT-4	1561	.....	1620
Chr106_AT-3	1561	.....T.....	1620
Chr106_B73	1621	TACTTGAAATGATGAACTAGATA-TTG-TGTTGCCAGCTGGCATCAAGACAAAGCTAA	1678
Chr106_KM	1621	.....	1678
Chr106_AT-4	1621	.....	1678
Chr106_AT-3	1621	.....T.....TCT.TC...CC	1678
Chr106_B73	1679	ATAGTCCTTTATTCATAATTGAGGAAAGACTGCGGCCATCTCTGATCTTTGAAAGCTGCA	1738
Chr106_KM	1679	.....	1738
Chr106_AT-4	1679	.....	1738
Chr106_AT-3	1679	.....	1738
Chr106_B73	1739	TTGGAAACAAAGCCTTTATCCACCTGTTAATAAGCTTTAGACAAATGIGCAAATTTC	1798
Chr106_KM	1739	.....	1798
Chr106_AT-4	1739	.....	1798
Chr106_AT-3	1739	.....	1798
Chr106_B73	1799	CACAAATGGACAAAAGCTTTCGACACTTCCACTTGACGCAAGCATTAAGGTTCTAATTTCT	1858
Chr106_KM	1799	.....A.....	1858
Chr106_AT-4	1799	.....A.....	1858
Chr106_AT-3	1799	.....	1858
Chr106_B73	1859	CACAAATGGACAAAAGCTTTCGACACTTCCACTTGACGCAAGCATTAAGGTTCTAATTTCT	1918
Chr106_KM	1859	.....	1918
Chr106_AT-4	1859	.....	1918
Chr106_AT-3	1859	.....	1918
Chr106_B73	1919	TTTGCAGAAATTGATGCTGTTAATTGAGAACAGAGAGGGGGCGCAGATACACAGCTTTA	1978
Chr106_KM	1919	.....	1978
Chr106_AT-4	1919	.....A.....	1978
Chr106_AT-3	1919	.....	1977
Chr106_B73	1979	ATGATTTCGAAATAGCAGTCGATGCTTTATCTGAGCACACGGCTGGCGACTTGGTA	2038
Chr106_KM	1979	.....	2038
Chr106_AT-4	1979	.....	2038
Chr106_AT-3	1978	.....	2037
Chr106_B73	2039	TCAACCTTACTTCTGCCTGATACAACTGATCTTCTTATGACACTGACTGAAATCCCTCAGA/GG	2098
Chr106_KM	2039	.....	2098
Chr106_AT-4	2039	.....	2098
Chr106_AT-3	2038	.....	2097

**Рис. 2.** Выравнивание нуклеотидных последовательностей транскриптов гена *Zm\_Chromosome 106* линий КМ, AT-3 и AT-4 и референсной линии B73. Точками показано нуклеотидное сходство с последовательностью референсной линии B73; ОНЗ у линий КМ, ЗМС-8 и ЗМС-П обозначены буквами; дефис обозначает отсутствие нуклеотида, рамка обозначает вставку нуклеотида; на рисунке представлены фрагменты гена, в которых обнаружены мутации.

B73	61	ACTGGGACTGTTGCCAACTCGAAACCTGGCTCATCATCGACGAACGAGGGGAAGCAACAG	120
AT-1	61	.....C.....	T 120
AT-3	61	.....	120
AT-4	61	.....C.....	T 119
B73	481	TCACCGCTGTTGGGGCCGCAGGAAGCAATCGGATCATTCGGGTCAATTGTGCTACC	540
AT-1	481	.....T.....	540
AT-3	481	.....	540
AT-4	480	.....T.....A.....	539
B73	781	ACTGTGAAAATTGGTCAATGAAAGAATTGGATATAATGTTGAAAAATCATATTCAATGG	840
AT-1	781	.....G.....	840
AT-3	781	.....	840
AT-4	780	.....G.....	839
B73	841	ACTGCCATCCATCAAAGTTCCAACGAGGAATATCCAGTTCCGGTCTTGACTGCTGCA	900
AT-1	841	.....	900
AT-3	841	....A.....	900
AT-4	840	.....	899
B73	961	GTAAGAATGCACTTTGCTTTGGAACCAAAACAGACAAGCGTAGGCCCTGGGAGGGG	1020
AT-1	961	.....T.....	1020
AT-3	961	.....	1020
AT-4	960	.....	1019
B73	1021	ACTGTTGATGTTCTTCAGAAGTACCCGGTGCCTAAAGTGTTCATTATGGTTATGAAATT	1080
AT-1	1021	.....	1080
AT-3	1021	.....C.....	1080
AT-4	1020	.....	1079
B73	1321	CAAGCTGCTGCGCCCGCCGGTGTGGTGCCGGTGCCTGCGCACGCCGACGCCGAC	1380
AT-1	1321	.....	1380
AT-3	1321	.....	1374
AT-4	1320	.....	1379
B73	1381	GCCTCACCGACACGCCCTCCCCCCCCCGGTTACATCCATCTACTCCGTCTCGTT	1440
AT-1	1381	.....T.....	C. 1440
AT-3	1375	.....	1434
AT-4	1380	.....T.....	C. 1439
B73	1441	GATTAGCTTACCCATTGGTATGTTGGTCAGAGTCGCCAGATCTAGTGTGGCTGA	1500
AT-1	1441	....C.....G.....	1500
AT-3	1435	.....	1494
AT-4	1440	....C.....G.....	1499
B73	1501	ACGTTGAATGTTAGGATGCTGCTGTTGTTATGCTCTGAGTCTTGAGTTCACTTTGTTAA	1560
AT-1	1501	....C.....T.....	1560
AT-3	1495	.....	1554
AT-4	1500	....C.....T.....	1559
B73	1561	TTTGCACCGTGGATGAGATGAATACTTGACGTTGCAA	1598
AT-1	1561	.....	1598
AT-3	1555	.....T.....	1592
AT-4	1560	.....	1597

**Рис. 3.** Множественное выравнивание нуклеотидных последовательностей гена *Zm\_Fie1* линий кукурузы саратовской селекции и референсной линии B73. Точками показано нуклеотидное сходство с последовательностью референсной линии B73; ОНЗ у линий AT-1, AT-3, AT-4 обозначены буквами; дефис обозначает отсутствие нуклеотида; на рисунке представлены фрагменты гена, в которых обнаружены мутации.

такой же процент (до 6) спонтанных проэмбрионально-обнаруживается у партеногенетической линии AT-3 [30]. У линии AT-1, исходной для линий AT-3 и AT-4, частота появления спонтанных зародышей составляла 82%, а ядер эндосперма – 52% [36]. Что является триггером начала спонтанного деления яйцеклетки и центральной клетки, пока неизвестно.

#### *Zm\_Hdt104*

Сиквенсы гена *Zm\_Hdt104*, кодирующего гистоновую деацетилазу линий кукурузы AT-3 и AT-4, представлены нами в базе данных GenBank (MW222955.1 (AT-3) и OK557951.1 (AT-4)). Нуклеотидные последовательности транскрипта гена *Zm\_Hdt104* линий AT-4 и AT-3 сравнивали с последовательностью мРНК *Zm\_Hdt104* референсной линии B73 (NM\_001148367.1). У линии AT-4 однонуклеотидных замен (ОНЗ) обнаружено не было, у линии AT-3 выявлено шесть ОНЗ. Как влияют

обнаруженные ОНЗ на функцию белка HDT104 у линии AT-3, пока неизвестно.

#### *Zm\_Ch106*

Ген кукурузы *Zm\_Ch106* кодирует белок, который участвует в ремоделировании хроматина. Транскрипты гена *Chr106* кукурузы секвенированы и представлены нами в базе данных GenBank (MW441336.1 – для линии KM и OL649771.1 – для линии AT-4). Полученные нуклеотидные последовательности транскрипта гена *Zm\_Ch106* линий KM, AT-3 и AT-4 сравнивали с нуклеотидной последовательностью мРНК *Zm\_Ch106* референсной линии B73. Количество ОНЗ на 1000 нуклеотидов у гена *Zm\_Ch106* у разных линий колеблется от 2.7 до 6.8. В белок-кодирующей части гена *Zm\_Ch106* у линии KM обнаружены семь ОНЗ, у линии AT-3 – десять ОНЗ и две делеции, у линии AT-4 – 15 ОНЗ (рис. 2). Как влияют обнаруженные ОНЗ на функцию белка CHR106 у линий AT, пока неизвестно.

<i>Fie2_B73</i>	TGGGAACCGAAGACAAAAGAACAGAGTCCTGGGGAGGGAAGCATCGATATCCTTCAGAAG	1260
<i>Fie2_AT-1</i>	TGGGAACCGAAGACAAAAGAACAGAGTCCTGGGGAGGGAAGCATCGATATCCTTCAGAAG	1260
<i>Fie2_AT-3</i>	TGGGAACCGAAGACAAAAGAACAGAGTCCTGGGGAGGGAAGCATCGATATCCTTCAGAAG	1260
<i>Fie2_AT-4</i>	TGGGAACCGAAGACAAAAGAACAGAGTCCTGGGGAGGGAAGCATCGATATCCTTCAGAAG	1260
	*****	*****
<i>Fie2_B73</i>	TATCCTGTCCCAGAAATGTGACATTGGTTTATCAAAATTTCATGTGATTTCACTTCAT	1320
<i>Fie2_AT-1</i>	TATCCTGTCCCAGAAATGTGACATTGGTTTATCAAAATTTCATGTGATTTCACTTCAT	1320
<i>Fie2_AT-3</i>	TATCCTGTCCCAGAAATGTGACATTGGTTTATCAAAATTTCATGTGATTTCACTTCAT	1320
<i>Fie2_AT-4</i>	TATCCTGTCCCAGAAATGTGACATTGGTTTATCAAAATTTCATGTGATTTCACTTCAT	1320
	*****	*****
<i>Fie2_B73</i>	CAGTTGGCGATAAGGCACCCGTGAAGGCAAATCTACGTGTGGGAAGTACAGTCCAGCCCT	1380
<i>Fie2_AT-1</i>	CAGTTGGCGATAAGGCACCCGTGAAGGCAAATCTACGTGTGGGAAGTACAGTCCAGCCCT	1380
<i>Fie2_AT-3</i>	CAGTTGGCGATAAGGCACCCGTGAAGGCAAATCTACGTGTGGGAAGTACAGTCCAGCCCT	1380
<i>Fie2_AT-4</i>	CAGTTGGCGATAAGGCACCCGTGAAGGCAAATCTACGTGTGGGAAGTACAGTCCAGCCCT	1380
	*****	*****
<i>Fie2_B73</i>	CCTGCTCTCATTGCTCGGCTGTATAATCAGCAGTGTAAATCGCCGATAAGACAAACTGCA	1440
<i>Fie2_AT-1</i>	CCTGCTCTCATTGCTCGGCTGTATAATCAGCAGTGTAAATCGCCGATAAGACAAACTGCA	1440
<i>Fie2_AT-3</i>	CCTGCTCTCATTGCTCGGCTGTATAATCAGCAGTGTAAATCGCCGATAAGACAAACTGCA	1440
<i>Fie2_AT-4</i>	CCTGCTCTCATTGCTCGGCTGTATAATCAGCAGTGTAAATCGCCGATAAGACAAACTGCA	1440
	*****	*****
<i>Fie2_B73</i>	GTGTCCTTCGATGGAAGCACAACTCCTGGAGCTGGTGAAGACGGCACCATCTGGCGGTGG	1500
<i>Fie2_AT-1</i>	GTGTCCTTCGATGGAAGCACAACTCCTGGAGCTGGTGAAGACGGCACCATCTGGCGGTGG	1500
<i>Fie2_AT-3</i>	GTGTCCTTCGATGGAAGCACAACTCCTGGAGCTGGTGAAGACGGCACCATCTGGCGGTGG	1500
<i>Fie2_AT-4</i>	GTGTCCTTCGATGGAAGCACAACTCCTGGAGCTGGTGAAGACGGCACCATCTGGCGGTGG	1500
	*****	*****
<i>Fie2_B73</i>	GATGAAGTGGACCACATCGAGCTCCAGAAACTGAAGAAGTGTGCGCGCTCAATGCTGGACT	1560
<i>Fie2_AT-1</i>	GATGAAGTGGACCACATCGAGCTCCAGAAACTGAAGAAGTGTGCGCGCTCAATGCTGGACT	1560
<i>Fie2_AT-3</i>	GATGAAGTGGACCACATCGAGCTCCAGAAACTGAAGAAGAAGTGTGCGCGCTCAATGCTGGACT	1560
<i>Fie2_AT-4</i>	GATGAAGTGGACCACATCGAGCTCCAGAAACTGAAGAAGAAGTGTGCGCGCTCAATGCTGGACT	1560
	*****	*****
<i>Fie2_B73</i>	GATGGTTACGCTCGGTTGGGGTTGCGATGGTTGAATCCGGTGGAAAAGTGCCACCTGG	1620
<i>Fie2_AT-1</i>	GATGGTTACGCTCGGTTGGGGTTGCGATGGTTGAATCCGGTGGAAAAGTGCCACCTGG	1620
<i>Fie2_AT-3</i>	GATGGTTACGCTCGGTTGGGGTTGCGATGGTTGAATCCGGTGGAAAAGTGCCACCTGG	1620
<i>Fie2_AT-4</i>	GATGGTTACGCTCGGTTGGGGTTGCGATGGTTGAATCCGGTGGAAAAGTGCCACCTGG	1620
	*****	*****

**Рис. 4.** Множественное выравнивание фрагмента нуклеотидных последовательностей гена *Zm\_Fie2* партеногенетических линий кукурузы AT саратовской селекции и референсной линии B73. Выделенные буквы обозначают нуклеотидные замены.

### *Zm\_Fie1*

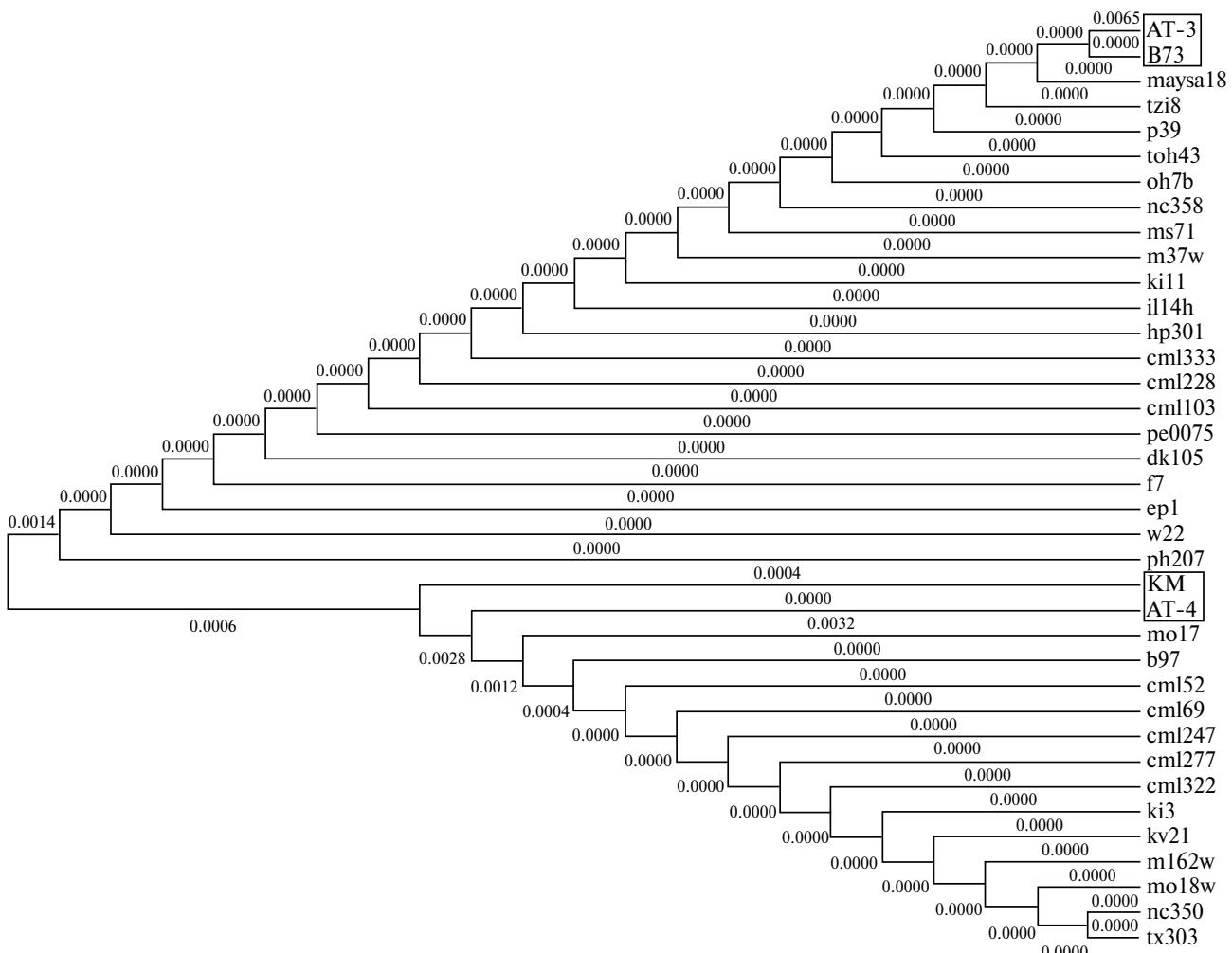
Ген *Zm\_Fie1* кукурузы экспрессируется в начале развития эндосперма и играет важную роль в его развитии [16]. Транскрипт гена *Zm\_Fie1* кукурузы линии AT-3 секвенирован и представлен нами в базе данных GenBank (MW222954.1), а для линий AT-1 и AT-4 транскрипты представлены на рис. 3.

Ген *Zm\_Fie1* у линии AT-1 имеет десять ОНЗ в белок-кодирующей области; у линии AT-3 – три ОНЗ в положениях, отличающихся от AT-1 и AT-4, и делецию из шести нуклеотидов в положении 1362–1367 (рис. 3). Линия AT-4 имеет 11 ОНЗ в белок-кодирующей области гена *Zm\_Fie1*. Количество ОНЗ на 1000 нуклеотидов у гена *Zm\_Fie1* исследованных линий колеблется от 1.9 до 7.5. Линии AT-1 и AT-4 имеют десять ОНЗ в одинаковых положениях, а линия AT-4 имеет дополнительную ОНЗ в положении 512 (рис. 3). Возможно, линия AT-4 унаследовала этот ген от линии AT-1 и приобрела дополнительную ОНЗ.

### *Zm\_Fie2*

Белок FIE2 контролирует экспрессию генов, связанных с началом развития эндосперма [15]. Транскрипты гена *Zm\_Fie2* кукурузы линий AT-1, AT-3 и AT-4 секвенированы и представлены нами на рис. 4. У гена *Zm\_Fie2* линий AT-1 и AT-4 в белок-кодирующей части обнаружено по одной ОНЗ, а у линии AT-3 наблюдаются две ОНЗ в белок-кодирующей части (рис. 4).

У белка FIE2 аминокислотные последовательности для всех трех линий полностью идентичны аминокислотной последовательности белка референсной линии B73 (данные не показаны). Количество ОНЗ на 1000 нуклеотидов у гена *Zm\_Fie2* исследованных линий колеблется от 0.6 до 1.2. Поскольку белок FIE2 контролирует экспрессию генов, связанных с началом развития эндосперма, в том числе, предположительно, гена *Zm\_Fie1*, вполне объяснима его консервативность и схожесть у непартеногенетических и партеногенетических линий кукурузы. Правда, консервативность белка FIE2 не объясняет экспрессии гена *Zm\_Fie1*,



**Рис. 5.** Филограмма линий кукурузы, построенная по гену *Zm\_Chrl06* с помощью программы MEGA 11. Квадратами отмечены исследованные в данной статье линии.

которую мы наблюдали у партеногенетической линии AT-4 без опыления, что не характерно для не-партеногенетических линий кукурузы [21].

Таким образом, можно заключить, что ген *Zm-Fie2* является высококонсервативным, и кодируемые им белки не различаются у партеногенетических и непартеногенетических линий, а спонтанное (без опыления) начало развития эндосперма контролируется другими генами.

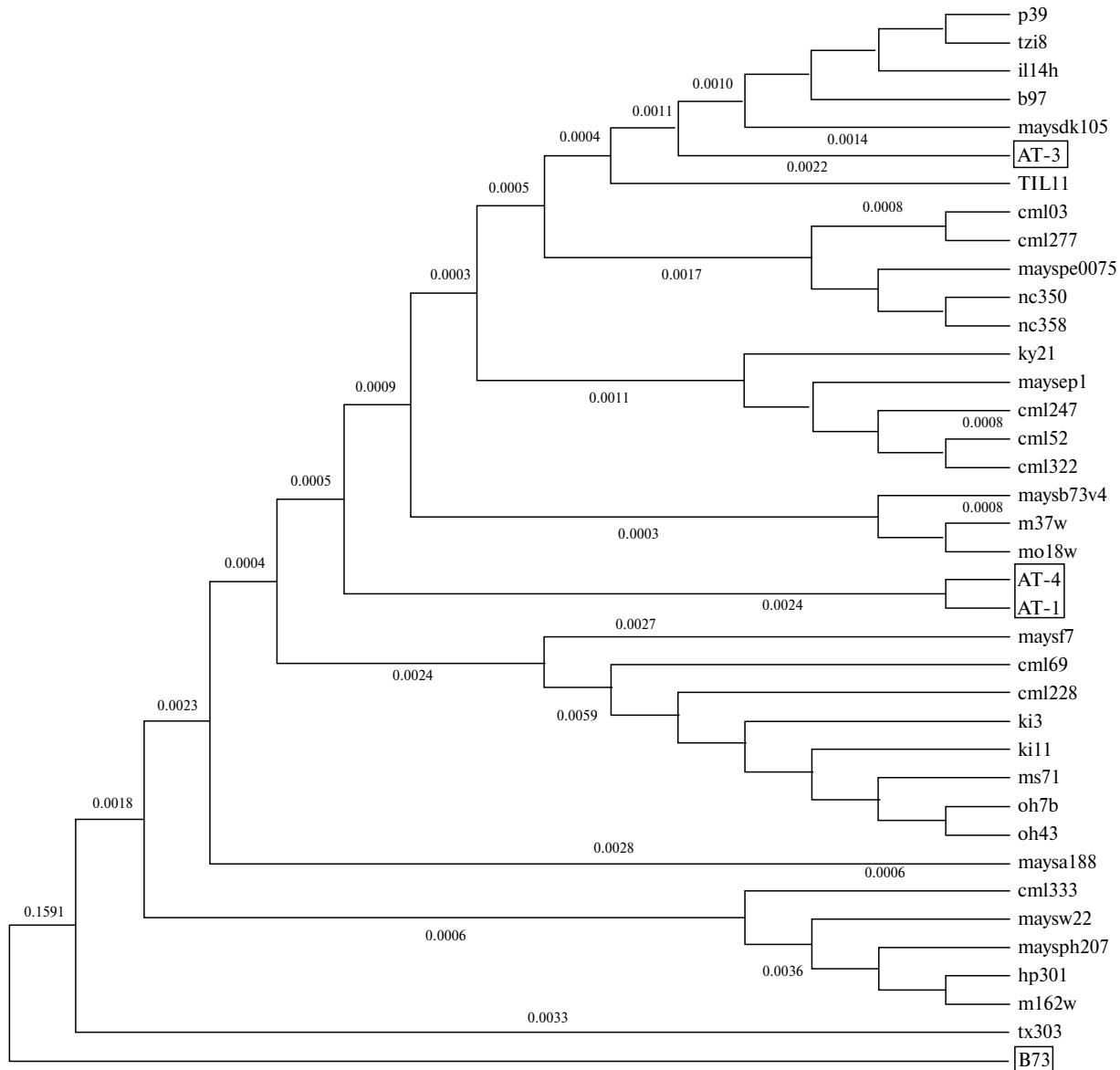
Для выявления филогенетического родства изучаемых линий кукурузы были построены филограммы с использованием последовательностей транскриптов генов *Zm\_Chrl06* (рис. 5) и *Zm\_Fie1* (рис. 6).

На филограмме, построенной на основе сходства последовательностей гена *Zm\_Chrl06*, линии кукурузы разделились на две группы с *p*-расстоянием (эволюционное расхождение) 0.0020. Партеногенетические линии АТ-3 и АТ-4 находятся в

разных группах. Линия АТ-3 и референсная линия В73 имеют  $p$ -расстояние 0.0065.

Линии AT-1 и AT-4 имеют много совпадений по ОНЗ у гена *Zm\_Fie1* (рис. 3) и входят в одну группу (рис. 6). Линия AT-3 от них сильно отличается и находится в отдельной группе, хотя обе линии (AT-3 и AT-4) являются прямыми потомками линии AT-1. То есть между AT-3 и AT-4 наблюдается разница как по гену *Zm\_Chr106* (рис. 5), так и по гену *Zm\_Fie1* (рис. 6). Линия AT-3 и референсная линия B73 имеют *p*-расстояние 0.1849. Возможно, делеция в шесть нуклеотидов в положении 1362–1367 (рис. 3) послужила причиной таких различий (рис. 6).

Таким образом, из четырех исследованных генов (*Hdt104*, *Chr106*, *Fie1* и *Fie2*) у партеногенетических линий кукурузы АТ-1, АТ-3, АТ-4 наиболее консервативными являются гены *Hdt104* и *Fie2*, нуклеотидные замены в которых маловероятно приводят к изменению функции соответствующих белков.



**Рис. 6.** Филограмма линий кукурузы, построенная по гену *Zm\_Fie1* с помощью программы MEGA 11. Квадратами отмечены исследованные в этой статье линии.

Авторы признательны Ю. В. Смолькиной за предоставленные образцы кукурузы.

Исследование выполнено за счет гранта Российского научного фонда (проект № 23-26-00101).

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

- Белова И., Тараканова Т., Абдырахманова Э. и др. Хромосомный контроль апомикса у гибридов кукурузы с гамагравсом // Генетика. 2010. Т. 46. № 9. С. 1188–1191 (Belova I.V., Tarakanova T.K., Abdyrahmanova E.A. et al. Chromosome control of apomixis in maize-gamagrass hybrids / Russ. J. Genet. 2010. V. 46. P. 1055–1057. <https://doi.org/10.1134/S1022795410090103>)
- Grimanelli D. Epigenetic regulation of reproductive development and the emergence of apomixes in angiosperms // Current Opinion in Plant Biology. 2012. V. 15. P. 57–62. <https://doi.org/10.1016/j.pbi.2011.10.002>

3. Koltunow A.M., Grossniklaus U. Apomixis: A developmental perspective // Annual Review of Plant Biology. 2003. Т. 54. № 1. С. 547–574.  
<https://doi.org/10.1146/annurev.applant.54.110901.160842>
4. Bicknell R., Koltunow A. Understanding apomixis: Recent advances and remaining conundrums // The Plant Cell. 2004. V. 16. P. 228–245.  
<https://doi.org/10.1105/tpc.017921>
5. Bradley J., Carman J., Jamison M., Naumova T. Heterochronic features of the female germline among several sexual diploid *Tripsacum* L. (Andropogoneae, Poaceae) // Sex. Plant Reprod. 2007. V. 20. P. 9–17.  
<https://doi.org/10.1007/s00497-006-0038-0>
6. Sauter M., Wiegen P., Lörz H., Kranz E. Cell cycle regulatory genes from maize are differentially controlled during fertilization and first embryonic cell division // Sex Plant Reprod. 1998. V. 11. P. 41–48.  
<https://doi.org/10.1007/s004970050119>
7. Liu X., Fu J., Gu D. et al. Genome-wide analysis of gene expression profiles during the kernel development of *Zea mays* // Genomics. 2008. V. 91. P. 378–387.  
<https://doi.org/10.1016/j.ygeno.2007.12.002>
8. Garcia-Aguilar M., Michaud C., Leblanc O., Grimanelli D. Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes // The Plant Cell. 2010. V. 22. P. 3249–3267.  
<https://doi.org/10.1105/tpc.109.072181>
9. Leblanc O., Grimanelli D., Hernandez-Rodriguez M. et al. Seed development and inheritance studies in apomictic maize – *Tripsacum* hybrids reveal barriers for the transfer of apomixis into sexual crops // Int. J. Dev. Biol. 2009. V. 53. P. 585–596.  
<https://doi.org/10.1387/ijdb.082813ol>
10. Matsuoka Y. Original matters: Lessons from the search for the wild ancestors of maize // Breeding Sci. 2005. V. 33. P. 383–390.  
<https://doi.org/10.1270/jsbbs.55.383>
11. Тырнов В.С., Еналеева Н.Х. Автономное развитие зародыша и эндосперма у кукурузы // Докл. АН СССР. 1983. Т. 272. № 3. С. 722–725.
12. Еналеева Н.Х., Тырнов В.С., Селиванова Л.П., Завалишина А.Н. Одинарное оплодотворение и проблема гаплоиндукиции у кукурузы // Докл. АН СССР. 1997. Т. 353. С. 405–407.
13. Kolesova A.Y., Tyrnov V.S. Embryological peculiarities of tetraploid parthenogenetic maize forms // Maize Genet. Cooperation Newsletter. 2012. V. 85. P. 65–66.
14. Гуторова О.В., Анастасова Н.В., Юдакова О.И. Создание генетически маркированных линий кукурузы с наследуемым и индуцированным типами партеногенеза // Изв. Самарского науч. центра Российской акад. наук. 2016. Т. 18. № 2. С. 341–344.
15. Danilevskaya O.N., Hermon P., Hantke S. et al. Duplicated fie genes in maize: Expression pattern and imprinting suggest distinct functions // Plant Cell. 2003. V. 15. P. 425–438.  
<https://doi.org/10.1105/tpc.006759>
16. Hermon P., Srilunchang K., Zou J. et al. Activation of the imprinted Polycomb group Fie1 gene in maize endosperm requires demethylation of the maternal allele // Plant Mol. Biol. 2007. V. 64. P. 387–395.  
<https://doi.org/10.1007/s11103-007-9160-0>
17. Makarevitch I., Eichten S.R., Briskine R. et al. Genomic distribution of maize facultative heterochromatin marked by trimethylation of H3K27 // Plant Cell. 2013. V. 25. P. 780–793.  
<https://doi.org/10.1105/tpc.112.106427>
18. Li Q., Eichten S.R., Hermanson P.J. et al. Genetic perturbation of the maize methylome // Plant Cell. 2014. V. 26. P. 4602–4616.  
<https://doi.org/10.1105/tpc.114.133140>
19. Чумаков М.И., Мазилов С.И. Генетический контроль гиногенеза у кукурузы (обзор) // Генетика. 2022. Т. 58. № 4. С. 388–397.  
doi: 10.31857/S001667582204004X (Chumakov M. I., Mazilov S. I. Genetic control of maize gynogenesis // Rus. J. Genet. 2022. V. 58. № 4. P. 384–392.  
<https://doi.org/10.1134/S1022795422040044>
20. Volokhina I., Gusev Y., Moiseeva Y. et al. Expression of genes coding for chromatin-modifying enzymes maize embryo sacs before and after pollination // Plant Gene. 2020. V. 22.  
<https://doi.org/10.1016/j.plgene.2020.100221>
21. Volokhina I., Gusev Y., Moiseeva Y. et al. Gene expression in parthenogenic maize proembryos // Plants. 2021. V. 10.  
<https://doi.org/10.3390/plants10050964>
22. Mozgova I., Kohler C., Hennig L. Keeping the gate closed: Functions of the polycomb repressive complex PRC2 in development // The Plant J. 2015. V. 83. P. 121–132.  
<https://doi.org/10.1111/tpj.12828>
23. Grossniklaus U., Vielle-Calzada J.P., Hoeppner M.A., Gagliano W.B. Maternal control of embryogenesis by MEDEA, a polycomb group gene in *Arabidopsis* // Science. 1998. V. 280. P. 446–450.  
<https://doi.org/10.1126/science.280.5362.446>
24. Luo M., Bilodeau P., Koltunow A. et al. Genes controlling fertilization-independent seed development in *Arabidopsis thaliana* // Proc. Nat. Acad. Sci. USA. 1999. V. 96. № 1. P. 296–301.  
<https://doi.org/10.1073/pnas.96.1.296>
25. Ohad N., Yadegari R., Margossian L. et al. Mutations in *FIE*, a WD Polycomb group gene, allow endosperm development without fertilization // Plant Cell. 1999. V. 11. P. 407–416.  
<https://doi.org/10.1105/tpc.11.3.407>
26. Luo M., Platten D., Chaudhury A. et al. Expression, imprinting, and evolution of rice homologs of the polycomb group genes // Mol. Plant. 2009. V. 2. № 4.

- P. 711–723.  
<https://doi.org/10.1093/mp/ssp036>
27. Wu X., Xie L., Sun X. et al. Mutation in Polycomb repressive complex 2 gene *OsFIE2* promotes asexual embryo formation in rice // Nat. Plants. 2023. V. 9. № 11. P. 1848–1861.  
<https://doi.org/10.1038/s41477-023-01536-4>
28. Enaleeva N.Ch., Tyrnov V.S. Cytological manifestation of apomixis in AT-1 plants of corn // Maize Genet. Cooperation Newsletter. 1997. № 71. P. 74–75.
29. Coe E.H. A line of maize with high haploid frequency // Am. Naturalist. 1959. V. 59. P. 381–382.  
<https://doi.org/10.1086/282098>
30. Анастасова Н.В., Гуторова О.В., Юдакова О.И., Смолькина Ю.В. Особенности строения и развития женских генеративных структур у линий кукурузы с наследуемым и индуцированным типами партеногенеза // Изв. Самарского науч. центра Российской акад. наук. 2017. Т. 19. № 2 (2). С. 216–219.
31. Tyrnov V.S. Producing of parthenogenetic forms of maize // Maize Genet. Cooperation Newsletter. 1997. V. 71. P. 73–74.
32. Tyrnov V.S., Smolkina Y.V., Titovets V.V. Estimation of parthenogenesis frequency on the grounds of genetical and embryological data // Maize Genet. Cooperation Newsletter. 2001. V. 75. P. 56–57.
33. Tamura K., Nei M., Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method // Proc. Nat. Acad. Sci. USA. 2004. V. 101 (30). P. 11030–11035.  
<https://doi.org/10.1073/pnas.04042061>
34. Tamura K., Stecher G., Kumar S. MEGA 11: Molecular evolutionary genetics analysis version 11 // Mol. Biol. Evol. 2021. V. 38 (7). P. 3022–3027.  
<https://doi.org/10.1093/molbev/msab120>
35. Смолькина Ю.В. Особенности развития завязей у партеногенетических линий кукурузы без опыления // Бюл. Ботан. сада Саратовского гос. ун-та. 2003. № 2. С. 197–201.
36. Еналеева Н.Х., Тырнов В.С. Цитологическое проявление элементов апомиксиса у линии кукурузы AT-1 и ее гибридов // Апомиксис у растений: состояние, проблемы и перспективы исследования. Саратов, 1994. С. 57–59.

## Partenogenesis Maize genes: Comparative Mutations Analysis

Ye. M. Moiseeva<sup>1</sup>, V. V. Fadeev<sup>1</sup>, Yu. V. Fadeeva<sup>1</sup>, S. I. Mazilov<sup>1</sup>,  
A. Y. Kolesova<sup>2</sup>, M. I. Chumakov<sup>1</sup>, \*

<sup>1</sup>Institute of Biochemistry and Physiology of Plants and Microorganisms, Federal Research Center  
“Saratov Scientific Center of the Russian Academy of Sciences”, Saratov, 410049 Russia

<sup>2</sup>Federal Center of Agriculture Research of the South-East Region, Saratov, 410010 Russia

\*e-mail: chumakov\_m@ibppm.ru

The article presents an analysis of the polymorphism of nucleotide sequences of the genes presumably associated with the parthenogenetic development of the embryo and endosperm in maize. Sequencing and subsequent multiple alignment of transcripts of the target genes (*Hdt104*, *Chr106*, *Fie1* and *Fie2*) studied in the work of the AT-1, AT-3 and AT-4 parthenogenetic maize lines and the reference line B73 determined the presence of SNP, deletions and insertions. The phylogenetic trees for the studied genes were constructed.

**Keywords:** maize, parthenogenesis, mutations, gene polymorphism.

## ГЕНЕТИЧЕСКАЯ ИЗМЕНЧИВОСТЬ И ФИЛОГЕНЕТИЧЕСКИЕ ВЗАИМООТНОШЕНИЯ ПОПУЛЯЦИЙ ОБЫКНОВЕННОЙ ЩУКИ (*Esox lucius* L.) ИЗ НЕКОТОРЫХ РЕК РОССИИ

© 2024 г. Л. Т. Бачевская<sup>1</sup>, В. В. Переверзева<sup>1,\*</sup>, А. А. Примак<sup>1</sup>, Г. А. Агапова<sup>1</sup>

<sup>1</sup>Институт биологических проблем Севера Дальневосточного отделения Российской академии наук,  
Магадан, 685000 Россия

\*e-mail: vvpereverzeva@mail.ru

Поступила в редакцию 10.04.2024 г.

После доработки 17.06.2024 г.

Принята к публикации 02.07.2024 г.

Получены данные о строении и изменчивости нуклеотидной последовательности гена цитохрома *b* mtДНК обыкновенной щуки *Esox lucius* L. из рек Колымского, Волжского, Донского бассейнов и р. Анадырь. Обнаружены уникальные гаплотипы, указывающие на ограниченный поток генов между исследованными популяциями. Щука из рек Волга и Анадырь имеет высокий уровень генетического разнообразия. Низкие значения этого показателя отмечены в некоторых популяциях Колымского и Донского бассейнов. Филогенетический анализ нуклеотидных последовательностей гена цитохрома *b* показал наличие трех генетических линий А, Б, В, имеющих общий корень происхождения. Высокий уровень дифференциации выделенных линий, по-видимому, обусловлен длительной репродуктивной изоляцией обыкновенной щуки в нескольких рефугиумах, которые были источником для циркумполярного распространения этого вида.

**Ключевые слова:** обыкновенная щука, *Esox lucius* L., митохондриальная ДНК (mtДНК), ген цитохрома *b* (*cytb*), генетическая изменчивость, филогенетический анализ.

**DOI:** 10.31857/S0016675824120068 **EDN:** WAKVGV

В истории раннекайнозойской пресноводной ихтиофауны Северного полушария, по мнению Е. К. Сычевской (1975) [1], Esocidae занимают особое место, будучи единственным семейством класса костных рыб, известным с эоценом на всех трех голарктических материках. Открытие и изучение новых форм щуковидных в палеоген – неогене Евразии и Северной Америки [1–4] привели к реконструкции достаточно сложной картины расселения этой группы. Одна из гипотез предлагала считать центром происхождения Esocidae Северную Азию, откуда они проникли в Европу в середине олигоцена. Предположительно на рубеже олигоцена – миоцена могла произойти первая миграция щуковых в Америку [1, 5, 6]. В дальнейшем активные тектонические движения, оледенения, происходившие в Евразии, изменили характер водоемов. Озерно-болотные сменились озерно-речными водоемами [7, 8], а климатические изменения и оледенения, по-видимому, привели к обеднению фауны Esocidae. Возникшая к тому времени в Европе и расселившаяся в Азии в конце позднего плиоцена *E. lucius* L. вытеснила олигоцен-миоценовую фауну азиатских щук. Вероятно, в дальнейшем *E. lucius* L. проникла в водоемы Северной Америки

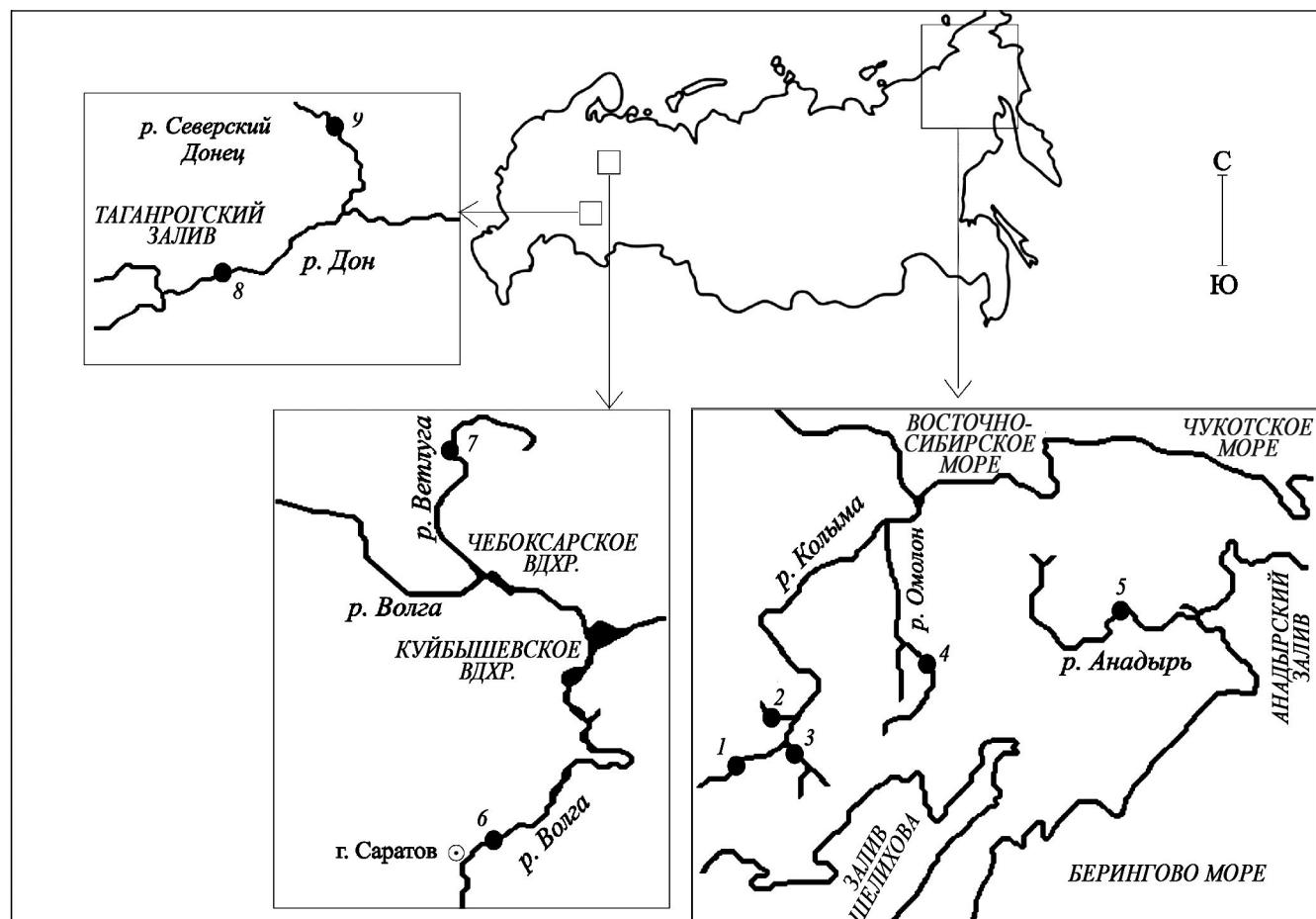
через Берингию [1, 4, 9]. Другая гипотеза о североамериканском происхождении Esocidae [10] получила наиболее широкое распространение после сенсационной находки в Западной Канаде *E. Tietmani*, датированной плиоценом [4], что значительно отодвинуло момент первого появления щук в данном регионе. Несмотря на это, М. В. Wilson [4] не исключал распространение щук из Евразии через Берингийскую сушу, а также допускал их автохтонное возникновение в Северной Америке или возможность проникновения в раннем палеогене из Европы. Обыкновенная щука *E. lucius* L. впервые достоверно отмечена в Европе в конце позднего плиоцена, а в Северной Америке – лишь с плейстоцена [1, 5], имеет циркумполярное распространение. Несомненно, ее расселение связано с похолоданием в бореальной области Северного полушария, что дало большие экологические преимущества этому холодолюбивому виду. Надо отметить, что обыкновенная щука обладает способностью адаптироваться к весьма разнообразным условиям обитания, и это способствовало ее экспансии и широкому географическому распространению. Ее экологическая пластичность – важный биологический фактор, позволивший виду освоить

целый ряд водоемов, расположенных в суровых климатических условиях северо-востока России, включая Чукотский автономный округ. При этом необходимо обратить внимание, что щука населяет водоемы региона мозаично и имеет прерванный ареал. Она отсутствует в водоемах южного побережья Чукотки, а в бассейнах рек Колыма и Анадырь широко распространена [11–14]. Экология вида, обитающего в этих водоемах, хорошо изучена [11, 12, 15–21]. Молекулярно-генетические исследования обыкновенной щуки из рек данного региона до недавнего времени не проводились. Но уже на первых этапах изучения mtДНК этого вида была отмечена дифференциация популяций из Колымского и Анадырского бассейнов и вместе с тем показана общность их происхождения [22]. Молекулярно-генетический анализ популяций из рек Колымского и Анадырского бассейнов выявил высокий уровень межгрупповой изменчивости и позволил сделать вывод о длительном периоде их изоляции при отсутствии возможности обмена генофондом [23]. С привлечением данных GenBank была проведена оценка эволюционного возраста митохондриального генофонда щуки из изученных

водоемов [23]. Приведенные выше сведения показали информативность исследований гена *cytb* обыкновенной щуки и обоснованность продолжения исследования с учетом слабой изученности полиморфизма генов mtДНК вида, населяющего не только северо-восточную, но и европейскую часть России. Таким образом, была расширена география исследований щуки, а целью работы стали определение генетической изменчивости гена *cytb* обыкновенной щуки из рек Волжского, Донского, Колымского и Анадырского бассейнов и анализ родственных внутривидовых взаимоотношений.

## МАТЕРИАЛЫ И МЕТОДЫ

С использованием молекулярно-генетических методов исследованы восемь выборок обыкновенной щуки из некоторых рек России (рис. 1). Биологический материал щук из рек Волжского, Донского и Колымского бассейнов взят из коллекции ИБПС ДВО РАН, находящейся в лаборатории ихтиологии. ДНК выделяли из фрагментов плавников рыб, хранившихся в 96%-ном этиловом спирте. Для работы также использованы уже



**Рис. 1.** Карта-схема района сбора материала. 1 – Колымское водохранилище, 2 – р. Эльген, 3 – р. Буюнда, 4 – р. Омолон, 5 – р. Анадырь, 6 – р. Волга, 7 – р. Ветлуга, 8 – р. Дон, 9 – р. Северский Донец.

опубликованные данные по генетическому разнообразию щуки из рек Колымского бассейна и р. Анадырь [22, 23]. Выделение тотальной ДНК, амплификацию нуклеотидной последовательности гена *cytb*, очистку продукта полимеразной цепной реакции и секвенирование полной нуклеотидной последовательности проводили с применением методик, представленных ранее [22]. Ген *cytb* картировали относительно полной нуклеотидной последовательности mtДНК обыкновенной щуки *Esox lucius* L., GenBank, NC004593 [24]. Обнаруженные нуклеотидные последовательности депонированы в GenBank. Их выравнивание и анализ проводили с использованием пакета программ MEGA 10.0.2.74 [25] и ARLEQUIN ver. 3.5 [26]. Соответствие характера нуклеотидных замен гипотезе нейтральности (для исследованного маркера) в выборках щуки определяли с помощью тестов D F. Tajima и *F<sub>s</sub>* Y.-X. Fu [27–29]. Детекция филогенетических связей гаплотипов гена *cytb* проводилась с использованием пакета программ MEGA 10.0.2.74 [25]. Для построения дендрограммы филогенетических отношений были привлечены нуклеотидные последовательности гена *cytb* обыкновенной щуки из GenBank [30, 31] и использована бипараметрическая модель дистанций Кимуры, учитывающая разную вероятность транзиций и трансверсий [25]. Дендрограмма гаплотипов гена *cytb* построена на основе метода максимального правдоподобия (ML). Оценка узлов ветвления определялась бутстреп-методом (1000 итераций). Для построения медианных сетей применяли алгоритм MJ (Median-Joining) программы Network 4.5.1.0 [32]. В качестве внешней группы использована последовательность нуклеотидов гена *cytb* амурской щуки *Esox reicherti*, GenBank, AY497444 [24].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Анализ ранее полученных [22, 23] и новых нуклеотидных последовательностей гена *cytb* обыкновенной щуки (рис. 2) показал наличие небольшого числа вариабельных позиций от общей его длины. Для популяций из рек Эльген, Буюнда, Омолон, Анадырь и Колыма оно составило 0.44%, равно как и для популяций из рек Волга, Ветлуга, Дон и Северский Донец. С учетом представленных в GenBank образцов нуклеотидных последовательностей гена *cytb* вида из водоемов России, Европы и Северной Америки число вариабельных позиций возросло и стало более значимым (4.04%). Среднее число попарных различий между нуклеотидными последовательностями [25] изученного гена щуки составило 0.4%, хотя интервал различий более широк (0–1.1%). Анализ мутаций, обнаруженных в нуклеотидных последовательностях гена *cytb* щуки из исследованных популяций, позволил отметить три транзиции в первой, две – во второй и шесть – в третьей позиции кодона. Таким образом, число

замен в третьем положении кодона изученного гена в 2 раза превысило таковое в первом положении. Известно, что третий нуклеотид большинства кодонов в транслируемых участках гена наиболее вариабелен из-за вырожденности кода [33]. Анализ, проведенный с учетом представленных в GenBank образцов нуклеотидных последовательностей гена *cytb* вида из водоемов России, не внес изменений в соотношение транзиций по третьей и первой позициям кодона изученного гена. В то же время было отмечено, что число транзиций в исследованных нуклеотидных последовательностях значительно больше, чем трансверсий (23 : 1), что вполне объяснимо [34, 35].

Продолжение исследований обыкновенной щуки из водоемов Колымского бассейна показало, что увеличение количества изученных особей не привело к обнаружению новых гаплотипов, а лишь незначительно изменило частоту распределения обнаруженных ранее (табл. 1). Идентичная картина наблюдалась в процессе изучения щуки из р. Анадырь. Таким образом, в выборках из водоемов Колымского бассейна и р. Анадырь обнаружено шесть гаплотипов. Молекулярно-генетический анализ образцов щуки из рек Волга, Ветлуга, Дон, Северский Донец позволил отметить полиморфизм и выделить также шесть вариантов нуклеотидной последовательности гена *cytb*, представленных на рис. 2. При этом следует подчеркнуть, что гаплотип Ecb1 в большинстве случаев имел высокую частоту распространения (табл. 1). В то же время обращает на себя внимание тот факт, что все гаплотипы, отмеченные у щуки из рек Волжского бассейна, не встречались в исследованных ранее популяциях. Они являются уникальными, наряду с вариантом Ecb6, который обнаружен только у омолонской щуки (табл. 1). Наличие уникальных гаплотипов, по-видимому, указывает на ограниченный поток генов между популяциями. Важно отметить, что щука из р. Ветлуги (приток р. Волги) характеризуется лишь двумя гаплотипами, тогда как у рыб в р. Волге их обнаружено пять и лишь один для них общий (табл. 1). По исследованному маркеру щука из рек Дон и Северский Донец представлена одним гаплотипом Ecb1. Все варианты нуклеотидной последовательности гена *cytb* обыкновенной щуки из рек Волга, Ветлуга, Дон и Северский Донец депонированы в GenBank (OQ547794 – OQ547797). Ряд исследованных нами популяций находятся на большом географическом удалении друг от друга. Они обладают разным набором гаплотипов, при этом для всех характерен невысокий уровень гаплотипического разнообразия (табл. 2). Остается неизвестным, связана ли низкая гаплотипическая изменчивость с малым эффективным размером исследованных популяций или она отражает исторические и демографические особенности их формирования. В то же время, бесспорно, что гаплотип

							1 1111111
		11111122	2233344444	5555677777	778999999990	0000111	
		4622566804	7801246778	1129301245	78311467891	3578134	
		6703325393	9718715170	3488054683	16727870874	0782410	
							ЛИНИЯ А
Ecb1	KT203375	GGCGATCAGT	CTGTCTCGCG	TAAC TG CACA	TGGATTT CG TT	GGGTTAT	
Ecb2	KT203376	.....	.....	.....	.....	.....	C..
Ecb3	KT203377	.....	.....	.....	.....	.....	T..
Ecb4	KT203378	.....	.....	.....	.....	.....	A..
Ecb5	KT203379	.....	.....	.....	.....	.....	G..
Ecb6	KU659805	.....	.....	.....	.....	.....	G..
Ecb7	OQ547794	.....	.....	T.. T..	.....	.....	
Ecb8	OQ547795	.....	C ..	.....	.....	.....	
Ecb9	OQ547796	.....	.....	.....	.....	.....	C..
Ecb10	OQ547797	.....	.....	.....	T..	.....	
KM281456	р. Волга	.....	.....	.....	.....	.....	A..
KM281465	р. Таз	A ..	.....	.....	.....	.....	
KM281466	р. Таз	.....	A ..	.....	.....	.....	G..
AY497446	Канада	.....	.....	.....	.....	.....	A..
AY497445	Канада	.....	.....	.....	G..	.....	
KM281477	Дунай	.....	.....	.....	.....	.....	C..
KM281463	Швеция	.....	G ..	.....	.....	.....	
KM281462	Швеция	.....	.....	A ..	.....	.....	C..
KM281461	р. Урал	.....	T ..	.....	.....	.....	
KM281459	Швеция	.....	.....	.....	A ..	.....	
KM281458	р. Кура	.....	.....	T ..	A ..	.....	
KM281468	оз. Гурон США	.....	.....	C ..	.....	.....	
KM281464	р. Урал	.....	.....	.....	.....	A ..	
KM281460	р. Урал	.....	C ..	.....	.....	.....	A..
KM281457	оз. Гурон США	.....	.....	.....	.....	.....	
AY497453	оз. Онтарио США	A ..	.....	.....	.....	.....	
ЛИНИЯ Б							
AY497451	Германия	.....	T ..	A ..	C.G ..	T ..	AC ..
AY497452	р. Нева	.....	T ..	A ..	.....	T ..	AC ..
KM281475	Аляска	.....	T ..	A ..	.....	T ..	AC ..
KM281474	Азорские о-ва	.....	T ..	T.A ..	A ..	.....	A ..
KM281473	Азорские о-ва	.....	T ..	TC ..	A ..	.....	C ..
KM281472	Швеция	.....	T ..	T ..	A ..	.....	ACC ..
AP004103		.....	T ..	T .. C .. A ..	T ..	T ..	AC ..
ЛИНИЯ В							
KM281478	Дунай	.....	AG.T ..	.....	T ..	.....	C.AC.C ..
KM281476	Дунай	.....	A .. T ..	.....	T ..	.....	C.AC ..
KM281467	Дунай	.....	A .. T ..	.....	T ..	.....	CAAC ..

Рис. 2. Нуклеотидные последовательности гена цитохрома *b* обыкновенной щуки из рек северо-востока России и варианты, представленные в GenBank. Сайты замен показаны от начала гена цитохрома *b* mtДНК.

Ecb1 гена *cytb*, преобладающий в большинстве из них, распространен как в европейских, так и в североамериканских популяциях щуки. Носители его обнаружены по Голарктическому региону, что свидетельствует о трансконтинентальном расселении вида [23, 30, 31].

Рассчитаны индексы молекулярного разнообразия, необходимые для анализа внутривидовой дифференциации и истории становления изученных популяций щуки (табл. 2). Среди популяций, входящих в чукотскую группу, выделяется анадырская щука, которая имеет более высокие индексы молекулярного разнообразия (табл. 2) по

сравнению с популяциями вида из рек Колымского бассейна. Щука из р. Анадырь, по-видимому, благополучно пережила сложные изменения рельефа и гидросети Чукотки. Наиболее драматичной была история развития р. Колымы [13, 36–41]. Возможно, именно это привело к снижению генетического разнообразия щуки Колымского бассейна. Восполнение разнообразия за счет мигрантов из обширного водоема, существовавшего в плиоцен-раннечетвертичную эпоху в пределах Среднесибирского плоскогорья, было нарушено из-за нового водораздела в виде Новосибирских островов [42, 43]. За счет мигрантов из р. Анадырь оно также

**Таблица 1.** Частота распределения гаплотипов гена *cyp19at1* мтДНК в популяциях обыкновенной щуки из водоемов России

Водоем, (число особей)	Ecb1	Ecb2	Ecb3	Ecb4	Ecb5	Ecb6	Ecb7	Ecb8	Ecb9	Ecb10	KM281456
Буюнда (42)	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Эльген (29)	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Омолон (14)	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
Колымское водохранилище (30)	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Анадырь (51)	0.6735	0.1633*	0.0408	0.0408**	0.0816	0.0	0.0	0.0	0.0	0.0	0.0
Волга (31)	0.7097	0.0	0.0	0.0	0.0	0.0	0.1290	0.0968	0.0	0.0323	0.0323
Ветлуга (8)	0.7500	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2500	0.0	0.0
Дон (11)	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Северский Донец (4)	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Примечание. С учетом гаплотипов из базы GenBank: \* – KM281469, \*\* – KM281470.

**Таблица 2.** Индексы молекулярного разнообразия гена *cytb* mtДНК в исследованных популяциях обыкновенной щуки

Водоем	<i>N</i>	<i>V</i>	$Pi \pm SD$	$\pi \pm SD$	$h \pm SD$	$D (P)$	$F_S (P)$
Буюнда	42	0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Эльген	29	0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Омолон	14	0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Колымское водохранилище	30	0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Колымский бассейн в целом	115	2	0.21571 ± 0.26125	0.0002 ± 0.0003	0.2157 ± 0.0462	0.1641 (0.7430)	0.71401 (0.4340)
Анадырь <sup>1</sup>	51	4	0.6337 ± 0.5039	0.00056 ± 0.0005	0.5498 ± 0.0719	-0.64878 (0.2900)	-1.33262 (0.2220)
Волга <sup>2</sup>	31	10	0.7742 ± 0.5810	0.00068 ± 0.00057	0.4839 ± 0.1012	-1.03191 (0.1770)	-1.26884 (0.1400)
Ветлуга	8	1	0.4286 ± 0.4286	0.00038 ± 0.00043	0.4286 ± 0.1687	-0.3335 (0.7900)	0.53626 (0.3980)
Дон	11	0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Северский Донец	4	0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0

Примечание. *N* – количество особей в выборке; *V* – число вариабельных сайтов в нуклеотидной последовательности гена *cytb*;  $Pi$  – среднее число поларных различий между гаплотипами;  $\pi$  – нуклеотидное разнообразие; *h* – гаплотипическое разнообразие; SD – стандартное отклонение; показатели рассчитаны с учетом гаплотипов из базы GenBank.<sup>1</sup> – KM281469, KM281470, <sup>2</sup> – KM281456; *D* – коэффициент теста Tajima;  $F_S$  – коэффициент теста  $F_u$ .

не могло быть восстановлено, так как связь между этими водоемами была нарушена и более не возобновлялась.

Всю гидографическую сеть европейской части бывшего СССР, по-видимому, следует рассматривать (с точки зрения возможности расселения рыб) как единую водную систему, так как только такой подход объясняет сходство видового состава в реках [44]. Формирование долин крупных рек европейской части России, к которым относятся Волга и Дон, происходило в разные геологические периоды. Река Дон имеет сложную картину развития, начиная с границ олигоцена [45]. Не менее сложную историю имеет р. Волга, которую принято считать более молодой по происхождению. Развитие и формирование этих водоемов тесно связано с эпохами неоднократных материковых оледенений и с динамикой (в период трансгрессий и регрессий) существовавших на юге Русской равнины морей и крупных пресных водоемов [45–48]. В процессе сложных геологических перестроек неизбежно снижение внутривидового и внутрипопуляционного генетического разнообразия. Анализ молекуллярно-генетических данных по щуке из рек Волга и Ветлуга показал более высокие значения генетического разнообразия (табл. 2), чем отмеченные для других исследованных популяций вида. Возможно, они избежали разрушительного геологического воздействия на их генетическое разнообразие, и/или для его восстановления было достаточно времени и источников для сохранения и реколонизации. Длительное обитание в динамичных условиях рефugiумов приводит к повышению генетического и фенотипического разнообразия их обитателей [31, 49, 50]. Одним из наиболее крупных рефugiумов являлся Понто-Каспий, в который впадали несколько рек, среди которых Волга и Урал. Следует отметить, что наряду с волжской популяцией уральская щука также имеет высокие показатели нуклеотидного разнообразия гена *cytb* [31, 51, 52]. Возможно, высокий уровень генетического разнообразия щуки из этих водоемов был предопределен тесной связью с указанным рефugiумом. Для анализа полученных данных на селективную нейтральность [27–29] использовали не только *D*-параметр Таджими, но и *Fs*-параметр Фу-теста, так как статистическая сила первого может ограничиваться за счет влияния демографических событий и отбора [28, 53]. Для волжской, ветлужской и анадырской популяций были отмечены отрицательные значения, которые статистически незначимы (табл. 2). По-видимому, наблюдаемый состав генофонда этих популяций щуки сформировался в результате синergии ряда процессов, таких как эффект основателя и отбор.

На протяжении четвертичного периода реки Русской равнины меняли направление течения, соединяясь то с одними, то с другими водоемами.

В дорисское время бассейн верхней Волги входил в состав рек Воронеж – Дон. В дальнейшем они потеряли естественную связь [44, 45, 48]. Учитывая это, можно было бы ожидать схожесть генетического разнообразия щуки из рассматриваемых рек. Однако полученные данные не подтверждают этот тезис. Популяции вида из Волжского и Донского бассейнов существенно отличаются показателями генетического разнообразия, а также индексами на селективную нейтральность (табл. 2). Отмеченный низкий уровень генетического разнообразия в популяциях Донского бассейна, вероятно, в значительной мере связан с его геологическими перестройками, включая последний ледниковый максимум (20–23 тыс. л. н.) и голоценовый период [54, 55]. Последствия оледенения в сочетании с эффектом основателя и стохастическими процессами (дрейф генов), по-видимому, привели к существенному снижению генетического разнообразия этих популяций. Для его восстановления в условиях ограниченного потока генов необходимо длительное время. Отмеченный факт, по-видимому, усугубляется еще и демографическими особенностями щуки, которая имеет низкий эффективный размер популяции [56].

Результаты филогенетического анализа показали наличие трех отчетливо выраженных линий – А, Б и В (рис. 3) в медианной сети (MJ-алгоритм). Следует подчеркнуть, что в линии Б обнаружены транзиции C165T, C279T, G471A, а для линии В характерны транзиции G123A, C327T, T771C (рис. 2). На основании отмеченных особенностей весьма осторожно можно предположить, что эти мутации могут быть генетическими маркерами линий Б и В. Однако только дальнейшее накопление материалов позволит утвердить или отвергнуть данное предположение. Существование филогенетических линий подтверждено анализом *p*-дистанций между нуклеотидными последовательностями гена *cytb*, который показал высокий уровень статистической значимости различий между линиями А и Б,  $P < 0.001$ , между линиями Б и В ( $P < 0.05$ ). На основании *p*-дистанций построена дендрограмма сходства (рис. 4) и выделены кластеры, имеющие высокие значения бутстреп-индексов (74, 84, 87% соответственно), свидетельствующие, по-видимому, о длительной репродуктивной изоляции щуки из этих линий. Рассматриваемые генетические линии имеют общий корень происхождения, при этом А, Б появились вследствие бифуркации линии В (рис. 3). Последняя характеризуется относительно небольшим разнообразием гаплотипов, имеющих ограниченную географию распространения (р. Дунай). Проводимые ранее исследования показали, что регион Дуная был важным рефugiумом во время плейстоценовых оледенений [10]. Учитывая, что линии А и Б произошли от общего предка из генетической линии В, можно

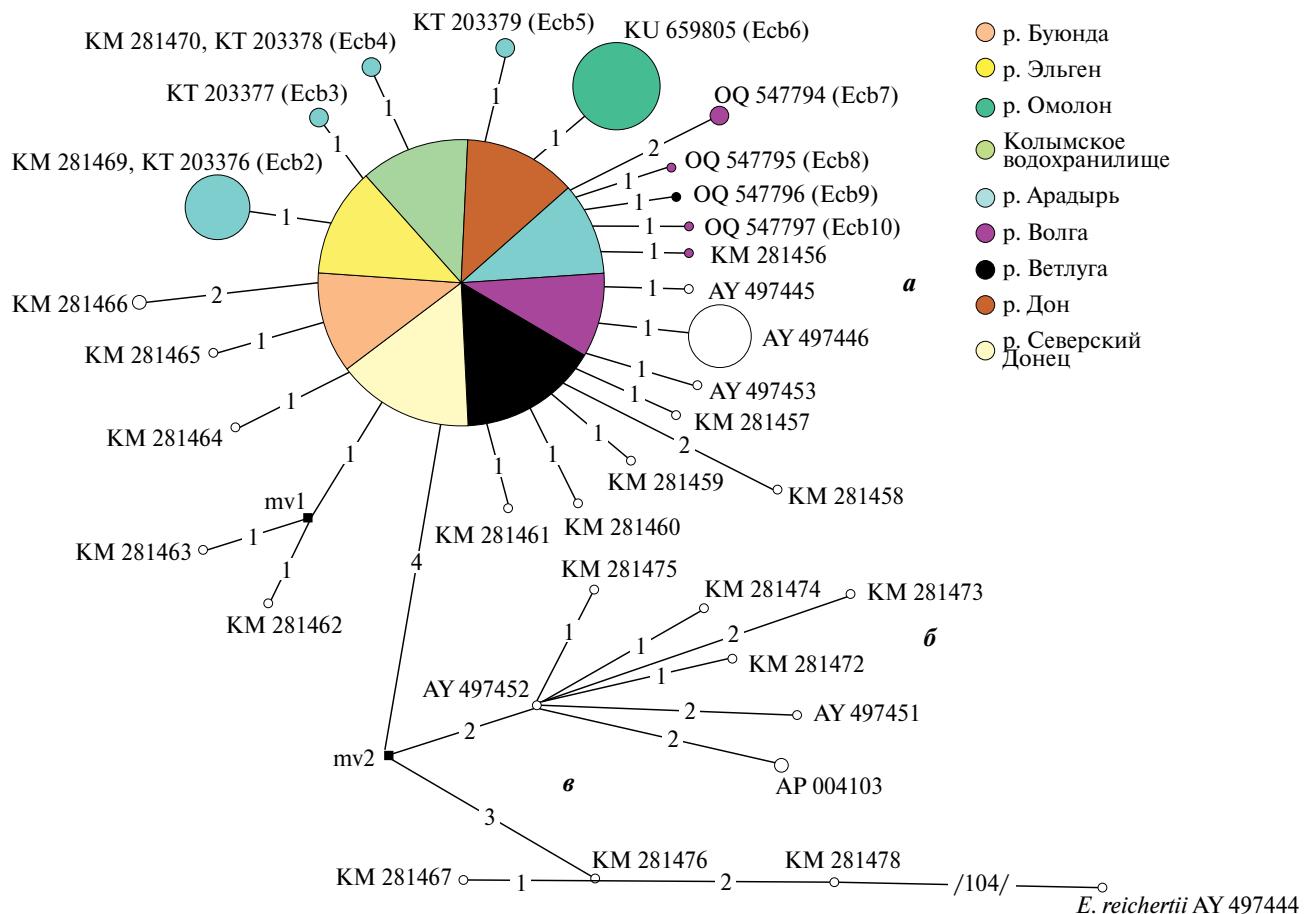
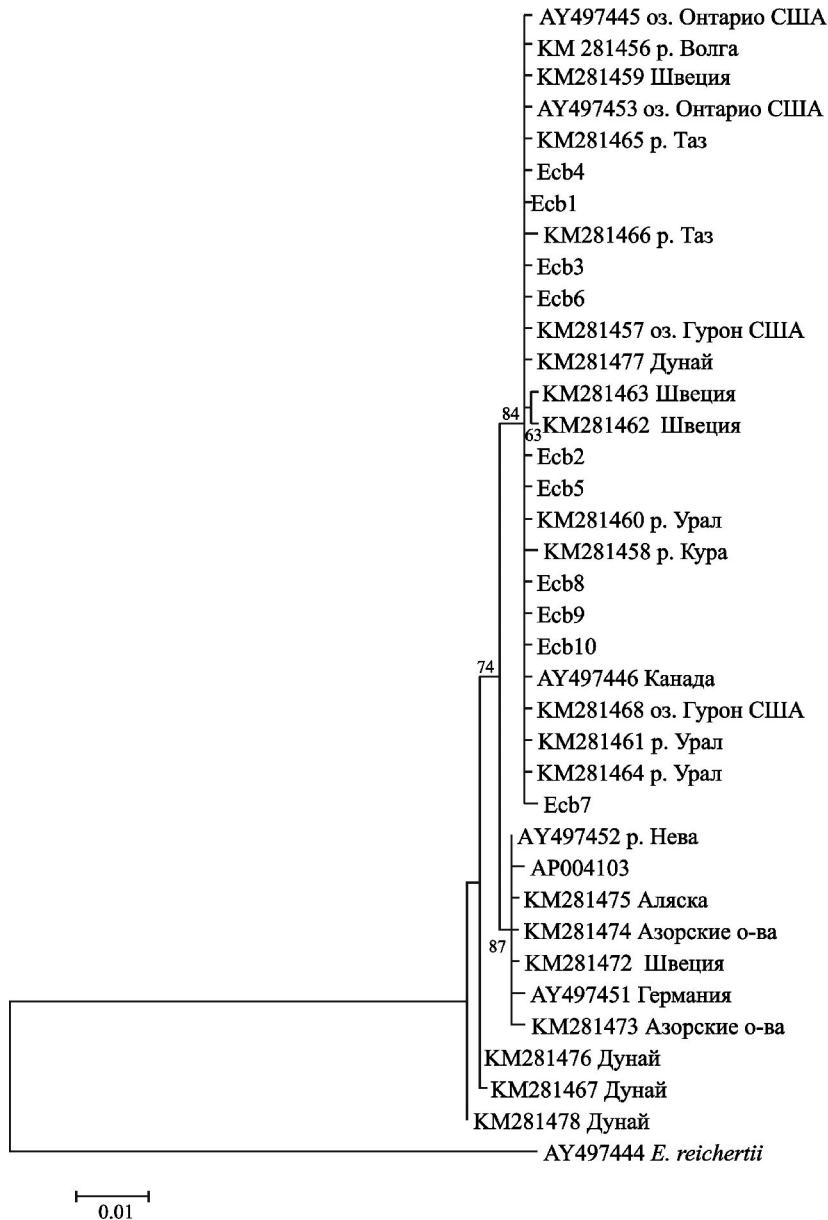


Рис. 3. Медианная сеть гаплотипов гена цитохрома *b* обыкновенной щуки из некоторых рек России и варианты, представленные в GenBank. На ветвях сети цифрами обозначено число мутаций; mv – медианный вектор. Размеры кружков соответствуют количеству гаплотипов. А, Б, В – филогенетические линии.

предположить, что расселение вида могло происходить из этого крупного водного бассейна. Расширению ареала вида способствовали также ледниковые озера, существовавшие в этом регионе в разные геологические периоды [57–60]. Филогенетическая линия Б обыкновенной щуки представлена гаплотипами, которые отмечены у щуки из водоемов Северной Америки, Центральной, Западной Европы и Азорских островов. Они произошли от варианта GenBank, AY497447, обнаруженного у щуки из р. Невы [31]. Несмотря на сложные условия, неоднократно повторявшиеся в истории региона, носители этого гаплотипа смогли выжить и значительно расширить свой ареал. Возможно, их выживанию и распространению способствовали крупные прогляциальные озера. Существование последних подтверждают результаты многочисленных геологических исследований [57–64]. Одни озера подпруживались краем Скандинавского щита, другие – частично Скандинавским и Баренцево-Карским покровом [62]. Среди таких бассейнов было

Балтийское ледниковое озеро, которое играло роль убежища для многих видов рыб. В период отступления ледника оно значительно расширялось в юго-восточном направлении и объединялось с другими многочисленными водоемами [64–66]. Сток из Балтийского ледникового озера в разные исторические периоды осуществлялся через разные проливы. Один из них проходил по территории современной Среднешведской низменности, что способствовало расселению ихтиофауны не только в северную, но и, возможно, в центральную часть Европы, в реках которой обнаружены носители гаплотипов из линии Б. Существование системы прогляциальных озер, занимавших значительные территории [62], видимо, способствовало распространению щуки не только в западном, но и в восточном направлении. Возможно дальнейшее исследование щуки из рек Сибири и Якутии (в настоящий момент мало изученной в молекулярно-генетическом направлении) позволит обнаружить носителей гаплотипов этой линии,



**Рис. 4.** ML-филогенетическое дерево нуклеотидных последовательностей гена цитохрома *b* обыкновенной щуки из некоторых рек России и варианты, представленные в GenBank. В узлах ветвления указаны бутстреп-индексы (> 50%).

прапородители которых, по-видимому, смогли в дальнейшем расселиться в них, а затем проникнуть в Северную Америку через Сибирско-Аляскинский речной комплекс [36]. Генетическая линия А (рис. 3) представлена значительным числом близкородственных гаплотипов с циркумполлярным распространением их носителей. Эта линия включает все обнаруженные нами варианты нуклеотидных последовательностей гена *cytb* и целый ряд гаплотипов, характеризующих обыкновенную щуку из рек Европы и Северной Америки. Они являются производными (путем одной или двух мутаций) от

Ecb1 (GenBank, KT203375, KM281455, KM281456, AY497447, AY497449, AY497450). Возможно, носители этого гаплотипа были наиболее многочисленными и экологически более пластичными, что обеспечило им широкое географическое распространение, в том числе и в реки Северной Америки. Генетическая линия А (как и линия Б) имеет выраженную “звездчатую” картину радиации, которая свидетельствует в пользу предположения о быстром и относительно недавнем расселении из ледниковых рефугиумов, которое происходило, по-видимому, волнобразно. Возможно, первая

волна переселения в североамериканские водоемы принадлежала представителям обыкновенной щуки из более архаичной (рис. 3) филогенетической линии Б. Следующая включала носителей гаплотипов из более “молодой” филогенетической линии А. Такое предположение объясняет наличие в североамериканских популяциях обыкновенной щуки гаплотипов, относящихся к обеим филогенетическим линиям, и не противоречит ранее высказанному мнению по поводу ее расселения [1, 4, 9, 10, 31]. Предложенные авторами сценарии расширения ареала обыкновенной щуки носят предположительный характер и могут изменяться по мере проведения дальнейших исследований и накопления данных.

Работа выполнена в рамках Государственного задания ФГБУН Институт биологических проблем Севера ДВО РАН (№ 1021060707934-2).

Настоящая статья не содержит каких-либо исследований с участием животных и человека. В качестве биоматериала в работе использовались застекленные фрагменты плавников рыб из ихтиологической коллекции ИБПС ДВО РАН.

Авторы заявляют об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Сычевская Е.К. Ископаемые щуковидные СССР и Монголии. М.: Наука, 1975. 115 с.
2. Сычевская Е.К. Роль Берингийской суши в расселении щуковидных // Берингия в кайнозое. Владивосток: Изд-во ДВНЦ АН СССР, 1976. С. 242–258.
3. Gaudnt J. Decouverte du plus ancien representant connu du genre *Esox lucius* L. (Poisson Teleosteen, Esocoidei) dans le Stavpien moyen du bassien d' Art (Vaucluse) // Geol. editer. 1978. V. 5. № 2. P. 257–268.
4. Wilson M.V. Oldest known *Esox lucius* L. (Pisces: Esocidae), part of a new Paleocene teleost fauna from Western Canada // Can. J. Earth Sci. 1980. V. 17. P. 307–312.
5. Сычевская Е.К. Пресноводная палеогеновая ихтиофауна СССР и Монголии. М.: Наука, 1986. 177 с.
6. Хистарова Г.Д. Рыбы из континентальных отложений позднего мела и кайнозоя Центрального и Восточного Казахстана // Ископаемая фауна и флора Центрального и Восточного Казахстана: Материалы по истории фауны и флоры Казахстана. Алма-Ата: Наука, 1971. Т. 5. С. 57–62.
7. Яковлев В.Н. Распространение пресноводных рыб неогена Голарктики и зоogeографическое районирование // Вопр. ихтиологии. 1961. Т. 1. № 2(19). Р. 209–220.
8. Яковлев В.Н. История формирования фаунистических комплексов пресноводных рыб // Вопр. ихтиологии. 1964. Т. 4. № 1(30). С. 10–22.
9. Пугачев О.Н. Паразиты пресноводных рыб северо-востока Азии. Л.: Зоол. ин-т АН СССР, 1984. 156 с.
10. Banarescu P. Principles and Problems of Zoogeography. Yugoslavia, Subotica, 1975. 214 p. Translated from: Principii si Probleme de Zoogeografie. Red. Acad. Rep. Soc. Romania, 1970.
11. Новиков А.С. Рыбы реки Колымы. М.: Наука, 1966. 135 с.
12. Кириллов Ф.Я. Рыбы Якутии. М.: Наука, 1972. 358 с.
13. Черешнев И.А. Происхождение пресноводной ихтиофауны районов Берингии // Биогеография Берингийского сектора Субарктики. Владивосток: Изд-во ДВНЦ АН СССР, 1986. С. 122–145.
14. Черешнев И.А. Биогеография пресноводных рыб Дальнего Востока России. Владивосток: Дальнаука, 1998. 130 с.
15. Берг Л.С. Рыбы пресных вод СССР и сопредельных стран. М.; Л.: Изд-во АН СССР, 1949. Ч. 2. С. 469–925.
16. Черешнев И.А., Шестаков А.В., Скопец М.Б. Определитель пресноводных рыб северо-востока России. Владивосток: Дальнаука, 2001. 129 с.
17. Черешнев И.А. Пресноводные рыбы Чукотки. Магадан: СВНЦ ДВО РАН, 2008. 324 с.
18. Грунин С.И., Бакланов М.А. Размерная изменчивость внешнеморфологических признаков сеголетков щуки *Esox lucius* L. // Вестн. Перм. гос. ун-та. 2004. Вып. 2: Биология. С. 99–101.
19. Грунин С.И. Динамика популяционных показателей щуки *Esox lucius* L. (Esocidae) р. Анадырь за период 1971–2010 гг. // Изв. ТИНРО. 2014. Т. 179. С. 45–54.
20. Хохлов Ю.Н., Хохлова Ю.Е. Особенности питания обыкновенной щуки *Esox lucius* L. в среднем течении р. Анадырь в осенний период // Вестн. СВНЦ ДВО РАН. 2006. № 2. С. 54–58.
21. Грунин С.И., Шестаков А.В. Современное состояние популяции щуки *Esox lucius* L. среднего течения р. Анадырь (Чукотка) // Вопр. рыболовства. 2011. Т. 12. № 3. С. 457–471.
22. Бачевская Л.Т., Переверзева В.В., Агапова Г.А. и др. Генетическое разнообразие популяций щуки обыкновенной *Esox lucius* L. из рек северо-востока России // Изв. РАН. Сер. биол. 2019. № 2. С. 154–161.  
doi 10.1134/S0002332919020036
23. Бачевская Л.Т., Переверзева В.В., Агапова Г.А. и др. Генетическая структура обыкновенной щуки (*Esox lucius* L.) из рек северо-востока России // Вестн. СВНЦ ДВО РАН. 2020. № 2. С. 98–106.  
doi 10.34078/1814-0998-2020-2-98-106

24. Ishiguro N.B., Miya M., Nishida M. Basal euteleostean relationships: A mitogenomic perspective on the phylogenetic reality of the 'Protacanthopterygii' // Mol. Phylogenet. Evol. 2003. V. 27. P. 476–488.
25. Tamura K., Stecher G., Peterson D. et al. MEGA-6: Molecular Evolutionary Genetics Analysis Version 6.0.2.74 // Mol. Biol. Evol. 2013. V. 30. P. 2725–2729. doi 10.1093/molbev/mst197
26. Excoffier L., Laval G., Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis // Evol. Bioinform. Online. 2005. V. 1. P. 47–50.
27. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism // Gen. Soc. Am. 1989. P. 585–595.
28. Fu Y., Li W. Statistic test of neutrality of mutations // Genetics. 1993. V. 133. P. 693–709.
29. Fu Y.-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection // Genetics. 1997. V. 147. P. 915–925.
30. Grande T., Laten H., Lopez J.A. Phylogenetic relationships of extant esocid species (Teleostei: Salmoniformes) based on morphological and molecular characters // Copeia. 2004. V. 4. P. 743–757.
31. Skog A., Vollestad L.A., Stenseth N.C. et al. Circumpolar phylogeography of the northern pike (*Esox lucius*) and its relationship to the Amur pike (*E. reicherti*) // Frontiers Zool. 2014. V. 11. P. 67. doi <http://www.frontiersinzoology.com/content/11/1/67>.
32. Bandelt H.J., Forster P., Röhl A. Median-Joining networks for inferring intraspecific phylogenies // Mol. Biol. Evol. 1999. V. 16. P. 37–48.
33. Zardoya R., Meyer A. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates // Mol. Biol. Evol. 1996. V. 13. № 7. P. 933–942.
34. Nei M. Molecular Evolutionary Genetics. N.Y.: Columb. univ. press, 1987. 512 p.
35. Nei M., Kumar S. Molecular Evolution and Phylogenetics. N.Y.: Oxford univ. press, 2000. 333 p.
36. Линдберг Г.У. Четвертичный период в свете биогеографических данных. М.; Л.: Изд-во АН СССР, 1955. 334 с.
37. Линдберг Г.У. Крупные колебания уровня океана в четвертичный период. Л.: Наука, 1972. 548 с.
38. Черешнев И.А. Биологическое разнообразие пресноводной ихтиофауны северо-востока России. Владивосток: Дальнаука, 1996. 198 с.
39. Баранова Ю.П., Бискэ С.Ф. Северо-восток СССР. М.: Наука, 1964. 350 с.
40. Байрон И.Ю., Миллер В.Г., Минченок В.Д. Развитие низовьев р. Колымы в плейстоцене // Геоморфология. 1977. № 2. С. 44–51.
41. Васьковский А.П. Краткий очерк растительности, климата и хронологии четвертичного периода в верховьях рек Колымы и Индигирки и на северном побережье Охотского моря // Ледниковый период на территории Европейской части СССР и Сибири. М.: Изд-во МГУ, 1959. С. 87–95.
42. Лунсгерггаузен Г.Ф. Геологическая история средней Лены и некоторые вопросы стратиграфии четвертичных отложений Восточной Сибири // Материалы Всес. совещания по изучению четвертич. периода. М.: Изд-во АН СССР, 1961. Ч. 3. С. 209–217.
43. Жуков В.В., Пинчук Л.Я., Пуминов А.П. Палеогеография северо-востока Средне-Сибирского плоскогорья и Лено-Анадырской низменности в кайнозое // Четвертичный период Сибири. М.: Наука, 1966. С. 264–272.
44. Лебедев В.Д. Пресноводная четвертичная ихтиофауна Европейской части СССР. М.: Изд-во МГУ, 1960. 404 с.
45. Горецкий Г.И. Палеопотаиологические эскизы Палео-Дона и Пра-Дона. Минск: Наука и техника, 1982. 248 с.
46. Родзянко Г.Н. История развития долины Дона в плиоцене. Возраст и генезис переуглублений на шельфах и история речных долин. М.: Наука, 1984. С. 64–76.
47. Горецкий Г.И. Аллювий великих антропогенных прарек Русской равнины" М.: Наука, 1964. 415 с.
48. Горецкий Г.И. Формирование долины р. Волги в раннем и среднем антропогене. Аллювий Пра-Волги. М.: Наука, 1968. 412 с.
49. Махров А.А. Снижение эволюционной пластичности в результате филогенетической иммобилизации и его экологическое значение // Сиб. экол. журн. 2019. Т. 26. № 5. С. 491–505. doi 10.15372/SEJ20190501
50. Шмальгаузен И.И. Проблема устойчивости органических форм (онтогенезов) в их историческом развитии // Журн. общей биологии. 1945. Т. 6. № 1. С. 3–25.
51. Bernatchez L., Wilson C. Comparative phylogeography of Nearctic and Palearctic fishes // Mol Ecol. 1998. V. 7. P. 431–452.
52. Nesbø C.L., Fosseheim T., Vollestad L.A., Jakobsen K.S. Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization // Mol. Ecol. 1999. V. 8. P. 1387–1404. doi: 10.1046/j.1365-294x.1999.00699.x
53. Simonsen K.L., Churchill G.A., Aquadro C.F. Properties of test of neutrality for DNA polymorphism data // Genetics. 1995. V. 141. P. 413–429.
54. Панин А.В., Сидорчук А.Ю., Власов М.В. Мощный поздневалдайский речной сток в бассейне Дона

- // Изв. РАН. Серия географическая. 2013. № 1. С. 118–129.
55. Матлахова Е.Ю., Панин А.В., Беляев В.Р., Борисова О.К. Развитие долины Верхнего Дона в конце позднего плейстоцена // Вестн. МГУ. Серия 5: География. 2019. № 3. С. 83–92.
  56. Miller L.M., Kapuscinski A.R. Historical analysis of genetic variation reveals low effective population size in a northern pike (*Esox lucius*) population // Genetics. 1997. V. 147. P. 1249–1258.
  57. Квасов Д.Д. Позднечетвертичная история крупных озер и внутренних морей Восточной Европы. Л.: Наука, 1975. 278 с.
  58. Лавров А.С. Позднеплейстоценовые подпрудные озера на северо-востоке Русской равнины // История озер в плейстоцене: Тез. докл. IV Всес. симпоз. по истории озер. Т. 2. Л., 1975. С. 119–127.
  59. Гросвальд М.Г. Покровные ледники континентальных шельфов. М.: Наука, 1983. 216 с.
  60. Гросвальд М.Г. Последнее оледенение плейстоцена // Атлас снежно-ледовых ресурсов мира. М. Ин-т географии, 1997. Т. 2. С. 104–139.
  61. Гросвальд М.Г. Евразийские гидросферные катастрофы и оледенение Арктики. М.: Научный мир, 1999. 120 с.
  62. Гросвальд М.Г. Оледенение Русского Севера и северо-востока в эпоху последнего великого походления: Материалы гляциологических исследований. Вып. 106. М.: Наука, 2009. 153 с.
  63. Liedtke H. Die nordischen Vereisungen in Mitteleuropa. 2te Auflage. Trier: Paulinus-Verlag, 1981. 308 p.
  64. Высоцкий Н.К. Очерк третичных и посттретичных образований Западной Сибири // Геол. иссл. и разведочные работы по линии Сибирской железной дороги. Вып. 5. СПб.: типогр. Стасюлевича, 1896. С. 69–94.
  65. Пирожников П.Л. К вопросу о происхождении северных элементов в фауне Каспия // Докл. АН СССР. 1937. Т. 15. № 8. С. 513–516.
  66. Belt T. The steppes of Southern Russia // Quarterly Journ. the Geol. Soc. of London. 1877. V. 33. P. 843–862.

## Genetic Variation and Phylogenetic Relationships of Northern Pike (*Esox lucius* L.) Populations from Some Rivers of Russia

L. T. Bachevskaya<sup>1</sup>, V. V. Pereverzeva<sup>1,\*</sup>, A. A. Primak<sup>1</sup>, G. A. Agapova<sup>1</sup>

<sup>1</sup>Institute of Biological Problems of the North, Far East Branch of the Russian Academy of Sciences,  
Magadan, 685000 Russia  
\*e-mail: vvpereverzeva@mail.ru

Data on the structure and variability of the nucleotide sequence of the mtDNA cytochrome *b* gene of the northern pike *Esox lucius* L. from the rivers of the Kolyma, Volga, Don basins and the Anadyr River were obtained. Unique haplotypes were found, indicating limited gene flow between the studied populations. Northern pikes from the Volga and the Anadyr have a high level of genetic diversity. Low genetic diversity was noted in some populations of the Kolyma and Don basins. Phylogenetic analysis of the nucleotide sequences of the cytochrome *b* gene showed the presence of three genetic lines – A, B, V, having a common root of origin. The high level of differentiation of the identified lines is apparently caused by the long-term reproductive isolation of the northern pike in several refugia, which were the source for the circumpolar distribution of this species.

**Keywords:** northern pike *Esox lucius* L., mitochondrial DNA (mtDNA), cytochrome *b* (*cytb*) gene, genetic variability, phylogenetic analysis.

## ОЦЕНКА ПОЛИМОРФИЗМА ГЕНОВ ИММУННОЙ СИСТЕМЫ *TNF* И *TLR2* У КРУПНОГО РОГАТОГО СКОТА

© 2024 г. Л. А. Калашникова<sup>1</sup>, И. Е. Багаль<sup>1,\*</sup>,  
Н. Е. Муругина<sup>2</sup>, В. Е. Калашников<sup>1</sup>, Р. Ю. Сенина<sup>1</sup>

<sup>1</sup>Всероссийский научно-исследовательский институт племенного дела, Московская область,  
г. Пушкино, 141212 Россия

<sup>2</sup>Государственный научный центр “Институт иммунологии”, Москва, 115522 Россия

\*e-mail: ladnatehplem@mail.ru

Поступила в редакцию 27.04.2024 г.

После доработки 07.06.2024 г.

Принята к публикации 31.07.2024 г.

Представлены результаты генотипирования генов фактора некроза опухоли (*TNF*) и толл-подобного рецептора 2 (*TLR2*) у 644 голов крупного рогатого скота шести пород, относящихся по происхождению к четырем разным генетическим корням. Исследованы холмогорская и голштинская породы черно-пестрого корня, айрширская порода красного корня, костромская и бурая швицкая породы бурого корня и палевая симментальская порода. Для выявления аллелей и генотипов использовали метод ПЦР-ПДРФ. Установлены достоверные различия частоты встречаемости аллелей генов *TNF* и *TLR2* между породами разных генетических корней ( $p < 0.001$ ) и между породами молочного и молочно-мясного направления продуктивности ( $p < 0.001$ ). У пяти пород скота из шести изученных по гену *TLR2* выявлено преобладание аллеля *T*, наиболее высокая его частота (0.99) установлена у айрширской породы. Стадо айрширских коров близко к мономорфному типу по гену *TLR2*, 98% животных имеют гомозиготный генотип *TT*. Более равномерное распределение аллелей и генотипов по гену *TLR2* установлено у животных черно-пестрого корня: у голштинского скота средняя частота аллеля *T* составила 0.78, у холмогорского – 0.83. Генотип *TT* выявлен у большинства голштинских (61%) и холмогорских (69%) коров. Частота аллеля *T* у симментальской породы составила 0.62, у костромской породы – 0.67. У симментальских и костромских коров чаще встречается гетерозиготный генотип *TG* (54 и 47% соответственно). Иная картина распределения частоты аллелей и генотипов выявлена у животных бурой швицкой породы. Частота аллеля *T* составляет в среднем 0.26. Более половины коров бурой швицкой породы (55%) имеют гомозиготный генотип *GG*. У всех пород по гену *TNF* преобладал аллель *B* с частотой встречаемости от 0.61 у голштинского скота до 0.98 у бурой швицкой породы. Высокая частота аллеля *B* установлена у костромской (0.94), айрширской (0.885) и симментальской (0.80) пород; у холмогорской породы выявлена средняя частота данного аллеля – 0.72. Наиболее высокая частота генотипа *BB* установлена в стадах бурой швицкой породы (96%). Выявлено преобладание гомозиготного генотипа *BB* у холмогорского (53%), симментальского (62%), айрширского (77%) и костромского скота (87.5%). У голштинских коров большинство животных (53%) представлены гетерозиготным генотипом *AB*. У трех пород – айрширской, костромской и бурой швицкой – не выявлен генотип *AA*.

**Ключевые слова:** крупный рогатый скот, гены иммунной системы, *TNF*, *TLR2*.

**DOI:** 10.31857/S0016675824120079 **EDN:** WAAWUE

Изменчивость генов иммунной системы играет важную роль в формировании иммунного ответа при развитии заболеваний крупного рогатого скота [1, 2]. Идентификация генетических вариантов и исследование их ассоциаций с устойчивостью к заболеваниям необходимы для селекционно-племенной работы, направленной на улучшение показателей здоровья животных и продуктивного

долголетия. Использование маркерных последовательностей, связанных с иммунитетом, будет способствовать улучшению показателей иммунного статуса животных.

Эффективность животноводства, объем и качество продукции можно значительно повысить за счет снижения потерь от болезней. Несмотря на

современные технологические достижения, одной из основных проблем в молочном скотоводстве является распространенность инфекций вымени, особенно в высокопродуктивных стадах, что ведет к увеличению затрат, связанных с лечением и профилактикой заболеваний, а также увеличивает риск ранней выбраковки коров [3, 4].

Толл-подобные рецепторы (TLR) участвуют в распознавании патоген-ассоциированных молекулярных паттернов и управляют самыми ранними стадиями иммунного ответа [5]. В геноме крупного рогатого скота были идентифицированы десять генов TLR. TLR2 крупного рогатого скота расположен на хромосоме 17 (NCBI ID 28154) [6]. По данным L.P. Zhang с соавт. [7], этот ген охватывает около 13.2 кбп геномной ДНК и состоит из двух экзонов и одного интрона. Первый экзон имеет длину 179 пн, а второй – 3333 пн. TLR2 расположен на поверхности клеток и распознает широкий спектр паттернов, ассоциированных с патогенами (PAMPs) [8–10].

TLR2 принимает активное участие в формировании устойчивости к маститу [2, 7, 11, 12]. Обнаружена выраженная экспрессия TLR2 при мастите, вызванном *S. aureus* [13]. Выявлены одноклеточные мутации гена TLR2, ассоциированные с паратуберкулезной инфекцией крупного рогатого скота [14, 15]. Установлены различия в иммунном ответе у разных пород крупного рогатого скота, которые могут способствовать повышенной или сниженной восприимчивости животных к инфекции [16]. Было показано, что локус SNP rs55617172 гена TLR2 достоверно ( $p < 0.01$ ) связан с устойчивостью к туберкулезу крупного рогатого скота [17].

При взаимодействии с возбудителем TLR2 посредством каскада реакций индуцирует экспрессию провоспалительных цитокинов (TNF и др.) [18]. Фактор некроза опухоли (TNF) был впервые идентифицирован как цитокин с противоопухолевым действием [19, 20]. Ген, кодирующий TNF, содержит четыре экзона, три интрона и расположен на хромосоме BTA23q22 (NCBI ID 280943) [21, 22].

TNF является одним из основных провоспалительных цитокинов. Он индуцирует высвобождение многих других цитокинов [23, 24] и является одним из ключевых медиаторов местного воспалительного иммунного ответа. Исследованиями ряда авторов определены одноклеточные замены (SNP) в гене TNF, ассоциированные с мастилом коров [12, 25–27]. Установлена связь между полиморфизмом гена TNF и лейкозом крупного рогатого скота [28, 29]. Показана взаимосвязь полиморфизмов гена TNF с риском развития туберкулеза крупного рогатого скота [30]. У голштинских коров было показано влияние SNP гена TNF на иммунную и репродуктивную функции [31, 32]. Кроме того, S. Kahl с соавт. установили связь

полиморфизма гена TNF с гиперчувствительностью к эндотоксину у телят [33].

Таким образом, гены паттерн-распознающих рецепторов и провоспалительных цитокинов благодаря своим биологическим функциям могут быть использованы в качестве потенциальных ДНК-маркеров иммунитета крупного рогатого скота.

Целью нашего исследования было изучение полиморфизма генов TNF и TLR2 у крупного рогатого скота разных пород.

## МАТЕРИАЛЫ И МЕТОДЫ

Для изучения полиморфизма генов TNF и TLR2 в лаборатории ДНК-технологий ФГБНУ ВНИИ племенного дела методом ПЦР-ПДРФ были исследованы пробы ДНК крупного рогатого скота шести пород из девяти племенных хозяйств РФ, принадлежащих четырем разным генетическим корням. Всего было исследовано 644 головы. Объектом исследований послужили животные молочного направления продуктивности: коровы холмогорской породы из Архангельской области (120 голов), голштинской породы (146 голов) зарубежной и отечественной селекции из трех регионов РФ – из Красноярского края (27 голов), из Курской области (40 голов), из двух хозяйств Московской области (39 и 40 голов), а также айрширской породы из Вологодской области (100 голов). Были изучены коровы молочно-мясного направления продуктивности: костромской породы (32 головы) из Костромской области, бурой швицкой породы (141 голова) из двух регионов РФ – Смоленской (51 голова) и Тульской областей (90 голов) и симментальской породы, разводимой в условиях Красноярского края (105 голов). Породная принадлежность определялась согласно первичным данным зоотехнического учета.

При генотипировании гена TLR2 (NC\_037344.1 rs55617172) два аллеля экзона 2, отличающиеся заменой тимина (T) на гуанин (G) в нуклеотидной цепи (T385G) с последующим изменением аспарагина (N) на глутамин (Q) в аминокислотной последовательности, выявляли по наличию сайта рестрикции EcoRV [7]. Одноклеточный полиморфизм гена TNF (NC\_037350.1) в экзоне 4 был изучен с использованием рестриктазы RsaI [27].

Для амплификации участков генов TNF и TLR2 методом ПЦР были синтезированы следующие пары праймеров:

TLR2 – F: 5'-AGGTCAAATCACTGGACAATG-3' и R: 5'-GAGATGTTCCCCAAGTGT-3';  
 TNF – F: 5'-AGAGTAGAACTGACAGGGTCG-3' и R: 5'-CTCGGCATAGTCCAGGTAG-3' [7, 27].

Продукт амплификации для TLR2 составил 448 пн, для гена TNF – 445 пн. С помощью

электрофореза в 3%-ном агарозном геле против маркера pUC19 DNA/MspI и компьютерной системы гель-документирования была определена длина фрагментов рестрикции.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

По результатам ПЦР-ПДРФ анализа для гена *TLR2* выявлено два аллеля *G* и *T* и три генотипа (рис. 1). Генотип *GG* имеет длину фрагмента 448 пн, генотип *TG* – три фрагмента длиной 448, 241 и 207 пн, генотип *TT* – два фрагмента длиной 241 и 201 пн. В табл. 1 представлены частоты аллелей и генотипов гена *TLR2*.

Анализ полученных данных показал, что из шести исследованных пород крупного рогатого скота у пяти пород – холмогорской, голштинской, айрширской, костромской и симментальской – преобладает аллель *T* гена *TLR2*. Породы молочного направления продуктивности отличаются от пород молочно-мясного направления более высокой частотой встречаемости аллеля *T* ( $p < 0.001$ ). У айрширской породы молочного направления продуктивности отмечена наиболее высокая частота аллеля *T* (0.990). Среди коров айрширской породы 98% имеют генотип *TT* и лишь 2% – генотип *TG*. У коров черно-пестрого корня молочного направления продуктивности частота встречаемости аллеля *T* варьирует от 0.780 у голштинской породы до 0.830 у холмогорских коров и достоверно отличается от частоты этого аллеля у айрширской породы ( $p < 0.001$ ). В стадах голштинской и холмогорской пород преобладает гомозиготный генотип *TT* (61 и 69% соответственно). Между популяциями голштинского скота частота этого генотипа варьирует

от 50% у животных датской селекции из племенненного хозяйства Курской области до 67% у отечественных голштинских коров из Московской области. Доля гетерозиготных животных с генотипом *TG* среди холмогорских коров (27%) из Архангельской области в целом ниже, чем в стадах голштинских коров (34%). Гомозиготный генотип *GG* в стадах холмогорской и голштинской пород имеют в среднем 4–5% коров.

У коров молочно-мясного направления продуктивности частота аллеля *T* гена *TLR2* варьирует между породами в достоверно значимых пределах ( $p < 0.001$ ). Наименьшая частота этого аллеля отмечена в стадах животных бурой швицкой породы (0.260), где преобладает генотип *GG* с частотой от 52 до 57%, в среднем 55%. От 35% до 42% животных бурой швицкой породы имеют гетерозиготный генотип, в среднем 38%. Частота аллеля *T* у коров костромской породы достигает 0.670, что достоверно выше ( $p < 0.001$ ), чем у коров бурой швицкой породы. Большинство коров костромской породы имеет гетерозиготный генотип (47%), частота генотипа *TT* составляет 44%. У коров симментальской породы палевого корня также преобладает аллель *T* (0.620) и гетерозиготный генотип (54%), на втором месте – генотип *TT* с частотой 35%.

Полученные нами данные по частоте аллелей и генотипов *TLR2* у голштинской породы в целом соответствуют результатам исследований полиморфизма гена *TLR2* у коров голштинской породы в Китае. L. P. Zhang с соавт. выявили низкую частоту генотипа *GG* (< 5%) для *TLR2 T385G* в популяции голштинского скота в Китае [7], частота аллеля *G* составила 0.24. У симментальской породы частота аллеля *G* в Китае выше и достигает 0.55 [7]. Полиморфизм *T385G* был достоверно связан с количеством соматических клеток (SCS),  $p < 0.05$ . Коровы с генотипами *TT* и *TG* имели больший показатель SCS, чем коровы с генотипом *GG* [7]. В то же время у помесного скота в Индии выявлена высокая частота встречаемости генотипа *TT SNP* (rs55617172) гена *TLR2* в группе животных, у которых было подтверждено отсутствие мастита [34].

Исследование локуса *TLR2-EcoRV* у местных пород крупного рогатого скота в Турции показало преобладание аллеля *C(G)* у большинства исследованных пород (от 51.39 до 68.97%), за исключением местной черной породы Native Black, где частота аллеля *A(T)* составила 54.29% [35]. Исследование P. Raufian с соавт. [12] показало, что в группе восприимчивых к маститу животных голштинской породы частота аллеля *K(T)* гена *TLR2* в 2.5 раза превышала частоту аллеля *A(G)*.

В результате генотипирования гена *TNF* было выявлено наличие двух аллелей и трех генотипов (рис. 2). Генотип *AA* соответствует на электрофорезе одному фрагменту ДНК длиной 445 пн,

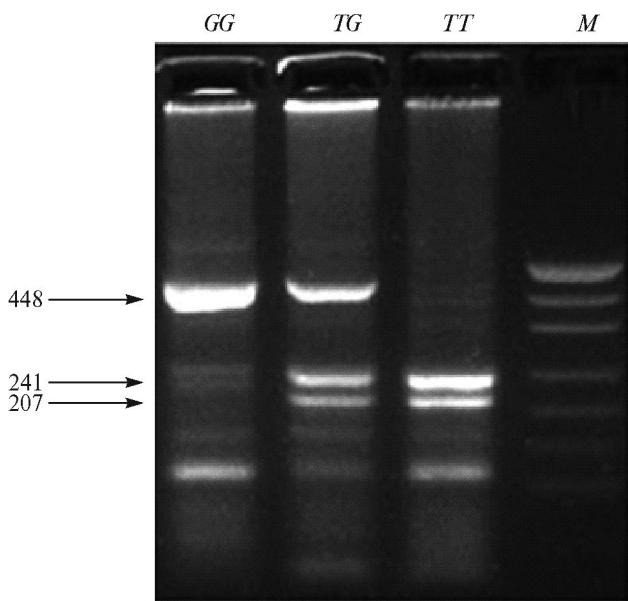
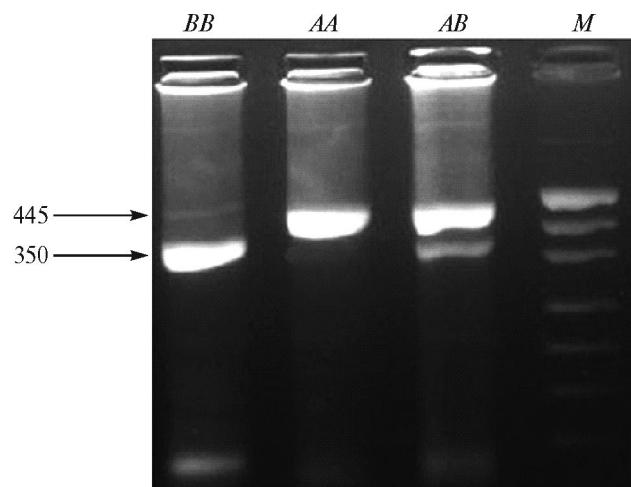


Рис. 1. Результаты электрофореза рестрикционных фрагментов гена *TLR2*. Маркер pUC19 / MspI.

Таблица 1. Полиморфизм гена *TLR2*

Хозяйство, область	<i>n</i>	Частота генотипа						Частота аллеля		<i>He</i>	$\chi^2$		
		<i>GG</i>		<i>TG</i>		<i>TT</i>		<i>G</i> $\pm$ <i>m<sub>G</sub></i>	<i>T</i> $\pm$ <i>m<sub>T</sub></i>				
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%						
Холмогорская порода													
ООО “Пежма”, Архангельская обл.	88	4	5	25	28	59	67	0.19 $\pm$ 0.03	0.81 $\pm$ 0.03	0.30	0.40		
ООО “Агрофирма “Холмогорская”, Архангельская обл.	32	1	3	7	22	24	75	0.14 $\pm$ 0.04	0.86 $\pm$ 0.04	0.24	0.29		
<b>В целом по холмогорской породе</b>	<b>120</b>	<b>5</b>	<b>4</b>	<b>32</b>	<b>27</b>	<b>83</b>	<b>69</b>	<b>0.17 <math>\pm</math> 0.02<sup>a***</sup></b> <b>b** c*** d***</b>	<b>0.83 <math>\pm</math> 0.02<sup>a***</sup></b> <b>c** d***</b>	<b>0.29</b>	<b>0.70</b>		
Голштинская порода													
ИП ГК(Ф)К Зубарева Н.В., Красноярский край	27	—	—	10	37	17	63	0.19 $\pm$ 0.05	0.81 $\pm$ 0.05	0.31	1.35		
ООО “Луч” Курская обл.	40	3	8	17	42	20	50	0.29 $\pm$ 0.05	0.71 $\pm$ 0.05	0.41	0.06		
ООО “Лесные поляны”, Московская обл.	39	1	2	12	31	26	67	0.18 $\pm$ 0.04	0.82 $\pm$ 0.04	0.30	0.78		
АО “Зеленоградское”, Московская обл.	40	3	8	11	27	26	65	0.21 $\pm$ 0.05	0.79 $\pm$ 0.05	0.34	1.27		
<b>В целом по голштинской породе</b>	<b>146</b>	<b>7</b>	<b>5</b>	<b>50</b>	<b>34</b>	<b>89</b>	<b>61</b>	<b>0.22 <math>\pm</math> 0.02<sup>e*** f*** g***</sup></b>	<b>0.78 <math>\pm</math> 0.02<sup>e*** f*** g*** h***</sup></b>	<b>0.34</b>	<b>0.00004</b>		
Айрширская порода													
СХПК “Племзавод Майский”, Вологодская обл.	100	0	0	2	2	98	98	0.01 $\pm$ 0.007 <sup>a*** e*** h*** i*** j***</sup>	0.99 $\pm$ 0.007 <sup>d*** f*** j*** k***</sup>	0.02	0.01		
Костромская порода													
СПК колхоз “Родина”, Костромская обл.	32	3	9	15	47	14	44	0.33 $\pm$ 0.06 <sup>b*** j*** k***</sup>	0.67 $\pm$ 0.06 <sup>c*** k*** l***</sup>	0.44	0.13		
Бурая швицкая порода													
АО “Смоленское”, Смоленская обл.	48	25	52	20	42	3	6	0.73 $\pm$ 0.05	0.27 $\pm$ 0.05	0.39	0.15		
ООО “Агрофармтрест”, Тульская обл.	90	51	57	32	35	7	8	0.74 $\pm$ 0.03	0.26 $\pm$ 0.03	0.38	0.39		
<b>В целом по бурой швицкой породе</b>	<b>138</b>	<b>76</b>	<b>55</b>	<b>52</b>	<b>38</b>	<b>10</b>	<b>7</b>	<b>0.74 <math>\pm</math> 0.03<sup>c*** f*** i*** k*** l***</sup></b>	<b>0.26 <math>\pm</math> 0.03<sup>b*** g*** j*** l*** m***</sup></b>	<b>0.39</b>	<b>0.07</b>		
Симментальская порода													
ЗАО “Сибирь-1”, Красноярский край	105	11	11	57	54	37	35	0.38 $\pm$ 0.03 <sup>d*** g*** h*** i*** l***</sup>	0.62 $\pm$ 0.03 <sup>a*** e*** h*** i*** m***</sup>	0.47	2.58		

Примечание.  $m_G(m_T)$  – ошибка частот аллелей,  $He$  – ожидаемая гетерозиготность,  $\chi^2$  – критерий соответствия. Пары различающихся генотипов при уровне достоверности  $**p < 0.01$ ,  $***p < 0.001$  обозначены буквенно-индексами а, б, с, д, е, ф, г, х, и, ж, к, л, м.



**Рис. 2.** Результаты электрофореза рестрикционных фрагментов гена *TNF*. Маркер pUC19 / *Msp*I.

генотип *AB* имеет три фрагмента длиной 445, 350 и 95 пн, а генотип *BB* – два фрагмента длиной 350 и 95 пн. Распределение частот аллелей и генотипов гена *TNF* у крупного рогатого скота представлено в табл. 2.

Согласно полученным данным, у всех шести изученных пород преобладал аллель *B* гена *TNF*. Наиболее высокие частоты встречаемости генотипа *BB* и аллеля *B* отмечены у пород бурого корня молочно-мясного направления продуктивности. Частота генотипа *BB* в стадах бурой швейцкой породы достигает в среднем 96%, частота аллеля *B* – 0.98. Среди коров отечественной костромской породы, также относящейся к бурому корню, 88% животных имеют генотип *BB*, частота аллеля *B* – 0.94. В стадах обеих пород бурого корня не обнаружены животные с генотипом *AA*. У коров симментальской породы, относящейся к палевому корню молочно-мясного направления продуктивности, частота аллеля *B* достоверно ниже, чем у пород бурого корня ( $p < 0.001$ ), и составляет 0.80, распределение генотипов гена *TNF*: *AA* – 3%, *AB* – 35%, *BB* – 62%.

Выявлены достоверные различия частот аллелей гена *TNF* у коров молочного направления продуктивности, относящихся к разным генетическим корням. У коров айрширской породы молочного направления продуктивности, представляющей красный корень, преобладает генотип *BB* с частотой 77% и аллель *B* с частотой 0.885, генотип *AA* не обнаружен. У коров черно-пестрого корня частота аллеля *B* достоверно ниже ( $p < 0.001$ ) и варьирует от 0.610 у голштинской породы до 0.720 у холмогорской породы. Частота аллеля *B* в стадах голштинской породы составляет от 0.540 у коров отечественной селекции до 0.650 у импортных животных. Частота гетерозиготного генотипа *AB* у голштинской породы варьирует в пределах от

40% у импортных животных до 62% у коров отечественной селекции, разводимых в хозяйствах Московской области, и в среднем составляет 53%. Наибольшее количество гомозигот *BB* обнаружено среди импортных голштинских животных, разводимых в Красноярском крае (41%) и Московской области (42%). Среди холмогорских коров также чаще встречается генотип *BB* (53%). От 5 до 18% голштинских коров имеют генотип *AA*, в среднем 12%; у коров холмогорской породы частота *AA* = 8%.

В исследовании P. Raufian с соавт. [12] выявлено преобладание генотипа *AA* гена *TNF* в группе устойчивых к маститу животных голштинской породы, в группе чувствительных к заболеванию животных чаще встречался генотип *BB*. Установлена достоверная корреляция между полиморфизмом гена *TNF* и *SCS* ( $p < 0.01$ ) и разность в экспрессии гена – самый высокий уровень экспрессии отмечен у животных с генотипом *BB*. Ученые полагают, что генотип *AA* может быть ассоциирован с относительной фенотипической устойчивостью к маститу в популяции голштинского скота [25]. А. J. Xu с соавт. выявили наибольшую частоту распространения гетерозиготного генотипа *AB* гена *TNF* у китайского голштинского скота. При этом у здорового скота преобладал аллель *A* (0.5237), а у животных, больных маститом, чаще встречался аллель *B* (0.65). Кроме того, была обнаружена статистически значимая связь ( $p < 0.05$ ) генотипа *BB* с высоким количеством соматических клеток *SCS* [27].

Таким образом, анализ результатов генотипирования SNP в ключевых генах иммунной системы *TNF* и *TLR2* у шести пород крупного рогатого скота разного происхождения выявил достоверные различия между породами разных направлений продуктивности ( $p < 0.001$ ) и между породами, относящимися к разным генетическим корням ( $p < 0.001$ ). По SNP T385G гена *TLR2* низкий уровень полиморфизма установлен у коров айрширской породы красного корня молочного направления продуктивности. Стадо айрширских коров близко к мономорфному типу по гену *TLR2*, поскольку 98% животных представлены одним гомозиготным генотипом *TT*.

Низкий уровень полиморфности выявлен по сайту *Rsa*I гена *TNF* в стадах животных бурого корня молочно-мясного направления продуктивности, где 96% животных представлены одним гомозиготным генотипом *BB*. Более равномерное распределение аллелей и генотипов гена *TNF* установлено в стадах голштинской и холмогорской пород черно-пестрого корня молочного направления продуктивности, где присутствует от 8 до 12% коров с генотипом *AA*, предположительно связанным с устойчивостью к маститу.

Таблица 2. Полиморфизм гена *TNF*

Хозяйство, область	n	Частота генотипа						Частота аллеля		<i>He</i>	$\chi^2$		
		<i>AA</i>		<i>AB</i>		<i>BB</i>		<i>A</i> $\pm$ <i>m<sub>A</sub></i>	<i>B</i> $\pm$ <i>m<sub>B</sub></i>				
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%						
Холмогорская порода													
ООО “Пежма”, Архангельская обл.	88	9	10	35	40	44	50	$0.30 \pm 0.04$	$0.70 \pm 0.04$	0.42	0.27		
ООО “Агрофирма “Холмогорская”, Архангельская обл.	32	1	3	12	38	19	59	$0.22 \pm 0.05$	$0.78 \pm 0.05$	0.34	0.30		
<b>В целом по холмогорской породе</b>	120	10	8	47	39	63	53	$0.28 \pm 0.03$ a** b*** c*** d***	$0.72 \pm 0.03$ a*** b*** c*** d***	0.40	0.09		
Голштинская порода													
ИП ГК(Ф)К Зубарева Н.В., Красноярский край	27	3	11	13	48	11	41	$0.35 \pm 0.07$	$0.65 \pm 0.07$	0.48	0.57		
ООО “Луч”, Курская обл.	40	2	5	24	60	14	35	$0.35 \pm 0.05$	$0.65 \pm 0.05$	0.46	4.06		
ООО “Лесные поляны”, Московская обл.	39	6	15	24	62	9	23	$0.46 \pm 0.06$	$0.54 \pm 0.06$	0.50	2.21		
АО “Зеленоградское”, Московская обл.	40	7	18	16	40	17	42	$0.38 \pm 0.05$	$0.62 \pm 0.05$	0.47	0.86		
<b>В целом по голштинской породе</b>	146	18	12	77	53	51	35	$0.39 \pm 0.03$ a** e*** f*** g*** h***	$0.61 \pm 0.03$ a*** f*** g*** h***	0.47	1.82		
Айрширская порода													
СХПК “Племзавод Майский”, Вологодская обл.	100	0	0	23	23	77	77	$0.115 \pm 0.022$ b*** e*** i*** j*	$0.885 \pm 0.022$ b*** e*** i*** j*	0.20	1.69		
Костромская порода													
СПК колхоз “Родина”, Костромская обл.	32	0	0	4	12.5	28	87.5	$0.06 \pm 0.03$ c*** f*** k**	$0.94 \pm 0.03$ f** k**	0.12	0.14		
Бурая швицкая порода													
АО “Смоленское”, Смоленская обл.	51	0	0	3	6	48	94	$0.03 \pm 0.01$	$0.97 \pm 0.01$	0.06	0.05		
ООО “Агрофармтрест”, Тульская обл.	90	0	0	2	2	88	98	$0.01 \pm 0.01$	$0.99 \pm 0.01$	0.02	0.04		
<b>В целом по бурой швицкой породе</b>	141	0	0	5	4	136	96	$0.02 \pm 0$ g*** i*** l***	$0.98 \pm 0.01$ g*** i*** l***	0.04	0.11		
Симментальская порода													
ЗАО “Сибирь-1”, Красноярский край	105	3	3	37	35	65	62	$0.20 \pm 0.03$ h*** j* k** l***	$0.80 \pm 0.03$ k** l***	0.33	0.71		

Примечание.  $m_A(m_B)$  – ошибка частот аллелей, *He* – ожидаемая гетерозиготность,  $\chi^2$  – критерий соответствия. Пары различающихся генотипов при уровне достоверности \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  обозначены буквенными индексами а, б, с, д, е, ф, г, х, и, ж, к, л.

В целом наличие выраженного полиморфизма по изученным генам иммунной системы у холмогорского и голштинского скота открывает возможность для дальнейших исследований и поиска ассоциаций с показателями здоровья животных, в частности с устойчивостью к маститу.

Исследование выполнено за счет гранта Российского научного фонда (проект № 23-26-00235).

Исследование одобрено Этическим комитетом ФГБНУ “Всероссийский научно-исследовательский институт племенного дела”, протокол № 1 от 02.02.2023 г.

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. *Bjelka M., Novák K.* Association of TLR gene variants in a Czech Red Pied cattle population with reproductive traits // *Vet. Immunol. Immunopathol.* 2020. V. 220. doi: 10.1016/j.vetimm.2019.109997
2. *Elmagharaby M.M., El-Nahas A.F., Fathala M.M. et al.* Association of toll-like receptors 2 and 6 polymorphism with clinical mastitis and production traits in Holstein cattle // *Iranian J. Veter. Res.* 2018. V. 19. № 3. P. 202–207.
3. *Puerto M.A., Shepley E., Cue R.I. et al.* The hidden cost of disease: I. Impact of the first incidence of mastitis on production and economic indicators of primiparous dairy cows // *J. Dairy Sci.* 2021. V. 104. № 7. P. 7932–7943. doi: 10.3168/jds.2020-19584
4. *Aghamohammadi M., Haine D., Kelton D.F. et al.* Herd-level mastitis-associated costs on canadian dairy farms // *Front. Veter. Sci.* 2018. V. 5. doi: 10.3389/fvets.2018.00100
5. *Trinchieri G., Sher A.* Cooperation of Toll-like receptor signals in innate immune defence // *Nat. Rev. Immunol.* 2007. V. 7. № 3. P. 179–190. doi:10.1038/nri2038
6. *McGuire K., Jones M., Werling D. et al.* Radiation hybrid mapping of all 10 characterized bovine Toll-like receptors // *Anim. Genet.* 2006. V. 37. № 1. P. 47–50.
7. *Zhang L.P., Gan Q.F., Ma T.H. et al.* Toll-like receptor 2 gene polymorphism and its relationship with SCS in dairy cattle // *Anim. Biotechnol.* 2009. V. 20. P. 87–95. doi: 10.1080/10495390902873096
8. *Takeda K., Akira S.* Toll-like receptors in innate immunity // *Int. Immunol.* 2005. V. 17. № 1. P. 1–14. doi: 10.1093/intimm/dxh186
9. *Akira S., Uematsu S., Takeuchi O.* Pathogen recognition and innate immunity // *Cell.* 2006. V. 124. № 4. P. 783–801. doi: 10.1016/j.cell.2006.02.015
10. *Ozinsky A., Underhill D.M., Fontenot J.D. et al.* The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between toll-like receptors // *Proc. Natl Acad. Sci. USA.* 2000. V. 97. № 25. P. 13766–13771. doi: 10.1073/pnas.250476497
11. *Inrnng Huang J., Liu L., Wang H. et al.* Variants and gene expression of the TLR2 gene and susceptibility to mastitis in cattle // *Asian J. Anim. Veter. Adv.* 2011. V. 6. P. 51–61. doi: 10.3923/ajava.2011.51.61
12. *Raufian P., Shodja Ghys J., Jafari R. et al.* Identification of genetic variation in two candidate genes of TLR2 and TNF $\alpha$  and its association with mastitis in Holstein Cattle // *Res. Anim. Prod. (Sci. Res.).* 2018. V. 8. P. 147–154. doi: 10.29252/rap.8.18.147
13. *Goldammer T., Zerbe H., Molenaar A. et al.* Mastitis increases mammary mRNA abundance of  $\beta$ -defensin 5, toll-like-receptor 2 (TLR2), and TLR4 but not TLR9 in cattle // *Clin. Diagn. Lab. Immunol.* 2004. V. 11. № 1. P. 174–185. doi: 10.1128/CDLI.11.1.174-185.2004
14. *Sadana T., Singh R., Singh Sh. et al.* Single nucleotide polymorphism of SLC11A1, CARD15, IFNG and TLR2 genes and their association with *Mycobacterium avium* subspecies paratuberculosis infection in native Indian cattle population // *Indian J. Biotechnol.* 2015. V. 14. P. 469–475.
15. *Koets A., Santema W., Mertens H. et al.* Susceptibility to paratuberculosis infection in cattle is associated with single nucleotide polymorphisms in Toll-like receptor 2 which modulate immune responses against *Mycobacterium avium* subspecies paratuberculosis // *Prev. Veter. Med.* 2010. V. 93. № 4. P. 305–315. doi: 10.1016/j.prevetmed.2009.11.008
16. *Bartens M.C., Gibson A.J., Etherington G.J. et al.* Single Nucleotide polymorphisms in the bovine TLR2 extracellular domain contribute to breed and species-specific innate immune functionality // *Front. Immunol.* 2021. V. 12. doi: 10.3389/fimmu.2021.764390
17. *Bhaladhare A., Sharma D., Kumar A. et al.* Single nucleotide polymorphisms in Toll-like receptor genes and case-control association studies with bovine tuberculosis // *Veter. World.* 2016. V. 9. № 5. doi: 10.14202/vetworld.2016.458-464
18. *Huang L.Y., Alberti J., Leifer C.A. et al.* Heat-killed *Brucella abortus* induces TNF and IL-12p40 by distinct MyD88-dependent pathways: TNF, unlike IL-12p40 secretion, is Toll-like receptor 2 dependent // *J. Immunol.* 2003. V. 171. № 3. P. 1441–1446. doi: 10.4049/jimmunol.171.3.1441

19. Aggarwal B.B., Kohr W.J., Hass P.E. et al. Human tumor necrosis factor. Production, purification, and characterization // *J. Biol. Chem.* 1985. V. 260. № 4. P. 2345–2354.
20. Pennica D., Nedwin G.E., Hayflick J.S. et al. Human tumour necrosis factor: Precursor structure, expression and homology to lymphotoxin // *Nature*. 1984. V. 312 (5996). P. 724–729.  
doi: 10.1038/312724a0
21. Agaba M., Kemp S.J., Barendse W., Teale A.J. Polymorphism at the bovine tumor necrosis factor alpha locus and assignment to BTA23 // *Mamm. Genome*. 1996. V. 7. P. 186–187.  
doi: 10.1007/s003359900051
22. Grosse W.M., Kappes S.M., Laegreid W.W. et al. Single nucleotide polymorphism (SNP) discovery and linkage mapping of bovine cytokine genes // *Mamm. Genome*. 1999. V. 10. № 11. P. 1062–1069.  
doi: 10.1007/s003359901162
23. Bannerman D.D. Pathogen-dependent induction of cytokines and other soluble inflammatory mediators during intramammary infection of dairy cows // *J. Anim. Sci.* 2009. V. 87 (13 Suppl.). P. 10–25.  
doi: 10.2527/jas.2008-1187
24. Bradley J.R. TNF-mediated inflammatory disease // *J. Pathol.* 2008. V. 214. № 2. P. 149–160.  
doi: 10.1002/path.2287
25. Raoofian P., Shoja J., Moghadam G.H., Javanmard A. Association investigation and gene expression in genotypes *TNF $\alpha$*  with somatic cell score in Holstein Dairy cattle // *Iranian J. Veter. Clin. Sci. Fall.* 2018. V. 12. № 2. P. 59–134.
26. Wojdak-Maksymiec K., Szyda J., Strabel T. Parity-dependent association between *TNF- $\alpha$*  and *LTF* gene polymorphisms and clinical mastitis in dairy cattle // *BMC Vet. Res.* 2013. V. 9.  
doi: 10.1186/1746-6148-9-114
27. Xu A.J., Liu X.L., Guo J.Z., Xia Z. Polymorphism of bovine TNF-a gene and its association with mastitis in Chinese Holstein cows // *Yi Chuan*. 2010. V. 32. № 9. P. 929–934.
28. Bojarojć-Nosowicz B., Kaczmarczyk E., Stachura A., Kotkiewicz M. Polymorphism in the promoter region of the tumor necrosis factor-alpha gene in cattle herds naturally infected and uninfected with the bovine leukemia virus // *Polish J. Vet. Sci.* 2011. V. 14. № 4. P. 671–673.  
doi: 10.2478/v10181-011-0101-0
29. Lendez P.A., Passucci J.A., Poli M.A. et al. Association of TNF- $\alpha$  gene promoter region polymorphisms in bovine leukemia virus (BLV)-infected cattle with different proviral loads // *Arch. of Virology*. 2015. V. 160. № 8. P. 2001–2007.  
doi: 10.1007/s00705-015-2448-5
30. Cheng Y., Huang C., Tsai H. Relationship of bovine TNF- $\alpha$  gene polymorphisms with the risk of bovine tuberculosis in Holstein cattle // *J. Vet. Med. Sci.* 2016. V. 78. № 5. P. 727–732.  
doi: 10.1292/jvms.15-0506
31. Kawasaki Y., Aoki Y., Magata F. et al. The effect of single nucleotide polymorphisms in the tumor necrosis factor- $\alpha$  gene on reproductive performance and immune function in dairy cattle // *J. Reprod. and Development*. 2014. V. 60. № 3. P. 173–178.  
doi: 10.1262/jrd.2013-140
32. Bimenova Zh.Zh., Makashev Y.K., Turyspayeva Sh.D. et al. The effect of the *TNF- $\alpha$*  gene alleles in Holstein cows on reproductive function // *J. Pharmaceutical Sci. and Res.* 2018. V. 10. № 4. P. 903–906.
33. Kahl S., Proszkowiec Wegla M.K., Connor E.E., Elsasser T.H. Association of tumor necrosis factor-alpha (TNF-alpha) gene promoter polymorphisms with TNF-alpha response to endotoxin (LPS) in calves // *J. Animal Sci. Abstract*. 2008. V. 86. P. 195.
34. Prebabathy T., Thanislass J., Dhanamal L. et al. Association between SNPs in *TLR2* gene segment corresponding to LRR functional domain of *TLR2* receptor and bovine mastitis // *Asian J. Animal Sci.* 2015. V. 9. P. 45–56.
35. Oner Y., Ata N., Karabaş M., Yilmaz O. Investigation of *TLR2* -EcoRV, NOD2- BsaHI ve IFN $\gamma$ - *HphI* allele frequency distribution among Turkish Native Cattle Breeds // *Turkish J. Agriculture. Food Sci. and Technol.* 2022. V. 10. № 6. P. 992–996.  
doi: 10.24925/turjaf.v10i6.992-996.4701

## Evaluation of Polymorphism of the Immune System Genes *TNF* and *TLR2* in Cattle

L. A. Kalashnikova<sup>1</sup>, I. E. Bagal<sup>1,\*</sup>, N. E. Murugina<sup>2</sup>, V. E. Kalashnikov<sup>1</sup>, R. Yu. Senina<sup>1</sup>

<sup>1</sup>All-Russian Research Institute of Animal Breeding, Moscow oblast, Pushkino, 141212 Russia

<sup>2</sup>National Research Center "Institute of Immunology", Moscow, 115522 Russia

\*e-mail: ladnatehplem@mail.ru

In this article the results of genotyping of tumor necrosis factor (*TNF*) and toll-like receptor 2 (*TLR2*) genes among 644 heads of cattle of six breeds belonging to four different genetic roots are presented. The Kholmogory and Holstein breeds of black-and-white root, the Ayrshire breed of red root, the Kostroma and Brown Swiss breed of brown root and the fawn Simmental breed were studied. The PCR-RFLP method was used to identify alleles and genotypes. Significant differences in allele frequency of the *TNF* and *TLR2* genes were observed between breeds of different genetic roots ( $p < 0.001$ ) and between breeds of dairy and dairy-meat productivity ( $p < 0.001$ ). In five of the six cattle breeds studied, the predominance of the *TLR2<sup>T</sup>* allele was revealed. The highest frequency of the *TLR2<sup>T</sup>* allele (0.99) was found in the Ayrshire breed. The herd of Ayrshire cows is close to being monomorphic for the *TLR2* gene, 98% of the animals have the homozygous *TLR2<sup>TT</sup>* genotype. A more uniform distribution of alleles and genotypes for the *TLR2* gene was found in black-and-white animals: in Holstein cattle the average frequency of the *TLR2<sup>T</sup>* allele was 0.78, in Kholmogory cattle it was 0.83. The *TLR2<sup>TT</sup>* genotype was detected in the majority of Holstein (61%) and Kholmogory (69%) cows. The frequency of the *TLR2<sup>T</sup>* allele in the Simmental breed was 0.62, in the Kostroma breed – 0.67. In Simmental and Kostroma cows, the heterozygous *TLR2<sup>TG</sup>* genotype is more common (54 and 47% respectively). A different picture of the distribution of allele and genotype frequencies was revealed in Brown Swiss animals. The *TLR2<sup>T</sup>* allele frequency averages 0.26. More than half of Brown Swiss cows (55%) are homozygous for *TLR2<sup>GG</sup>*. In all breeds, the *TNF<sup>B</sup>* allele predominated with a frequency of occurrence ranging from 0.61 in Holstein cattle to 0.98 in the Brown Swiss breed. A high frequency of the *TNF<sup>B</sup>* allele was found in the Kostroma (0.94), Ayrshire (0.885) and Simmental (0.80) breeds. In the Kholmogory breed, the average frequency of the *TNF<sup>B</sup>* allele was 0.72. The highest frequency of the *TNF<sup>BB</sup>* genotype was found in Brown Swiss herds (96%). The predominance of the homozygous *TNF<sup>BB</sup>* genotype was revealed in Kholmogory (53%), Simmental (62%), Ayrshire (77%) and Kostroma cattle (87.5%). In Holstein cows, the majority of animals (53%) carry the heterozygous *TNF<sup>AB</sup>* genotype. In three breeds – Ayrshire, Kostroma and Brown Swiss, the *TNF<sup>AA</sup>* genotype was not detected.

**Keywords:** cattle, immune system genes, *TNF*, *TLR2*.

## ХАРАКТЕРИСТИКА ПО STR-МАРКЕРАМ СЕРОГО УКРАИНСКОГО СКОТА, РАЗВОДИМОГО В РОССИЙСКОЙ ФЕДЕРАЦИИ

© 2024 г. А. С. Мокеев<sup>1,\*</sup>, Н. Н. Фурса<sup>1</sup>, С. В. Бекетов<sup>2,\*\*</sup>, Г. Р. Свищева<sup>2</sup>,  
А. А. Онохов<sup>2</sup>, Ю. А. Столповский<sup>2</sup>

<sup>1</sup>Институт животноводства степных районов им. М. Ф. Иванова “Аскания-Нова”

Национального научного селекционно-генетического центра по овцеводству Национальной академии аграрных наук, Аскания-Нова, 275230 Россия

<sup>2</sup>Институт общей генетики им. Н. И. Вавилова Российской академии наук, Москва, 119991 Россия

\*e-mail: mokeev.as@mail.ru

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

Поступила в редакцию 31.05.2024 г.

После доработки 19.06.2024 г.

Принята к публикации 28.06.2024 г.

С использованием 14 микросателлитных маркеров (*BM1824, BM2113, CSRM60, CSSM66, ETH3, ETH10, ETH225, ILSTS006, INRA023, SPS115, TGLA53, TGLA122, TGLA126, TGLA227*) проанализированы выборки стад серого украинского скота ИЖСР им. М. Ф. Иванова “Аскания-Нова” (Херсонская обл.) ( $n = 101$ ) и экспериментального хозяйства с. Черга (Республика Алтай) ( $n = 41$ ). При рассмотрении  $F$ -статистик Райта по исследованным локусам величина  $F_{IS} = -0.0285$  и указывает на избыток гетерозиготных генотипов в популяции серого украинского скота. При этом мала вероятность встречи аллелей общего предка  $F_{IT} = 0.1161$ , а уровень  $F_{ST} = 0.1394$  свидетельствует о средней дивергенции субпопуляций, наибольший вклад в которую вносит локус *Eth3*. При этом количество выявляемых полиморфных локусов – 95.64%, уровень аллельного –  $A_R = 7.66$ , генетического разнообразия –  $H_E = 0.76$  были больше в алтайской субпопуляции, чем в херсонской: 58.13, 4.41, 0.61% соответственно. В то же время херсонское стадо характеризовалось лучшей выравненностью, однородностью и консолидированностью, что позволяет рассматривать его в качестве источника ценного племенного материала при разведении серой украинской породы.

**Ключевые слова:** серый украинский скот, субпопуляции, микросателлиты, генотипирование, популяционно-генетическая характеристика.

**DOI:** 10.31857/S0016675824120081 **EDN:** WAAGJV

Серая украинская порода, прежние названия: серый степной, южно-русский, малороссийский, черкасский, черноморский и кубано-черноморский скот, является одной из древнейших пород, которая ведет свое происхождение от серого степного скота, включающего подольскую и романскую породы, в прошлом широко распространенных в южной части Европы, преимущественно в степной зоне Средиземноморья и Причерноморья. В настоящее время это одна из немногих реликтовых аборигенных пород, сохранившихся до наших дней. До начала двадцатого века серый украинский скот являлся основной рабочей и тягловой породой в Малороссии, а также ценился за мясные качества. Молочная продуктивность коров серой украинской породы небольшая и в среднем составляет 2950 кг, содержание жира в молоке – около 4.5% и белка – 3.45%. С конца XIX в. в связи с освоением

степных целинных земель Северного Причерноморья и Приазовья отмечалось постепенное снижение численности серого украинского скота, которое усилилось в начале XX в. на фоне коллективизации и механизации сельского хозяйства [1].

Однако уже в 1960-х гг. в Советском Союзе была принята государственная программа по сохранению серой украинской породы, в частности были сформированы два генофондных стада: первое в ОПХ “Поливановка” Украинского НИИ разведения и искусственного осеменения КРС (Днепропетровская обл.) (680 голов) и второе в опытном хозяйстве Украинского НИИ животноводства степных районов им. М. Ф. Иванова “Аскания-Нова” (Херсонская обл.) (около 2000 голов). В 1982 г. небольшая группа животных (34 особи) была перевезена из Аскании-Нова в с. Черга (Республика

Алтай) в экспериментальное хозяйство Института цитологии и генетики СО АН СССР [1, 2].

Тем не менее, несмотря на проводимые охранные мероприятия, общая численность серого украинского скота к 1980 г. снизилась до 1000 животных, а в 2007 г. – до 962 особей. В этом же году статус породы был оценен ФАО как критический [3] и к 2022 г. насчитывалось уже 738 особей [4].

Что касается непосредственно Российской Федерации, то серый украинский скот представлен здесь двумя субпопуляциями: около 260 животных содержится в Аскания-Нова и около 100 – в Алтайском экспериментальном племенном центре в с. Черга (Республика Алтай). Учитывая небольшую общую численность серого украинского скота, а также то, что генетическая структура этой породы в сравнительном аспекте мало изучена, цель настоящей работы заключается в том, чтобы с использованием STR-маркеров дать характеристику состояния генофондов двух локальных географически удаленных субпопуляций серого украинского скота, разводимых на территории Российской Федерации. При этом важно отметить, что STR-маркирование до настоящего времени остается наиболее доступным, достаточно информативным и востребованным методом при проведении популяционно-генетических исследований [5, 6].

## МАТЕРИАЛЫ И МЕТОДЫ

Материалом для исследований являлись образцы биоматериала (ушные выщипы) отдельных особей серого украинского скота ИЖСР им. М. Ф. Иванова (Херсонская обл.) (Askan,  $n = 101$ ), включая 97 коров и 4 быка (рис. 1, 2), полученные с использованием технологии FlexoPlusGeno (Caisley, Германия) [7]. Также исследовались пробы ДНК серого украинского скота экспериментального хозяйства с. Черга (Республика Алтай, Шебалинский р-н) (Altai,  $n = 41$ ) из коллекции лаборатории сравнительной генетики животных ИОГен РАН им. Н.И. Вавилова (Москва). ДНК из образцов



Рис. 1. Серый украинский скот (коровы).

хрящевой ткани наружного уха выделяли с помощью набора реагентов “ДНК-Экстрон-2” (ООО “Синтол”, Россия). Животные обеих рассматриваемых субпопуляций были прогенотипированы по 14 STR-маркерам (*BM1824, BM2113, CSRM60, CSSM66, ETH3, ETH10, ETH225, ILSTS006, INRA023, SPS115, TGLA53, TGLA122, TGLA126, TGLA227*).

Мультиплексный ПЦР-анализ микросателлитных локусов, содержащих короткие tandemные повторы, проводили с использованием набора COrDIS Cattle (ООО “Гордиз”, Россия) согласно инструкции производителя. После начальной денатурации (94°C, 3 мин) амплификацию фрагментов осуществляли в следующем температурно-временном режиме: 4 цикла при 98°C 30 с, 59°C 120 с, 72°C 90 с, 6 – при 94°C 30 с, 59°C 120 с, 72°C 90 с, 18 – при 90°C 30 с, 59°C 120 с, 72°C 75 с и окончательным удлинением при 68°C 15 мин. Полимеразную цепную реакцию выполняли на термоциклиере Veriti (Thermo Fisher Scientific, США). Капиллярный электрофорез продуктов ПЦР-реакции осуществляли на генетическом анализаторе ABI 3130XL (Applied Biosystems, США) с последующим определением длин аллелей микросателлитов в программном обеспечении GeneMapper v. 4.0. (Applied Biosystems, США).

Основные популяционно-генетические характеристики для каждой выборки были рассчитаны в R-пакете degenet [8]. Частоты аллелей, аллельное богатство и количество частных аллелей оценивали с помощью R-пакета PopGenReport. Для каждой пары “локус–популяция” индекс фиксации, а также наблюдаемую и ожидаемую гетерозиготности оценивали с помощью функции *divBasic* из R-пакета divRsity [9]. *F*-статистики ( $F_{IT}$ ,  $F_{ST}$  и  $F_{IS}$ ) для каждого локуса рассчитывали с помощью функции  $F_{ST}$  из R-пакета pegas. Дендрограмма, основанная на генетических расстояниях Нея, была построена с помощью алгоритма объединения соседей (NJ), реализованного в R-пакете poppr, анализ главных компонент по исследованным локусам (PCA



Рис. 2. Серый украинский скот (бык).

— Principal Component Analysis) проводили с помощью программного обеспечения PopGenReport [10, 11].

## РЕЗУЛЬТАТЫ

В ходе сравнительного анализа двух географически удаленных популяций серого украинского скота было установлено, что количество полиморфных локусов — 95.64%, и уровень аллельного разнообразия — 7.66, были больше в алтайской субпопуляции, по сравнению с асканийской: 58.13 и 4.41% соответственно (табл. 1).

В то же время значения наблюдаемой ( $H_O$ ) и ожидаемой ( $H_E$ ) гетерозиготностей оказались одинаковыми в алтайской субпопуляции ( $H_O = 0.76$ ,  $H_E = 0.76$ ) и достаточно близкими между собой в асканийской ( $H_O = 0.64$ ,  $H_E = 0.61$ ). При этом в последней при избытке гетерозигот уровень генетической изменчивости был несколько меньше —  $H_E = 0.61$ , по сравнению с алтайской —  $H_E = 0.76$ . На этом фоне рассчитанный коэффициент инбридинга  $F_{IS}$  алтайской популяции был равен 0, в то время как у асканийского скота отмечали незначительный избыток гетерозигот —  $F_{IS} = -0.05$  (табл. 1).

**Таблица 1.** Параметры аллельного и генетического разнообразия субпопуляций серого украинского скота

Локус	Субпопуляция									
	асканийская (Askan)					алтайская (Altai)				
	P, %	$A_R$	$H_O$	$H_E$	$F_{IS}$	P, %	$A_R$	$H_O$	$H_E$	$F_{IS}$
<i>ETH3</i>	66.67	3.7	0.3	0.31	0.03	83.33	5	0.83	0.74	-0.12
<i>CSSM66</i>	60	4.5	0.69	0.58	-0.19	90	8.59	0.88	0.82	-0.07
<i>INRA23</i>	60	5.86	0.83	0.76	-0.09	90	8.24	0.8	0.83	0.04
<i>ILSTS6</i>	75	5.63	0.66	0.64	-0.03	100	6.96	0.76	0.79	0.04
<i>TGLA227</i>	60	5.71	0.77	0.77	0	100	9.75	0.73	0.82	0.11
<i>TGLA126</i>	83.33	4.53	0.73	0.64	-0.14	100	5.69	0.73	0.61	-0.20
<i>TGLA122</i>	33.33	4.42	0.53	0.58	0.09	86.67	11.9	0.83	0.83	0
<i>SPS115</i>	50	3.02	0.54	0.5	-0.08	100	6.6	0.54	0.61	0.11
<i>ETH225</i>	55.56	4.57	0.73	0.66	-0.11	88.89	7.22	0.71	0.78	0.09
<i>TGLA53</i>	46.15	4.88	0.64	0.64	0	100	11.76	0.76	0.84	0.10
<i>CSRM60</i>	66.67	3.95	0.59	0.61	0.03	100	5.96	0.83	0.79	-0.05
<i>BM2113</i>	57.14	3.94	0.6	0.59	-0.02	100	6.99	0.83	0.83	0
<i>BM1824</i>	57.14	4.00	0.69	0.62	-0.11	100	6.22	0.68	0.71	0.04
<i>ETH10</i>	42.86	3.00	0.64	0.63	-0.02	100	6.43	0.68	0.66	-0.03
Средняя	58.13	4.41	0.64	0.61	-0.05	95.64	7.66	0.76	0.76	0

Примечание.  $P$  — доля полиморфных локусов;  $A_R$  — аллельное разнообразие;  $H_O$  — наблюдаемая гетерозиготность;  $H_E$  — ожидаемая гетерозиготность;  $F_{IS}$  — коэффициент инбридинга.

В ходе дальнейшего анализа на уникальность выявляемых аллелей было обнаружено 50 приват-аллелей в алтайской субпопуляции и 6 — в асканийской (табл. 2). В то же время в алтайской субпопуляции серого скота — только 10 уникальных аллелей (20% от их общего числа) превышали вероятность обнаружения (равную 0.05); в асканийской субпопуляции число таких аллелей составило 3 или 50%. Причем, если в алтайской субпопуляции приватные аллели выявлялись преимущественно с низкой частотой и в узких пределах обнаружения — 0.06–0.14, то у асканийского скота, несмотря на их немногочисленность, приват аллели встречались чаще и с более широким диапазоном варирирования — 0.16–0.39 (табл. 2).

При рассмотрении  $F$ -статистик по исследованным локусам обеих субпопуляций значение  $F_{IS} = -0.0285$  указывает на избыток гетерозиготных генотипов с малой вероятностью встречи аллелей общего предка в рассматриваемых популяциях  $F_{IT} = 0.1161$ , а уровень  $F_{ST} = 0.1394$  свидетельствует об их средней дивергенции, наибольший вклад в которую вносит локус *Eth3* (табл. 3). Согласно дендрограмме Neighbor-Joining взаимоотношений можно видеть два больших кластера,

**Таблица 2.** Приват-аллели в локусах и субпопуляциях серого украинского скота

Локус	Субпопуляция	Аллель	Частота, %	Локус	Субпопуляция	Аллель	Частота, %
<i>ETH3</i>	Askan	115	26	<i>TGLA122</i>	Altai	175	4
<i>ETH3</i>	Altai	121	9	<i>TGLA122</i>	Askan	179	3
<i>ETH3</i>	Altai	127	13	<i>TGLA122</i>	Altai	183	1
<i>CSSM66</i>	Altai	181	1	<i>TGLA122</i>	Altai	185	4
<i>CSSM66</i>	Altai	187	3	<i>SPS115</i>	Altai	246	1
<i>CSSM66</i>	Altai	195	5	<i>SPS115</i>	Altai	252	4
<i>CSSM66</i>	Altai	197	6	<i>SPS115</i>	Altai	258	1
<i>CSSM66</i>	Askan	199	1	<i>SPS115</i>	Altai	262	1
<i>INRA23</i>	Altai	196	5	<i>ETH225</i>	Altai	142	5
<i>INRA23</i>	Altai	200	13	<i>ETH225</i>	Askan	146	16
<i>INRA23</i>	Askan	202	39	<i>ETH225</i>	Altai	152	3
<i>INRA23</i>	Altai	210	4	<i>ETH225</i>	Altai	154	1
<i>INRA23</i>	Altai	216	1	<i>ETH225</i>	Altai	156	1
<i>ILSTS6</i>	Altai	282	1	<i>TGLA53</i>	Altai	154	4
<i>ILSTS6</i>	Altai	300	1	<i>TGLA53</i>	Altai	162	2
<i>TGLA227</i>	Altai	079	3	<i>TGLA53</i>	Altai	164	4
<i>TGLA227</i>	Altai	091	5	<i>TGLA53</i>	Altai	176	3
<i>TGLA227</i>	Altai	097	14	<i>TGLA53</i>	Altai	178	4
<i>TGLA227</i>	Altai	101	3	<i>TGLA53</i>	Altai	182	2
<i>TGLA126</i>	Altai	123	2	<i>TGLA53</i>	Altai	186	1
<i>TGLA122</i>	Altai	139	1	<i>CSRM60</i>	Altai	096	7
<i>TGLA122</i>	Altai	141	10	<i>CSRM60</i>	Altai	104	3
<i>TGLA122</i>	Altai	145	2	<i>BM2113</i>	Altai	125	4
<i>TGLA122</i>	Altai	153	2	<i>BM2113</i>	Altai	131	14
<i>TGLA122</i>	Altai	161	5	<i>BM2113</i>	Altai	135	12
<i>TGLA122</i>	Altai	163	4	<i>BM1824</i>	Altai	176	1
<i>TGLA122</i>	Askan	169	3	<i>BM1824</i>	Altai	184	1
<i>TGLA122</i>	Altai	173	3	<i>BM1824</i>	Altai	186	7

сформированных алтайской и асканийской субпопуляциями (рис. 3). При этом серый украинский скот из Черги делится еще на два достаточно удаленных подклUSTERов (или две линии разведения), в один из которых попадают пять особей асканийской субпопуляции, что подтверждает их отдаленные родственные связи [12]. Что же касается асканийской субпопуляции, то она образует куст из нескольких сближенных подклUSTERов (всего восемь родственных групп), отражающих ее множественную линейную структуру.

При последующем рассмотрении тех же самых субпопуляций с использованием метода главных компонент (PCA) (рис. 4) наблюдается их четкое разграничение с выраженной неоднородностью особей в алтайской и консолидацией животных в асканийской субпопуляции.

## ОБСУЖДЕНИЕ

Можно видеть, что при общем рассмотрении популяций серого украинского скота с использованием *F*-статистик Райта отмечается незначительный

**Таблица 3.** *F*-статистики по исследованным локусам субпопуляций серого украинского скота

Локус	$F_{IT}$	$F_{ST}$	$F_{IS}$
<i>ETH3</i>	0.20948280	0.23848947	-0.038090963
<i>CSSM66</i>	0.08202644	0.19849845	-0.145317266
<i>INRA23</i>	0.07071799	0.11239222	-0.046951181
<i>ILSTS6</i>	0.17892123	0.18257128	-0.004465281
<i>TGLA227</i>	0.11005591	0.07895635	0.033765563
<i>TGLA126</i>	0.08620622	0.21275222	-0.160744810
<i>TGLA122</i>	0.23889019	0.19515622	0.054338459
<i>SPS115</i>	0.01416733	0.02679428	-0.012974590
<i>ETH225</i>	0.07487413	0.10363076	-0.032081235
<i>TGLA53</i>	0.18498442	0.15169638	0.039240709
<i>CSRM60</i>	0.13075568	0.12568836	0.005795789
<i>BM2113</i>	0.12618519	0.13532227	-0.010567033
<i>BM1824</i>	0.08929996	0.14285507	-0.062480814
<i>ETH10</i>	0.02944797	0.04738194	-0.018825989
Средняя	0.116143961	0.139441805	-0.028525617

избыток гетерозигот с отсутствием инбридинга  $F_{IT} = 0.116 < F_{ST}$  (табл. 1), причем в большей степени это относится к херсонской популяции. Все это указывает на то, что в обоих стадах серого украинского скота ведется работа по сохранению генетического разнообразия с той лишь разницей, что более низкий уровень аллельного разнообразия ( $A_R$ ) при большей генетической изменчивости ( $H_E$ ), обнаруживаемый в асканийской субпопуляции, свидетельствует о более интенсивном в ней фенотипическом отборе. Об этом же свидетельствует и большая консолидированность асканийской субпопуляции по данным РСА-анализа.

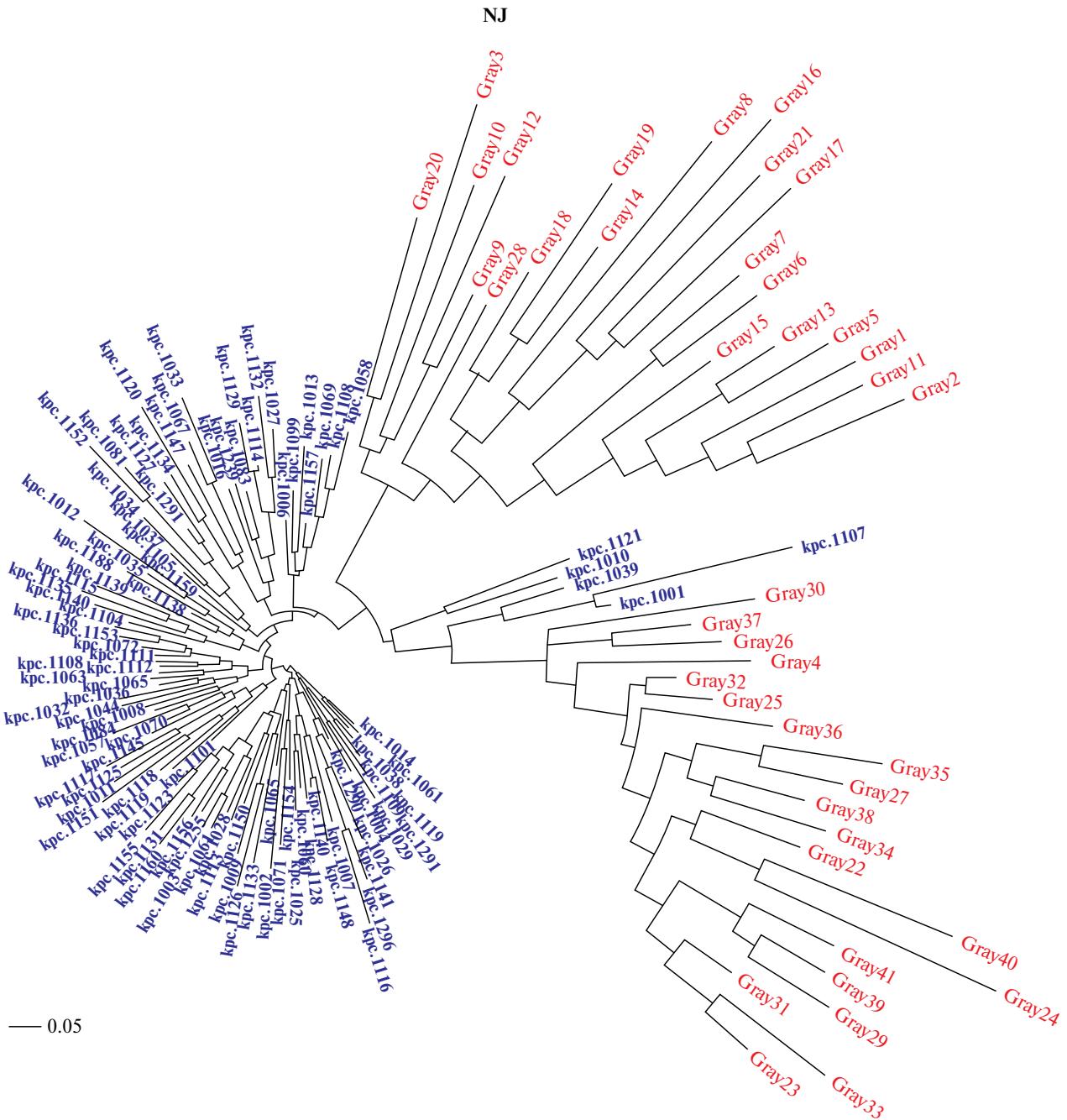
В связи с вышесказанным интересно отметить, что в ранее проведенных исследованиях при оценке аллельного полиморфизма по гену *BoLA-DRB3* с применением индекса Шеннона – Винера ( $H'$ ) самые высокие его значения (3.77 и 3.4) были выявлены у красно-пестрой и голштинских пород КРС и самое низкое – у якутской породы (1.74). Для серого украинского скота асканийской субпопуляции аллельный полиморфизм *BoLA-DRB3* составил 3.26, что сближает его по этому показателю со специализированными молочными породами [13].

В свою очередь, расчет значения  $F_{ST}$  (табл. 3), отражающего степень подразделенности субпопуляций, показал, что 86.1% выявленной у серого украинского скота изменчивости реализуется внутри популяции, включающей асканийскую и алтайскую субпопуляции, и 13.4% ( $F_{ST} = 0.134$ ) приходится на долю межпопуляционной составляющей.

В связи с этим интерес представляет оценка величины  $M$  [14], характеризующей интенсивность потока генов между субпопуляциями. В нашем случае  $M = 1.61$  и  $2M > 1$ , что свидетельствует об отсутствии тенденции к дивергенции рассматриваемых субпопуляций по аллельным частотам [15].

Что же касается величины ожидаемой генетической гетерозиготности ( $H_E$ ), то характерная гетерогенность генетического разнообразия, обнаруживаемая у серого украинского скота, свойственна и для других пород КРС, в том числе региональных и трансграничных. Например, по черно-пестрой породе в разных субпопуляциях величина ожидаемой гетерозиготности  $H_E$  варьировала от минимальной 0.417 до максимальной 0.751 (коэффициент вариации  $Cv = 19.7\%$ ) при среднем значении  $0.608 \pm 0.042$  и 95% доверительном интервале (ДИ) =  $0.526 \div 0.690$ . По голштинской породе значения  $H_E$  изменились от 0.430 до 0.891 (коэффициент вариации 19.4%) при среднем –  $0.639 \pm 0.026$  (95% ДИ,  $0.588 \div 0.690$ ) [16].

Причем одни авторы подобный внутривидовой разброс значений  $H_E$  связывают с использованием в вычислениях малочисленных нерепрезентативных выборок, в то время как другие склонны объяснять его характером динамики происходящих во времени внутривидовых процессов [12, 17]. Принимая во внимание данные точки зрения и с учетом объема выборочных данных, использованных в настоящем исследовании, можно было бы говорить об асканийской субпопуляции как

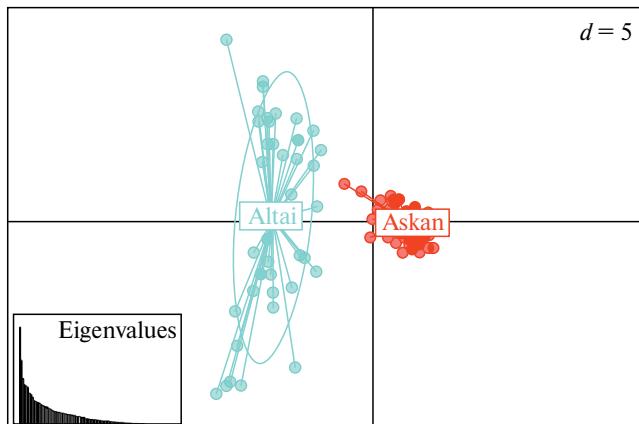


**Рис. 3.** NJ-дендограмма взаимоотношений исследуемых субпопуляций серого украинского скота, построенная на основе попарных генетических дистанций  $D_A$  Нея. Синим цветом обозначена асканийская субпопуляция, красным – алтайская.

наиболее объективно отражающей фактическую генетическую изменчивость серого украинского скота. Наблюдаемые различия  $H_E$  в двух рассматриваемых субпопуляциях также можно объяснить уменьшением поголовья в асканийском стаде и его увеличением в алтайском. Причем сходная картина наблюдалась ранее на Украине в отдельных субпопуляциях серого украинского скота (ООО

“Голосеево” и КСП “Вороньков”), где с ростом численности животных увеличивался показатель генотипической изменчивости ( $H_E$ ) до 0.80 и 0.81 соответственно [18].

Таким образом, по результатам сравнительного анализа асканийского и алтайского серого украинского скота в настоящей работе установлено, что рассматриваемые популяции по изучаемым



**Рис. 4.** Проекция особей исследуемых субпопуляций серого украинского скота на плоскости двух координат по данным PCA-анализа.

микросателлитным локусам характеризуются отсутствием инбридинга, отдаленными родственными связями и различались между собой показателями аллельного и генетического разнообразия: большим для алтайской микропопуляции и меньшим для херсонской. Эти данные свидетельствуют о том, что алтайская популяция по своим генетическим характеристикам была менее подвержена селекционному давлению. В то же время асканийское стадо серого украинского скота при большем линейном разнообразии и избыточной гетерозиготности отличается выравненностью, однородностью и консолидированностью, что позволяет рассматривать его в качестве источника ценного племенного материала при разведении этой породы.

Исследования выполнены при финансовой поддержке РНФ – проект № 23-16-00059 и Государственного задания Министерства науки и высшего образования РФ “Оценка генетического потенциала национальных пород крупного рогатого скота” № 122020800034-4.

Исследование одобрено Этическим комитетом Института общей генетики им. Н. И. Вавилова РАН 17 июня 2024 г., протокол № 3.

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. Столповский Ю.А. Серая украинская порода // Генетические ресурсы крупного рогатого скота: редкие и исчезающие отечественные породы. М.: Наука, 1993. С. 94–111.
2. Glazko V.I. Podolic cattle in the Ukraine and eastern territories // Stočarstvo: Časopis za unapređenje stočarstva. 2001. V. 55. № 1. P. 33–60.
3. The state of the world's animal genetic resources for food and agriculture. Rome, FAO, 2007. 511 p.
4. Ukrainian grey/Ukraine (cattle). Domestic animal diversity information system of the food and agriculture organization of the United Nations. Accessed April 2023. <https://www.fao.org/dad-is/browse-by-country-and-species/en/>
5. Dakin E.E., Avise J.C. Microsatellite null alleles in percentage analysis // Heredity. 2004. V. 93. P. 504–509. doi: 10.1038/sj.hdy.6800545
6. Olschewsky A., Hinrichs D. An overview of the use of genotyping techniques for assessing genetic diversity in local farm animal breeds // Animals. 2021. V. 11. № 7. P. 2016. <https://www.mdpi.com>. doi: 10.3390/ani11072016
7. Бекетов С.В., Семина М.Т., Мокеев А.С. и др. Перспективы применения технологии “генетического биркования” в животноводстве // Главный зоотехник. 2024. № 5. С. 3–15.
8. Jombart T. Adegenet: A R package for the multivariate analysis of genetic markers // Bioinformatics. 2008. V. 24. P. 1403–1405. doi: 10.1093/bioinformatics/btn129
9. Keenan K., McGinnity P., Cross T.F. et al. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors // Methods Ecol. and Evol. 2013. № 4. P. 782–788. doi:10.1111/2041-210X.12067
10. Adamack A.T., Gruber B. PopGenReport: Simplifying basic population genetic analyses in R. Methods // Ecology and Evolution. 2014. № 5. P. 384–387. doi: 10.1111/2041-210X.12158
11. Gruber B., Adamack A.T. landgenreport: A new R function to simplify landscape genetic analysis using resistance surface layers // Mol. Ecol. Resources. 2015. V. 15. P. 1172–1178. doi: 10.1111/1755-0998.12381
12. Кушнир А.В., Глазко В.И. Серый украинский скот и его близкородственные формы // Сиб. экол. журнал. 2009. Т. 16. № 3. С. 495–506.
13. Копылов К.В., Стародуб Л.Ф., Мохначева Н.Б., Супрович Н.П. Особенности изменчивости генома крупного рогатого скота серой украинской породы по цито- и ДНК-маркерам // Акт. пробл. интенсивного развития животноводства. 2019. Вып. 1. С. 60–69.
14. Holsinger K.E. Lecture notes in population genetics. Stanford: University of Connecticut, 2010. 275 p. doi: 10.6084/M9.FIGSHARE.100687

15. Кузнецов В.М. F-статистики Райта: оценка и интерпретация // Пробл. биол. продуктивных животных. 2014. № 4. С. 80–104.
16. Кузнецов В.М. Снижает ли кроссбридинг генетическое разнообразие? Разведение и сохранение пород молочного скота в России. Киров: НИИСХ Северо-Востока, 2017. 104 с.
17. Кузнецов В.М. Сравнение методов оценки генетической дифференциации популяций по микросателлитным маркерам // Зоотехния. 2020. Т. 21. № 2. С. 169–182. doi:10.30766/2072-9081.2020.21.2.169-182
18. Гузеев Ю.В., Мельник О.В., Спиридонос В.Г., Мельничук С.Д. Сравнительный анализ генетической структуры микропопуляции серой украинской породы крупного рогатого скота по ДНК-маркерам // Науковий вісник ЛНУВМБТ імені С. З. Гжицького. 2015. Т. 17. № 3(63). С. 166–171.

## Characteristics by STR Markers of Gray Ukrainian Cattle Breed in the Russian Federation

**A. S. Mokeev<sup>a,\*</sup>, N. N. Fursa<sup>a</sup>, S. V. Beketov<sup>b, \*\*</sup>, G. R. Svishcheva<sup>b</sup>,  
A. A. Onokhov<sup>b</sup>, Yu. A. Stolpovsky<sup>b</sup>**

<sup>a</sup>Ivanov Institute of Livestock Husbandry of Steppe Regions “Askania-Nova” – National Scientific Selection and Genetic Center for Sheep Breeding of the National Academy of Agrarian Sciences, Askania-Nova, 275230 Russia

<sup>b</sup>Vavilov Institute of General Genetics of Russian Academy of Sciences, Moscow, 119991 Russia

\*e-mail: mokeev.as@mail.ru

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

Using 14 microsatellite markers (*BM1824, BM2113, CSRM60, CSSM66, ETH3, ETH10, ETH225, ILSTS006, INRA023, SPS115, TGLA53, TGLA122, TGLA126, TGLA227*), samples of Ukrainian Grey cattle herds of the M.F. Ivanov Institute “Askania-Nova” (Kherson region) ( $n = 101$ ) and experimental farm of Cherga (Altai Republic) ( $n = 41$ ). When considering Wright's F-statistics for the studied loci, the value  $F_{IS} = -0.0285$  indicates an excess of heterozygous genotypes in the population of Ukrainian Grey cattle, with a low probability of encountering alleles of the common ancestor  $F_{IT} = 0.1161$ , and the level  $F_{ST} = 0.1394$  indicates an average divergence of subpopulations, the largest contribution to which is contributed by the Eth3 locus. At the same time, the number of detected polymorphic loci – 95.64%, the level of allelic diversity –  $A_R = 7.66$  and genetic diversity –  $H_E = 0.76$  were higher in the Altai subpopulation than in the Kherson subpopulation – 58.13%, 4.41, 0.61, respectively. At the same time, the Kherson herd was characterized by better evenness, homogeneity and consolidation, which allows us to consider it as a source of valuable breeding material for breeding the Ukrainian Grey breed.

**Keywords:** Ukrainian Grey cattle, subpopulations, microsatellites, genotyping, population genetic characteristics.

## ПОПУЛЯЦИОННО-ГЕНЕТИЧЕСКАЯ СТРУКТУРА НАСЕЛЕНИЯ СИБИРИ ПО ДАННЫМ О ЧАСТОТАХ ПОЛИМОРФНЫХ ВАРИАНТОВ ГЕНОВ ТРАНСПОРТА И РЕЦЕПЦИИ ВИТАМИНА D

© 2024 г. Б. А. Тхоренко<sup>1,\*</sup>, А. В. Мейер<sup>1</sup>, Г. В. Вавин<sup>1</sup>, Д. О. Имекина<sup>1</sup>,  
М. В. Ульянова<sup>1</sup>, Ф. А. Лузина<sup>2</sup>, М. Б. Лавряшина<sup>1</sup>

<sup>1</sup> Кемеровский государственный медицинский университет, Кемерово, 650056 Россия

<sup>2</sup> Научно-исследовательский институт комплексных проблем гигиены и профессиональных заболеваний, Новокузнецк, 654041 Россия

\*e-mail: tba2008@mail.ru

Поступила в редакцию 21.05.2024 г.

После доработки 09.08.2024 г.

Принята к публикации 12.08.2024 г.

Генофонды народов Сибири – потенциальный источник данных о составах гаплотипов полиморфных вариантов генов, сформировавшихся в ходе эволюции популяций под влиянием специфического комплекса природно-климатических факторов, в условиях относительной географической, культурной изоляции и традиционных практик жизнеобеспечения. С целью изучения генетической адаптации к недостаточной обеспеченности витамином D (низкий уровень инсоляции и доступности нутриентов, содержащих холекальциферол и эргокальциферол) в выборках коренного тюркоязычного (телеуты, томские татары, шорцы) и пришлого восточнославянского (русские) населения Сибири проведено исследование частот полиморфных вариантов (14 SNP) генов, отвечающих за транспорт витамина D – ген *GC* (Group-Specific Component), структуру его основного рецептора – ген *VDR* (vitamin D receptor) и рецепторов-партнеров *RXR* (retinoid X receptor) – гены *RXRA* и *RXRG*. Суммарный объем выборки составил 411 человек. Анализ частот полиморфных вариантов продемонстрировал специфические черты генофондов коренных популяций Сибири по сравнению с русским населением, общемировыми частотами и частотами, характерными для популяций Европы и Восточной Азии. При попарном сравнении с выборкой русских статистически значимые отличия выявлены в отношении частот rs7041 (шорцы), rs3847987 (томские татары, телеуты, шорцы), rs9409929 (томские татары, телеуты), rs877954 (томские татары), rs283696 (телеуты). По данным усредненных генетических расстояний (*d*) минимальные дистанции отмечены между телеутами и томскими татарами. Выборка русских в целом оказалась наиболее удаленной в генетическом пространстве исследованных народов Сибири. Исследование частот гаплотипов полиморфных вариантов генов транспорта и рецепции витамина D продемонстрировало накопление в генофонде коренных популяций SNP, ассоциированных, по данным литературы, с более высоким уровнем сывороточного витамина D. По генам *GC* и *VDR* гаплотипический профиль изученных полиморфных вариантов характеризовался выраженной спецификой и отличался по составу от такового у русского населения. Проведенное исследование позволяет сделать заключение о том, что структура генофондов коренных народов Сибири сохраняет своеобразие, отражающее историю их сложения и демонстрирующее черты генетической адаптации коренных популяций к условиям обитания.

**Ключевые слова:** витамин D, гены, коренные народы, полиморфизм, популяция, рецепторы, русские, Сибирь.

**DOI:** 10.31857/S0016675824120092 **EDN:** VZZYTO

Витамин D в основной активной форме (кальциферол, эргокальциферол [1]) и в виде метаболитов (около 50 вариантов с различиями по функциональной активности и тканям-мишеням) – регулятор широкого спектра процессов (ремоделирование хроматина, клеточный рост и

дифференцировка, фосфорно-кальциевый обмен, иммунный ответ и др.), действующий на геномном и негеномном уровнях [2–5]. Эффективность трансдукции витамина D детерминируется балансом взаимодействий эндогенных и экзогенных факторов, обусловливающих его метаболизм,

транспорт и рецепцию [5, 6]. Природно-климатические условия, а также особенности питания, определяемые доступностью нутриентов, этнокультурными и религиозными традициями, формируют пул экзогенных факторов [7]. К эндогенным относятся особенности составов микробиомных сообществ желудочно-кишечного тракта [8–10]. Однако ведущую роль в этом ключе, несомненно, играют генные комплексы, сложившиеся в процессе приспособления индивидуальных геномов к среде обитания и закрепившиеся в генофондах популяций [11, 12].

Подавляющее большинство регионов России и Сибирь в их числе относятся к территориям, для населения которых характерен недостаток или даже дефицит витамина D [13, 14]. Это обусловлено коротким периодом инсоляции на достаточном уровне и низкой обеспеченностью нутриентами, содержащими холекальциферол (D3) и эргокальциферол (D2). В этом контексте значительный интерес представляют генофонды коренных сибирских народов. Их сложение до недавнего времени проходило под влиянием специфического комплекса природно-климатических факторов, в условиях относительной географической, культурной изоляции и особенностей питания. Ряд работ [15–19] демонстрирует специфику структуры генофондов коренных популяций, свидетельствующую о приспособлении к условиям окружающей природной среды. Прицельное изучение особенностей генофондов народов Сибири может послужить источником данных о вариантах аллелей, встречающихся в популяции с высокой частотой. Дальнейшие исследования *in silico* на клеточных моделях или иными способами – путь к прецизионной медицине и персонализированному нутрициологическому подходу [20, 21] для профилактики и коррекции алиментарных заболеваний.

В рамках сформулированной задачи в выборках коренного тюркоязычного (телеуты, томские татары, шорцы) и пришлого восточнославянского (русские) населения Сибири проведено исследование частот полиморфных вариантов генов, продукты которых отвечают за транспорт витамина D – ген *GC* (Group-Specific Component), а также за

структурную его основного рецептора VDR (vitamin D receptor) – ген *VDR* и рецепторов-партнеров RXR (retinoid X receptor) – гены *RXRA* и *RXRG*.

## МАТЕРИАЛЫ И МЕТОДЫ

Материалом для исследования послужили образцы ДНК населения Сибири – русских ( $n = 100$ ) и коренных народов – телеутов ( $n = 98$ ), томских татар ( $n = 100$ ), шорцев ( $n = 113$ ). Обследование коренных групп населения осуществлено в ходе экспедиций в места их компактного расселения в административных границах Кемеровской (Беловский, Междуреченский, Таштагольский районы) и Томской (Томский район) областей в период с 2010 по 2023 гг. Русские представлены жителями г. Кемерово и Кемеровской области. Суммарный объем выборки составил 411 человек.

Исследование проведено под контролем комитета по этике и доказательности медицинских научных исследований ФГБОУ ВО КемГМУ Минздрава России (протокол № 285 от 13.04.2022 г.). Биологические материалы получены в сопровождении генеалогических анкет, демографических данных и информированных согласий на участие в исследовании. Выборки сформированы с учетом пола и этнической принадлежности (по самоопределению) согласно правилам, обоснованным в работе Е. В. Балановской и соавт. [22]. Потомков межнациональных браков из анализа исключали.

ДНК получали методом фенол-хлороформной экстракции. Генотипирование проводили с помощью ПЦР в режиме реального времени по панели 14 полиморфных вариантов генов, обеспечивающих транспорт (*GC*), связывание (*VDR*) и сигнaling (*RXRA*, *RXRG*) витамина D (табл. 1).

Выбор полиморфных вариантов генов, потенциально ассоциированных с уровнем сывороточного витамина D (25(OH)D) и витамина D в активной форме (1-25(OH)<sub>2</sub>D), осуществлен с учетом популяционной частоты минорного аллеля (MAF) (Ensembl, <http://www.ensembl.org>) и литературных данных о связи полиморфного варианта со статусом витамина D. Локализацию SNP в структуре гена указывали по данным NCBI,

**Таблица 1.** Характеристика исследованных генов системы витамина D

Ген	Наименование продукта гена	Хромосомная локализация	Экзоны, <i>n</i>	SNP, тыс.*
<i>GC</i>	GC vitamin D binding protein	4q13.3	15	25.1
<i>VDR</i>	vitamin D receptor	12q13.11	12	25.5
<i>RXRA</i>	retinoid X receptor alpha	9q34.2	12	51.3
<i>RXRG</i>	retinoid X receptor gamma	1q23.3	12	17.0

Примечание. \* – по данным NCBI (<http://www.ncbi.nlm.nih.gov/gene>).

<https://www.ncbi.nlm.nih.gov> (в качестве референсной последовательности использованы данные MANE Project v1.3), аннотацию – с помощью Ensembl VEP, <http://www.ensembl.org/info/docs/tools/ver/index.html>.

Расчет аллельных, гаплотипических частот, показателей гетерозиготности и равновесия Харди – Вайнберга осуществляли при помощи программы Genepop 4.7 (<https://genepop.curtin.edu.au/>) и веб-инструмента SNPstat (<https://www.snpstats.net/start.htm>). Различия частот считали статистически значимыми при уровне  $p < 0.05$ . Для сравнения привлекали общемировые частоты исследованной панели SNP, а также частоты в популяциях Европы и Восточной Азии. Источником информации послужили ALFA dataset (<https://ncbiinsights.ncbi.nlm.nih.gov/2020/03/26/alfa/>) и 1000Genomes (<https://www.internationalgenome.org/>).

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Анализ соответствия частот полиморфных вариантов генов распределению Харди–Вайнберга выявил в исследованных выборках ряд отклонений от состояния равновесия. У русских Сибири в одном случае – в отношении rs3847987 в гене *VDR*. Что касается коренных сибирских народов, то спектр таких SNP оказался несколько шире: от двух – у шорцев до четырёх – у телеутов. Отметим, что исследование всех SNP проведено на единых выборках, с использованием единых методов, реактивов и аппаратуры. Проверка воспроизводимости генотипирования (10% образцов, а в случае rs7041 – totally в выборках телеутов и томских татар) подтвердила результаты в 100% случаев. Анализ направленности отклонения показателей гетерозиготности выявил общую для всех популяций тенденцию к некоторому снижению уровня гетерозиготности, которое оказалось более выражено у телеутов (табл. 2). Известно, что общая численность этого народа на протяжении XIX–XXI вв. не превышала 2–2.5 тыс. человек [23, 24], а индекс эндогамии в местах компактного расселения в период с 1940-х по 2010-е гг. возрос с 58 до 89% [24]. Небольшая численность и высокая

частота однонациональных браков – основа роста инбридинга и снижения гетерозиготности в популяции телеутов. Кроме того, ранее проведенные нами исследования [24, 25] позволяют предположить, что зафиксированная в настоящем исследовании динамика популяционной структуры телеутов – это результат дрейфа генов вследствие генетико-демографических процессов, связанных с переходом на суженный тип воспроизводства, со снижением fertильности и рождаемости, а также с усилением брачной ассортативности по этническому признаку. Их совокупный эффект в исследованной выборке, по всей видимости, превысил влияние межэтнического смешения, так как, несмотря на принципы формирования выборок, исключающие из исследования потомков от смешанных браков, нельзя не принимать в расчет возможности наличия таковых за пределами трех поколений.

Исследование частот полиморфных вариантов сформированной панели генов *GC*, *VDR*, *RXRA* и *RXRG* продемонстрировало своеобразие генофонда коренных популяций Сибири по сравнению с русским населением, а также с общемировыми частотами (TOT – Total) и частотами, характерными для популяций Европы (EUR – Europa) и Восточной Азии (EAS – East Asian) (табл. 3–5).

Сравнение частоты полиморфных вариантов гена *GC* в выборках коренных народов с русским населением Сибири выявило статистически значимые отличия ( $p < 0.05$ ) в отношении частоты аллелей SNP rs7041 (шорцы). В целом в выборках шорцев и телеутов отмечена общая тенденция – меньшая распространенность аллелей, ассоциированных с пониженным по данным литературы (табл. 3) содержанием витамина D в сыворотке крови, в том числе из-за сниженного аффинитета продукта гена и следовательно способности к связыванию и транспорту данного витамина. В выборках томских татар и русских частоты аллелей гена *GC*, ассоциированные с потенциально худшим адаптационным потенциалом к условиям недостаточной обеспеченности витамином D, более распространены и несколько превышают

**Таблица 2.** Усредненные показатели гетерозиготности в исследованных выборках по данным полиморфных вариантов генов *GC*, *VDR*, *RXRA* и *RXRG*

Национальность	<i>No</i>	<i>He</i>	<i>D</i>
Русские	0.407	0.413	-0.017
Телеуты	0.363	0.409	-0.101
Томские татары	0.363	0.396	-0.090
Шорцы	0.382	0.404	-0.049

Примечание. *No* – наблюдаемая гетерозиготность, *He* – ожидаемая гетерозиготность, *D* – показатель отклонения гетерозиготности.

**Таблица 3.** Характеристика исследованных полиморфных вариантов гена *GC*

SNP	Аннотация SNP*	Выборки	$\chi^2_{x\text{-в}}$	Частота аллеля	$\chi^2$
rs7041 A>C экзон 12	Сниженное содержание 25(OH)D в крови [26]	Русские	0.40	(C) 0.572	
		Телеуты	<b>12.65</b>	0.463	2.27
		Татары	<b>20.35</b>	0.635	0.80
		Шорцы	2.39	0.389	<b>6.40</b>
		TOT/EUR/EAS		0.537/0.566/0.264	
rs4588 G>T экзон 12	Сниженное содержание 25(OH)D в крови [27]	Русские	0.11	(T) 0.317	
		Телеуты	0.01	0.252	0.99
		Татары	0.09	0.328	0.03
		Шорцы	0.12	0.265	0.62
		TOT/EUR/EAS		0.270/0.281/0.288	
rs3755967 C>T инtron 13	Сниженное содержание 25(OH)D в крови [28]	Русские	0.02	(T) 0.317	
		Телеуты	<b>3.99</b>	0.251	0.25
		Татары	0.56	0.278	0.35
		Шорцы	0.08	0.244	1.24
		TOT/EUR/EAS		0.170/0.218/0.260	

Примечание. Для табл. 3–5: \* – данные литературы,  $\chi^2_{x\text{-в}}$  – критерий оценки соответствия частот распределению Харди–Вайнберга,  $\chi^2$  – критерий для попарного сравнения частот. Полужирным шрифтом отмечены статистически значимые величины при сравнении коренных групп с выборкой русских; TOT – общемировые частоты, EUR – частоты популяций Европы, EAS – частоты популяций Восточной Азии.

среднемировые значения и показатели, характерные для популяций Европы и Восточной Азии (табл. 3).

В отношении генетических частот *VDR* значимые различия ( $p < 0.05$ ) выявлены для rs3847987 при сравнении выборки русских с выборками коренных народов Сибири. По данному полиморфному варианту, как и по всем другим, проанализированным в исследовании SNP *VDR*, выборка русских характеризовалась более низкими частотами альтернативных аллелей, в том числе по сравнению со среднемировыми и среднеевропейскими показателями (табл. 4). Для коренного населения Сибири в целом выявлены частоты, близкие (rs7975232, rs2228570, rs7968585) и превышающие (rs3847987) общемировые, европейские и восточноазиатские.

Исследованная панель полиморфных вариантов по генам ретиноидных рецепторов *RXRA* и *RXRG* оказалась более информативной (табл. 5). Статистически значимые различия получены при сравнении выборок коренных сибирских народов с русскими по rs9409929 и rs877954 гена *RXRA*, а также rs283696 гена *RXRG*. У телеутов чаще ( $p < 0.05$ ) фиксировался аллель C rs283696. Данный аллельный вариант оказался более распространенным во всех исследованных выборках сибирских народов. У телеутов и томских татар продемонстрирована сниженная ( $p < 0.05$ ) частота аллеля A rs9409929, для которого показан более высокий

уровень содержания активной формы витамина D (1.25(OH)<sub>2</sub>D) в крови. У томских татар также реже выявлялся аллель G rs877954, в отношении которого также есть данные литературы [5] об ассоциации с уровнем витамина D в сыворотке крови.

По данным о частотах аллелей методом Nei была построена матрица генетических расстояний, проиллюстрированная дендрограммой (рис. 1). Ожидаемо наиболее дистанционированной в общем генетическом пространстве оказалась выборка русских – усредненное значение  $d$ , характеризующее удаленность русских от всех исследованных популяций, составило 0.028. Минимальные генетические расстояния получены между телеутами – томскими татарами (0.008), а также шорцами – телеутами (0.009). Полученные результаты отражают этногенез, географическую локализацию этнических популяций и межэтнические контакты сибирских народов. Так, по данным истории и этнографии, в сложении томских татар приняли участие выездные телеуты, откочевавшие с князем Иркой Удековским на новые территории. А интенсивные межэтнические связи шорцев и телеутов определены близостью расположения мест компактного расселения.

Для получения данных о потенциале популяций в контексте генетической адаптации к недостаточной обеспеченности витамином D был проведен анализ частот гаплотипов полиморфных вариантов

Таблица 4. Характеристика исследованных полиморфных вариантов гена *VDR*

SNP	Аннотация SNP*	Выборки	$\chi^2_{\text{х-в}}$	Частота аллеля	$\chi^2$
rs1544410 (BsmI) C>T инtron 9	Сниженное содержание 25(OH)D в крови [29]	Русские	0.25	(T) 0.315	
		Телеуты	<b>4.36</b>	0.274	0.37
		Татары	1.29	0.208	2.84
		Шорцы	0.19	0.295	0.08
		TOT/EUR/EAS		0.388/0.398/0.064	
rs731236 (TaqI) A>G экзон 10	Сниженное содержание 25(OH)D в крови [30]	Русские	0.26	(G) 0.307	
		Телеуты	2.58	0.236	1.12
		Татары	<b>4.49</b>	0.234	1.26
		Шорцы	0.06	0.293	0.04
		TOT/EUR/EAS		0.384/0.398/0.067	
rs7975232 (ApaI) C>A инtron 9	Сниженное содержание 25(OH)D в крови [31]	Русские	0.89	(A) 0.390	
		Телеуты	2.17	0.511	2.67
		Татары	1.96	0.437	0.43
		Шорцы	2.92	0.469	1.09
		TOT/EUR/EAS		0.554/0.537/0.290	
rs2228570 (FokI) A>G экзон 3	Повышенное содержание 25(OH)D в крови [32]	Русские	0.40	(G) 0.572	
		Телеуты	3.39	0.623	0.49
		Татары	0.01	0.619	0.43
		Шорцы	0.21	0.543	0.15
		TOT/EUR/EAS		0.611/0.612/0.568	
rs3847987 C>A экзон 10	Ассоциации с уровнем триглицеридов [33]	Русские	<b>6.80</b>	(A) 0.116	
		Телеуты	1.27	0.268	<b>7.24</b>
		Татары	2.99	0.258	<b>6.39</b>
		Шорцы	<b>4.21</b>	0.217	<b>3.52</b>
		TOT/EUR/EAS		0.124/0.125/0.216	
rs7968585 C>T экзон 3	Сниженное содержание 25(OH)D в крови [34]	Русские	1.86	(T) 0.440	
		Телеуты	<b>4.83</b>	0.511	2.68
		Татары	2.14	0.573	0.43
		Шорцы	0.01	0.522	1.09
		TOT/EUR/EAS		0.520/0.519/0.275	

генов. Частоты наиболее распространенных вариантов “Топ-3” показаны на столбчатых гистограммах (рис. 2).

Значительный интерес в этом ключе представляется продукт гена *GC* – DBP (vitamin D binding protein, витамин D-связывающий белок), который несет уникальный сайт связывания для всех метаболитов витамина D. Комплекс DBP–25(OH)D образует

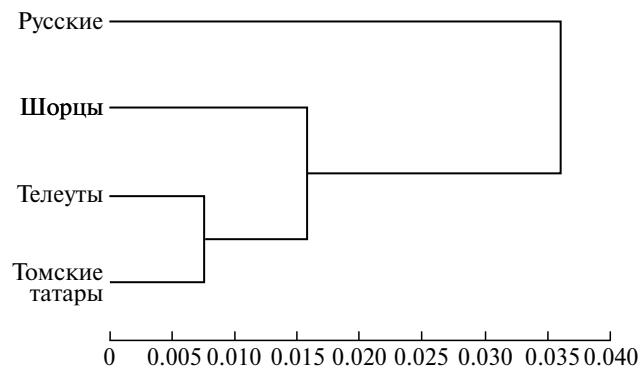
циркулирующий резервуар витамина, предотвращающий гиповитаминоз D, когда поступление нового витамина D недостаточно. Таким образом, DBP регулирует биодоступность всех метаболитов витамина D для тканей и клеток организма.

Ген *GC* содержит несколько функциональных полиморфных вариантов С доказанным по данным литературы [26,27] влиянием D в сыворотке

**Таблица 5.** Характеристика исследованных полиморфных вариантов генов *RXR*

SNP	Аннотация SNP*	Выборки	$\chi^2_{\text{х-в}}$	Частота аллеля	$\chi^2$
<i>RXRA</i> rs9409929 $G > A$ инtron 1	Повышенный уровень 1,25(OH) <sub>2</sub> D в крови [35]	русские	0.04	(A) 0.430	
		телеуты	0.38	0.250	<b>6.86</b>
		татары	0.74	0.276	<b>4.93</b>
		шорцы	0.67	0.398	0.14
		TOT/EUR/EAS		0.340/0.353/0.217	
rs3132299 $C > G$ инtron 2	Ассоциации с уровнем витамина D в крови [5]	русские	0.08	(G) 0.842	
		телеуты	0.40	0.794	0.74
		татары	1.94	0.822	0.14
		шорцы	2.28	0.801	0.54
		TOT/EUR/EAS		0.843/0.825/0.806	
rs877954 $A > G$ инtron 9	Ассоциации с уровнем витамина D в крови [5]	русские	0.36	(G) 0.697	
		телеуты	1.77	0.578	2.93
		татары	0.29	0.526	<b>5.88</b>
		шорцы	0.03	0.734	0.32
		TOT/EUR/EAS		0.633/0.658/0.662	
<i>RXRG</i> rs2651860 $A > C$ инtron 4	Ассоциация с дислипидемией [36]	русские	1.73	(C) 0.182	
		телеуты	0.43	0.247	1.17
		татары	2.07	0.197	0.07
		шорцы	0.26	0.226	0.56
		TOT/EUR/EAS		0.149/0.205/0.101	
rs283696 $T > C$ инtron 9	Ассоциация с дислипидемией [36]	русские	0.010	(C) 0.752	
		телеуты	0.23	0.873	<b>4.58</b>
		татары	0.36	0.821	1.37
		шорцы	3.07	0.857	3.34
		TOT/EUR/EAS		0.872/0.796/0.927	

крови. Наиболее важными и изученными являются два – rs7041 и rs4588. Первый – rs7041 (с. 1296A > C) кодирует глутаминовую кислоту вместо аспарагиновой в положении 432 (p.Asp432Glu), второй – rs4588 (с.1307G > T) – лизин вместо треонина в положении 436 (p.Thr436Lys). Сочетание этих полиморфных вариантов приводит к трем основным и наиболее распространенным изоформам гена *GC*: *GC1s* (rs7041C–rs4588G, кодирующим 432Glu/436Thr); *GC1f* (rs7041A–rs4588G, кодирующим 432Asp/436Thr); *GC2* (rs7041A–rs4588T, кодирующим 432Asp/436Lys). Эти три изоформы генерируют шесть различных фенотипов (*GC1f-GC1f*, *GC1f-GC1s*, *GC1s-Gc1s*, *GC1f-GC2*, *GC1s-GC2* и *GC2-GC2*) и модулируют концентрацию циркулирующего DBP и следовательно биодоступность витамина D. Еще одним SNP стал инtronный



**Рис. 1.** Положение генофондов исследованных народов в общем генетическом пространстве по данным панели 14 SNP генов системы витамина D (метод Уорда).

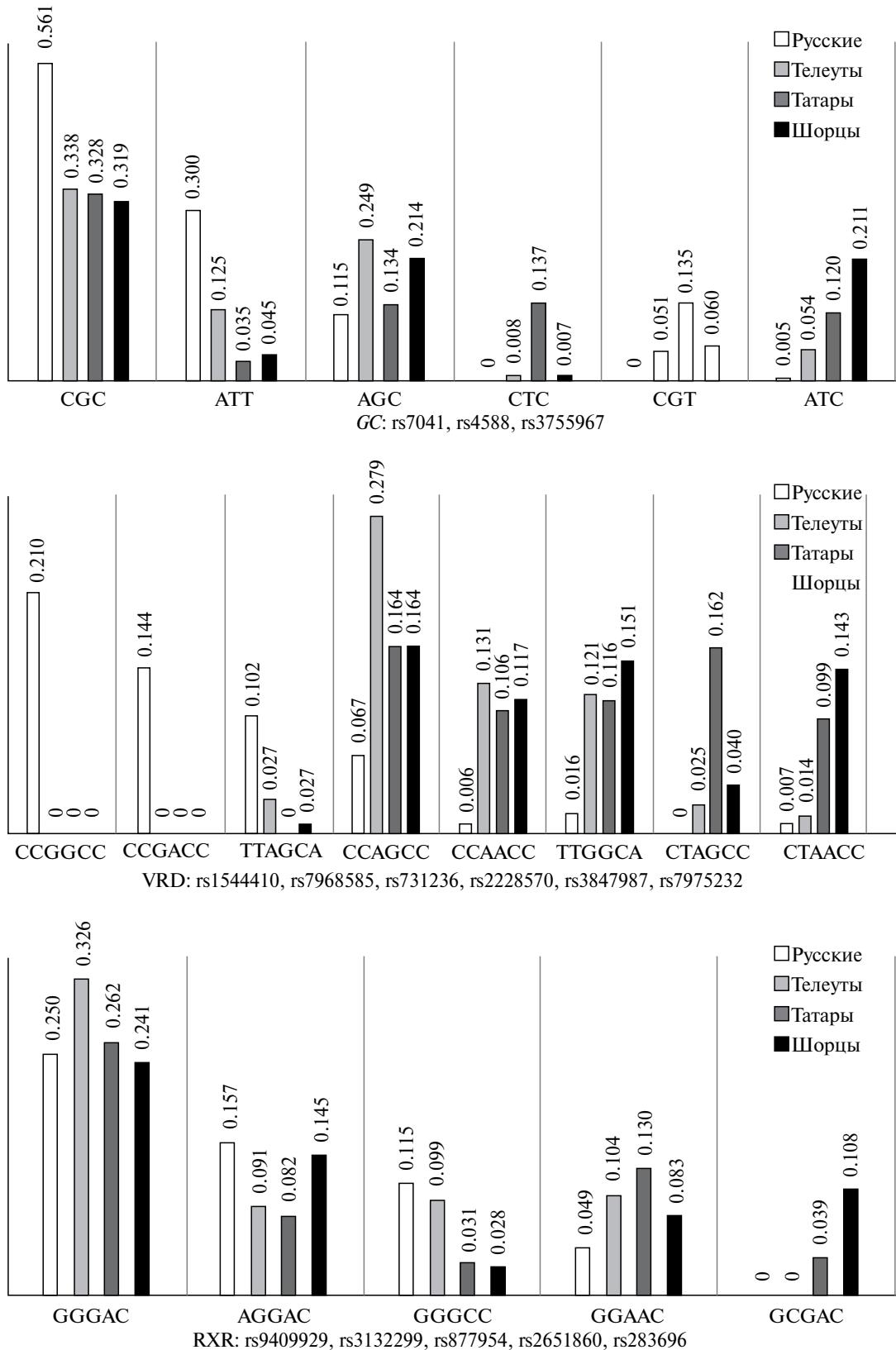


Рис. 2. Частоты распространенных гаплотипов полиморфных вариантов исследованного комплекса генов в выборках коренных народов и русских Сибири.

вариант rs3755967, для которого также установлена ассоциация со сниженным уровнем витамина D [28].

Анализ составов гаплотипов исследованных полиморфных вариантов гена *GC* свидетельствует о генетической адаптации популяций коренных сибирских народов к условиям недостаточной обеспеченности витамином D. В их генофондах отмечено накопление вариантов гена, кодирующего белок с лучшей способностью к связыванию и удержанию данного витамина. Отметим, что во всех проанализированных выборках самым распространенным оказался *GC1s* в составе гаплотипа CGC, однако в выборке русских вторым по частоте был гаплотип ATT (*GC2* и мажорный вариант – аллель *T* по rs3755967). Для белка фенотипа *GC2-2* установлена сниженная аффинность к 25(OH)D. В группах коренных народов Сибири частота такого варианта ниже (в 2–10 раз), а гаплотип AGC, включающий три исходных “предковых” варианта гена, напротив, занимает в “Топ-3” вторую или третью позицию.

В отношении гена *VDR* у русских и коренных народов Сибири выявлен совершенно разный гаплотипический профиль. Продукт гена *VDR* – рецептор витамина D – экспрессируется в различных клетках и тканях с максимальным уровнем в тонкой, двенадцатиперстной и толстой кишке, почках и коже. В клетке он концентрируется в ядре, цитозоле и на цитоплазматической мембране, реализуя регуляцию на геномном и негеномном уровнях. Белок VDR специфически связывает активную форму витамина D (1,25(OH)<sub>2</sub>D) и опосредует ее действие – активирует экспрессию широкого спектра генов-мишеней с элементом VDRE в промоторе, в том числе генов иммунного ответа и генов, поддерживающих клеточные процессы роста и дифференцировки.

Распространенные в выборке русских гаплотипы гена *VDR* – CCGGCC, CCGACC и TTAGCA – в исследованных нами выборках коренных народов не зарегистрированы или крайне редки. В составе первых двух гаплотипов содержится аллель *G* rs7968585, ассоциированный по данным литературы со сниженным уровнем витамина D (25(OH)D), и/или *G* rs2228570 (повышенное содержание 25(OH)D в сыворотке крови). Для третьего гаплотипа можно предположить негативный эффект – три полиморфных варианта потенциально ассоциированы со сниженным содержанием сывороточного витамина D (25(OH)D). В генофондах сибирских татар, телеутов и шорцев самым частым оказался гаплотип CCAGCC, включающий варианты “нормы” по исследованному комплексу SNP и один мажорный вариант – rs2228570, кодирующий укороченную форму рецептора FokI VDRA. Отметим, что существуют три изоформы рецептора витамина D, различающиеся по длине: VDRA

(427 а.о.), VDRB1 (477 а.о.) и FokI VDRA (424 а.о.). Для последней – укороченной – формы, детерминируемой на генетическом уровне SNP rs2228570, характерен высокий уровень трансляции, что предполагает повышенную чувствительность тканей и клеток-мишеней к лиганду – витамину D – в условиях его недостаточного поступления.

Что касается сочетаний исследованных полиморфных вариантов генов *RXRA* и *RXRG*, то их распределение в исследованных выборках населения Сибири оказалось более однородным, несмотря на выявленные статистически значимые различия при попарном сравнении частот отдельных SNP. Вероятно, это связано с полифункциональностью ретиноидных рецепторов, которые способны образовывать гетеродимеры с широким спектром рецепторов-партнеров (VDR, PPAR, FXR, LXR, TR), модифицируя их сигналинг, обеспечивающий экспрессию генов и метаболизм веществ в клетке. Самым частым вариантом во всех выборках без исключения стал GGGAC. В составе этого комплекса присутствуют два потенциально “негативных” SNP, ассоциированных со сниженным уровнем витамина D – rs3132299, rs877954, и два варианта “нормы” по риску дислипидемии – rs2651860, rs283696. В выборке шорцев вторым по распространенности стал гаплотип AGGAC, содержащий полиморфный вариант rs9409929 (аллель *A*, по некоторым данным, ассоциирован с повышенным уровнем сывороточного витамина D).

Исследование структуры генофондов коренного и пришлого населения Сибири по данным 14 полиморфных вариантов генов системы витамина D позволяет сделать заключение о том, что структура генофондов коренных народов Сибири сохраняет своеобразие, отражающее историю сложения и демонстрирующее черты генетической адаптации коренных популяций к условиям среды обитания.

Данное исследование выполнено за счет средств гранта Российского научного фонда № 22-25-20209, <https://rscf.ru/project/22-25-20209> и Министерства науки и высшего образования Кузбасса.

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. Все обследованные – совершеннолетние.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. IUPAC-IUB joint commission on biochemical nomenclature (jcbn). Nomenclature of vitamin D. Recommendations 1981 // Mol. Cell. Biochem. 1982. V. 49. № 3. P. 177–181. doi: 10.1007/BF00231180
2. Громова О.А., Торшин И.Ю., Гильельс А.В. и др. Метаболиты витамина D: роль в диагностике и терапии витамин-D-зависимых патологий // Фармакокинетика и фармакодинамика. 2016. № 4. С. 9–18.
3. Бочкинникова А.Г., Тришкин А.Г., Пустотина О.А., Мозес В. “D-статус” пациенток с репродуктивными нарушениями // Фундаментальная и клиническая медицина. 2018. Т. 3. № 3. С. 6–11.
4. Christakos S., Dhawan P., Verstuyf A. et al. Vitamin D: Metabolism, molecular mechanism of action, and pleiotropic effects // Physiol. Rev. 2016. V. 96. № 1. P. 365–408. doi: 10.1152/physrev.00014.2015
5. Voltan G., Cannito M., Ferrarese M. et al. Vitamin D: An overview of gene regulation, ranging from metabolism to genomic effects // Genes (Basel). 2023. V. 14. № 9. doi: 10.3390/genes14091691
6. Водолазская А.Н., Орлова С.В., Батышева Т.Т. и др. Факторы, влияющие на биодоступность витамина D // Мед. алфавит. 2023. № 29. С. 55–60. doi: 10.33667/2078-5631-2023-29-55-60
7. Козлов А.И., Вершубская Г.Г. 25-гидроксивитамин D в различных группах населения Севера России // Физиология человека. 2019. Т. 45. № 5. С. 125–136. doi: 10.1134/S0131164619050060
8. Luthold R.V., Fernandes G.R., Franco-de-Moraes A.C. et al. Gut microbiota interactions with the immunomodulatory role of vitamin D in normal individuals // Metabolism. 2017. № 69. P. 76–86. doi: 10.1016/j.metabol.2017.01.007
9. Bakke D., Sun J. Ancient nuclear receptor VDR with new functions: microbiome and inflammation // Inflamm. Bowel. Dis. 2018. V. 24. № 6. P. 1149–1154. doi: 10.1093/ibd/izy092
10. Boughanem H., Ruiz-Limón P., Pilo J. et al. Linking serum vitamin D levels with gut microbiota after 1-year lifestyle intervention with mediterranean diet in patients with obesity and metabolic syndrome: A nested cross-sectional and prospective study // Gut Microbes. 2023. V. 15. № 2. doi: 10.1080/19490976.2023.2249150
11. Козлов А.И., Вершубская Г.Г., Лавряшина М.Б., Остроухова И.О. Отражение особенностей традиционного питания в генофондах народов с лесо-таёжным типом природопользования // Вестник МГУ. Серия 23: Антропология. 2020. № 3. С. 46–56. doi: 10.32521/2074-8132.2020.3.046-056
12. Manousaki D., Mitchell R., Dudding T. et al. Genome-wide association study for vitamin D levels reveals 69 independent loci // Am. J. Hum. Genet. 2020. V. 106. № 3. P. 327–337. doi: 10.1016/j.ajhg.2020.01.017
13. Вильмс Е.А., Добровольская Е.В., Турчанинов Д.В. и др. Обеспеченность взрослого населения Западной Сибири витамином D: данные популяционного исследования // Вопросы питания. 2019. Т. 88. № 4. С. 75–82. doi: 10.24411/0042-8833-2019-10044
14. Петрушкина А.А., Пигарова Е.А., Рожинская Л.Я. Эпидемиология дефицита витамина D в Российской Федерации // Остеопороз и остеопатии. 2018. Т. 21. № 3. С. 15–20. doi: 10.14341/osteo10038
15. Малярчук Б.А. Полиморфизм гена рецептора витамина D у коренного населения Сибири // Вестник Северо-Восточного науч. центра ДВО РАН. 2020. № 3. С. 120–127. doi: 10.34078/1814-0998-2020-3-120-127
16. Малярчук Б.А., Деренко М.В., Денисова Г.А. Адаптивные изменения генов десатурации жирных кислот у коренного населения Северо-Востока Сибири // Генетика. 2021. Т. 57. № 12. С. 1458–1464. (Malyarchuk B.A., Derenko M.V., Denisova G.A. Adaptive changes in fatty acid desaturation genes in indigenous populations of Northeast Siberia // Rus. J. Genetics. 2021. V. 57. № 12. P. 1461–1466.) doi: 10.31857/S0016675821120109
17. Козлов А.И., Пылев В.Ю., Вершубская Г.Г., Балановская Е.В. Клинальная изменчивость генетических детерминант трегалазной недостаточности в популяциях Южной Сибири, Казахстана, Центральной Азии и Монголии // Вестник МГУ. Серия 23: Антропология. 2023. № 3. С. 63–71. doi: 10.32521/2074-8132.2023.3.063-071
18. Малярчук Б.А., Деренко М.В. Оценка роли отбора в эволюции митохондриальных геномов коренного населения Сибири // Вавил. журнал генетики и селекции. 2023. Т. 27. № 3. С. 218–223. doi: 10.18699/VJGB-23-28
19. Козлов А.И., Малярчук Б.А., Лавряшина М.Б., Вершубская Г.Г. Нарушения усвоения сахарозы подтверждают своеобразие генетической истории эскимосов // Вестник МГУ. Серия 23: Антропология. 2023. № 2. С. 82–91. doi: 10.32521/2074-8132.2023.2.082-091
20. Доскина Е.В. Роль различных форм витамина D в лечении пациентов с дефицитом витамина D (клинический случай) // Эндокринология. Новости. Мнения. Обучение. 2021. Т. 10. № 2(35). С. 123–129. doi: 10.33029/2304-9529-2021-10-2-123-129
21. Carlberg C., Raczyk M., Zawrotna N. Vitamin D: A master example of nutrigenomics // Redox. Biol. 2023. V. 62. P. 102695. doi: 10.1016/j.redox.2023.102695

22. Балановская Е.В., Жабагин М.К., Агджоян А.Т. и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине // Генетика. 2016. Т. 52. № 12. С. 1371–1387. (Balanovska E.V., Zhabagin M.K. Agdzhoyan A.T. et al. Population biobanks: organizational models and prospects of application in gene geography and personalized medicine // Rus. J. Genetics. 2016. V. 52. № 12. P. 1227–1243.) doi: 10.7868/S001667581612002X
23. Функ Д.А. Бачатские телеуты в XVIII – первой четверти XX века: историко-этнографическое исследование. М.: Ин-т этнологии и антропологии им. Н.Н. Миклухо-Маклая РАН, 1993. 325 с.
24. Лавряшина М.Б., Ульянова М.В., Поддубиков В.В. и др. Мониторинг сельских популяций шорцев и телеутов: воспроизводство, среда, гены // Генетика человека и патология: Сб. научных трудов. 2017. №. 11. С. 49–51.
25. Лузина Ф.А., Колбаско А.В., Лотош О.Е. Демографическая структура и её динамика у телеутов // Бюл. СО РАМН. 2006. № 3. С. 60–63.
26. Hibler E.A., Hu C., Jurutka P.W. et al. Polymorphic variation in the GC and CASR genes and associations with vitamin D metabolite concentration and meta-chronous colorectal neoplasia // Cancer Epidemiol. Biomarkers Prev. 2012. V. 21. № 2. P. 368–375. doi: 10.1158/1055-9965.EPI-11-0916
27. Abbas S., Linseisen J., Slanger T. et al. The GC2 allele of the vitamin D binding protein is associated with a decreased postmenopausal breast cancer risk, independent of the vitamin D status // Cancer Epidemiol. Biomarkers Prev. 2008. V. 17. № 6. P. 1339–1343. doi: 10.1158/1055-9965.EPI-08-0162
28. Gibbs D.C., Fedirko V., Um C. et al. Associations of circulating 25-hydroxyvitamin D3 concentrations with incident, sporadic colorectal adenoma risk according to common vitamin D-binding protein isoforms // Am. J. Epidemiol. 2018. V. 187. № 9. P. 1923–1930. doi: 10.1093/aje/kwy102
29. Ashok N., Saraswathy R. Association of polymorphisms of vitamin D gene in children with asthma and allergic rhinitis – Hospital based study // Heliyon. 2023. V. 10. № 1. doi: 10.1016/j.heliyon. 2023.e23673
30. Ferraz R.S., Silva C.S., Cavalcante G.C. et al. Variants in the VDR gene may influence 25(OH)D levels in type 1 diabetes mellitus in a brazilian population // Nutrients. 2022. V. 14. № 5. doi: 10.3390/nu14051010
31. Zhao X.Q., Wan H.Y., He S.Y. et al. Vitamin D receptor genetic polymorphisms associate with a decreased susceptibility to extremity osteomyelitis partly by inhibiting macrophage apoptosis through inhibition of excessive ROS production via VDR-Bmi1 signaling // Front. Physiol. 2022. № 13. doi: 10.3389/fphys.2022.808272
32. Jia J., Tang Y., Shen C. et al. Vitamin D receptor polymorphism rs2228570 is significantly associated with risk of dyslipidemia and serum LDL levels in Chinese Han population // Lipids Health Dis. 2018. V. 17. № 1. P. 193. doi: 10.1186/s12944-018-0819-0
33. Zhang D., Cheng C., Wang Y. et al. The influence of VDR polymorphisms on the type 2 diabetes susceptibility in Chinese: An interaction with hypertriglyceridemia // Mol. Genet. Genomics. 2021. V. 296. № 4. P. 837–844. doi: 10.1007/s00438-021-01784-z
34. Levin G.P., Robinson-Cohen C., de Boer I.H. et al. Genetic variants and associations of 25-hydroxyvitamin D concentrations with major clinical outcomes // Jama. 2012. V. 308. № 18. P. 1898–1905. doi: 10.1001/jama.2012.17304
35. Hibler E.A., Jurutka P.W., Egan J.B. et al. Association between polymorphic variation in VDR and RXRA and circulating levels of vitamin D metabolites // The J. Steroid Biochemistry and Mol. Biology. 2010. V. 121. № 1–2. P. 438–441. doi: 10.1016/j.jsbmb.2010.03.05
36. Sentinelli F., Minicocci I., Montali A. et al. Association of RXR-gamma gene variants with familial combined hyperlipidemia: genotype and haplotype analysis // J. Lipids. 2013. № 2013. doi: 10.1155/2013/51794

## Population and Genetic Structure of the Siberian Population According to Data on the Frequencies of Polymorphic Variants of Vitamin D Transport and Reception Genes

B. A. Tkhorenko<sup>1,\*</sup>, A. V. Meyer<sup>1</sup>, G. V. Vavin<sup>1</sup>, D. O. Imekina<sup>1</sup>,  
M. V. Ulyanova<sup>1</sup>, F. A. Luzina<sup>2</sup>, M. B. Lavryashina<sup>1</sup>

<sup>1</sup>Kemerovo State Medical University of the Ministry of Health of the Russian Federation, Kemerovo, 650056 Russia

<sup>2</sup>Research Institute for Complex Problems of Hygiene and Occupational Diseases, Novokuznetsk, 654041 Russia

\*e-mail: tba2008@mail.ru

The gene pools of Siberian peoples are a potential source of data on the compositions of haplotypes of polymorphic gene variants formed during the evolution of populations under the influence of a specific set of natural and climatic factors, in conditions of relative geographical and cultural isolation, as well as traditional subsistence practices. In order to study genetic adaptation to insufficient supply of vitamin D (low level of insolation and availability of nutrients containing cholecalciferol and ergocalciferol) in samples of indigenous Turkic-speaking (Teleuts, Tomsk Tatars, Shors) and foreign East Slavic (Russians) population of Siberia, a study of the frequencies of polymorphic variants (14 SNPs) of genes responsible for transport of vitamin D – *GC* (Group-Specific Component), the structure of its main receptor – gene *VDR* (vitamin D receptor) and the RXR receptor (retinoid X receptor) in form of genes *RXRA* and *RXRG* was conducted. The total sample size was 411 people. The analysis of polymorphic variant frequencies demonstrated specific features of the gene pools of Siberian indigenous populations compared to the Russian population, global frequencies and frequencies characteristic of populations of Europe and East Asia. In pairwise comparison with the sample of Russians, statistically significant differences were found in terms of frequencies rs7041 (Shors), rs3847987 (Tomsk Tatars, Teleuts, Shors), rs9409929 (Tomsk Tatars, Teleuts), rs877954 (Tomsk Tatars), rs283696 (Teleuts). According to data of average genetic distances (*d*), the minimum distances were noted between Teleuts and Tomsk Tatars. The sample of Russians in general turned out to be the most distant in the genetic space of the studied peoples of Siberia. The study of haplotype frequencies of polymorphic variants of vitamin D transport and receptor genes demonstrated the accumulation of SNP in the indigenous gene pool associated, according to the literature, with a higher levels of serum vitamin D. According to the *GC* and *VDR* genes, the haplotypic profile of polymorphic variants was characterized by pronounced specificity and differed in composition from that of the Russian population. The study allows us to conclude that structure of gene pools of indigenous peoples of Siberia retains a distinctiveness that reflects the history of their composition and demonstrates the features of the genetic adaptation of indigenous populations to environmental conditions.

**Keywords:** vitamin D, genes, indigenous peoples, polymorphism, population, receptors, Russians, Siberia.

## О СТАБИЛИЗАЦИИ ТЕМПА ДИВЕРГЕНЦИИ ИЗОНИМИИ

© 2024 г. В. П. Пасеков<sup>1,\*</sup>

<sup>1</sup>Федеральный исследовательский центр “Информатика и управление” Российской академии наук,  
Москва, 119991 Россия  
<sup>\*</sup>e-mail: pass40@mail.ru

Поступила в редакцию 19.06.2024 г.

После доработки 22.07.2024 г.

Принята к публикации 24.07.2024 г.

Проведен теоретический анализ фамильного состояния популяции (вектора концентраций однофамильцев в мужском компоненте популяции) и его динамики в результате случайного фамильного дрейфа. Используется аппроксимация такого процесса моделью Райта – Фишера популяции с неперекрывающимися поколениями, неподверженной давлению отбора, т. е. последовательностью вложенных случайных выборок с возвращением из совокупности фамилий отцов. Размер выборки равен  $N/2$  согласно численности мужского компонента в популяции размера  $N$ . В одной и той же популяции одновременно протекают процессы случайного дрейфа как фамилий, так и генов. Их кардинальное различие в том, что размер выборки фамилий вчетверо меньше, чем выборки аллелей аутосомного локуса. Анализ случайного дрейфа упрощается при переходе от координат-концентраций к квадратным корням из них. При смене поколений состояние получает выборочное отклонение, измеряемое угловым расстоянием, а его средний квадрат дает темп дивергенции, стабилизирующийся в новых координатах. Данна адаптация (применительно к анализу фамильного дрейфа) известного в популяционной генетике результата о характере дивергенции на этапе относительно малого по сравнению с размером популяции количества поколений. Дивергенция фамилий протекает в 4 раза быстрее дивергенции концентраций аллелей.

**Ключевые слова:** случайный фамильный дрейф, дивергенция концентраций фамилий и аллелей, изонимия, угловые расстояния, стабилизация темпа дивергенции.

**DOI:** 10.31857/S00166758241200103 **EDN:** VZVXWJ

Изучение фамильной структуры популяций человека интересно не только само по себе, но и как отражение действующих на уровне популяции процессов, как отражение происхождения популяций и как косвенное свидетельство характера генетической структуры (см. [1], где имеется обширная библиография). Дело не только в том, что фамилии могут наследоваться патрилинейно и передаваться сходно с генами негомологичного участка Y-хромосомы, но и в том, что характер типичных популяционных процессов (миграция, изоляция, популяционные волны численности и др.) близким образом влияет на распределение генов и на распределение фамилий. Сходство в передаче потомкам фамилии и генов позволяет использовать фамильные данные при изучении структуры ДНК Y-хромосомы (см. обзор [2, 3]) и в ряде случаев сузить круг фамилий подозреваемых в криминалистике. К настоящему времени проведены широкие исследования фамильной структуры во многих странах и их внутренних регионах, в том числе в России (см., например, [4 с картографическим анализом, 5]). Количество соответствующих

работ перевалило за половину тысячи, и обзор современного состояния данной области заслуживает отдельной публикации, а здесь мы ограничились ссылками преимущественно на монографии, но упомянем посвященную библиографию работу [6], в которой источники сгруппированы по изучаемым странам.

Отметим, что с термином изонимия связаны оставшиеся за рамками настоящей статьи популярные подходы, основанные на использовании данных по частоте браков между однофамильцами для оценивания коэффициента инбридинга в популяции [7–9] (см. критические замечания в [10]). Мы не рассматриваем используемые в публикациях такие характеристики фамильной структуры популяции, как индекс случайной изонимии и показатели разнообразия фамилий. При анализе распределения фамилий наш фокус лежит на других подходах и методах, применяемых в популяционной генетике. Конечно, при этом требуются определенные коррекции в методах исследования и в интерпретации результатов. Цель настоящей

работы состоит в адаптации методов популяционно-генетического анализа применительно к изучению фамильной структуры и ее связи с генетической структурой, а также обоснования теоретического фундамента таких методов.

Для достижения указанной цели используем упрощенную модель случайного фамильного дрейфа в популяции с неперекрывающимися поколениями [11–13]. Согласно закономерностям репродукции при оплодотворении зигота получает случайным образом один из двух аллелей аутосомного локуса отца и один от матери, т. е. генотип потомка представляет собой случайную выборку аллеля от отца и аллеля от матери. На популяционном уровне при случайном комбинировании генотипов родителей при неперекрывающихся поколениях генетический состав популяции потомков является результатом случайного выбора аллелей из родительской популяции. Аналогично фамильный состав потомков формируется как случайная выборка фамилий из мужской составляющей родительской популяции. Данная модель в популяционной генетике известна как модель Райта – Фишера. В ряду неперекрывающихся поколений мы получаем последовательность вложенных выборок. Динамику фамильного состава (изонимии), изменяющегося в результате выборочных ошибок при “копировании” родительского состава, назовем по аналогии с генным дрейфом процессом фамильного дрейфа.

Использование предположения о неперекрывающихся поколениях, когда речь идет о популяциях человека, проблематично, так как входит в противоречие с реальным положением вещей. Однако допустить такое использование можно на основе многочисленных результатов изучения с его помощью разнообразных реальных популяций. Достаточно вспомнить проверку закона Харди–Вайнберга, полученного в своей классической форме для популяций с неперекрывающимися поколениями. Отметим также, что многие выводы при изучении разнообразных природных популяций получены с помощью приложения результатов непрерывной аппроксимации для дискретных моделей популяций с неперекрывающимися поколениями.

Другая проблема использования модели случайного дрейфа связана с тем, что реальные популяции подвержены одновременному давлению нескольких факторов микрэволюции. Тем не менее такое использование оправдано, так как при сравнении по результатам давления систематических факторов и случайного дрейфа последний доминирует на относительно небольших промежутках времени [14, 15]. В данном контексте фамильное состояние очередной популяции потомков моделируется как результат случайной выборки фамилий из их совокупности в мужском компоненте родительской популяции.

Дальнейшее изложение придерживается следующего плана. Сначала формулируются основные понятия, используемые при изучении фамильной структуры. Затем обсуждается переход от традиционных фамильных состояний популяции в терминах концентраций однофамильцев к состояниям с координатами в виде квадратных корней из концентраций. Далее обосновывается аппроксимация распределения фамильных состояний популяции, описываемых угловым отклонением  $\theta$  от начального состояния, нормальным распределением. Преимущество новых координат состоит в достижении независимости от состояния популяции эффектов случайного дрейфа.

Кратко коснемся обозначений. Названия векторов и матриц набраны полужирным шрифтом (заглавными буквами для матриц, матрица с элементами  $a_{ij}$  обозначается как  $[a_{ij}]$ ). К обозначениям фамильных аналогов популяционно-генетических характеристик добавлено окончание  $s$  (для дисперсий  $I_s$  и углов  $\Theta_s$  соответственно). Символ  $E$  относится к операции получения среднего значения (*математического ожидания*). Когда у  $E$  имеется нижний индекс, то подразумевается, что усреднение производится по переменной, обозначаемой этим индексом. Расстояние между точками  $x$  и  $y$  в Евклидовом пространстве обозначаем как  $|x - y|$ . Знак тождества “=” используется в смысле равенства по определению. Символ  $\blacktriangleleft$  отмечает конец доказательства.

## ОСОБЕННОСТИ ВЫБОРОЧНОГО ДРЕЙФА ФАМИЛИЙ В ОДНОЙ ИЗОЛИРОВАННОЙ ПОПУЛЯЦИИ

*Фамильное состояние популяции* определяется как набор (вектор) концентраций групп мужчин-однофамильцев, короче концентраций фамилий в популяции. В модели процесса случайного дрейфа фамилий последовательность фамильных состояний по неперекрывающимся поколениям представляет собой цепь результатов вложенных случайных выборок с возвращением из фамилий мужских компонентов соответствующих родительских популяций. Вероятность появления определенной фамилии при извлечении выборочной единицы (у нас сына) равна концентрации этой фамилии среди родителей (среди глав семей). Формально каждая выборка рассматривается как мужская составляющая популяции в очередном поколении, а последовательность выборок в этой схеме определяет динамику изонимии в ряду неперекрывающихся поколений. Распределение состава выборки (распределение возможного фамильного состава популяции в следующем поколении, т. е. концентраций фамилий) является полиномиальным (мультиномиальным).

Любое состояние популяции как совокупности, состоящей из групп однотипных объектов (у нас групп однофамильцев), можно геометрически представить в Евклидовом пространстве как точку (вектор из начала координат)  $\mathbf{x}$  на части гиперплоскости над полуосями неотрицательных координат (см. рис. 1), координаты точки  $\mathbf{x}$  равны концентрациям групп  $\{x_i\}$ . Эта гиперплоскость отсекает единичные отрезки на осях координат и состоит из множества точек  $\mathbf{x}$  таких, что в случае  $k$  групп

$$\mathbf{x} = (x_1, x_2, \dots, x_k)^T, \quad x_i \geq 0, \quad \sum_{i=1}^k x_i = (\mathbf{x}, \mathbf{e}) = 1, \quad \mathbf{e} \equiv (1, 1, \dots, 1)^T.$$

Здесь  $T$  – символ транспонирования,  $(\mathbf{x}, \mathbf{e})$  – скалярное произведение вектора-состояния  $\mathbf{x}$  и вектора нормали  $\mathbf{e}$  к рассматриваемой плоскости  $(\mathbf{x} - \mathbf{p}, \mathbf{e}) = 0$ , к плоскости отклонений  $\mathbf{x}$  от (начального) состояния  $\mathbf{p}$ . На границе множества состояний (фазового пространства) концентрация одной из групп равна нулю. Выборочное отклонение  $\mathbf{x}$  от  $\mathbf{p}$  можно охарактеризовать квадратом Евклидова расстояния  $|\mathbf{x} - \mathbf{p}|$  между  $\mathbf{x}$  и  $\mathbf{p}$ .

$$|\mathbf{x} - \mathbf{p}|^2 \equiv (\mathbf{x} - \mathbf{p}, \mathbf{x} - \mathbf{p}) = \sum_{i=1}^k (x_i - p_i)^2.$$

*Задача настоящей работы состоит в упрощении анализа динамики фамильного и генетического состояний популяции с неперекрывающимися поколениями, изменяющихся в результате случайного дрейфа. Анализ как бы обращает нас к модели случайного генного дрейфа с дискретным временем, для которой большинство результатов выведены с использованием аппроксимации непрерывными аналогами. Получаемые выводы могут использоваться при оценивании инбридинга.*

Повторим, что процесс случайного фамильного дрейфа популяции с неперекрывающимися поколениями является последовательностью выборочных изменений фамильного состояния при смене поколений (последовательностью вложенных случайных выборок с возвращением). Состояние популяции с  $k$  вариантами фамилий в следующем (первом) поколении представляет собой результат случайной выборки с возвращением фамилий из множества фамилий мужчин родительской популяции. Размер выборки фамилий равен  $N(1)/2$ , где  $N(1)$  – численность диплоидной популяции в первом поколении, а  $N(1)/2$  – численность ее мужского компонента, передающего свои фамилии по поколениям.

Выборка является случайной при независимых выборах фамилии для каждого потомка. Повторим, что ее размер равен  $N(1)/2$ . Хотя мы далее интерпретируем  $N$  как размер популяции, ключевым является размер выборки мужчин, носителей наследуемых фамилий (с учетом дополнительных поправок

его можно назвать эффективным дисперсионным размером мужского компонента, в нашем случае он взят для простоты равным  $N/2$ ). Интерпретация  $N$  как общего размера популяции условна, и  $N$  фактически играет роль параметра. Чем меньше размер  $N/2$  мужского компонента популяции, тем более интенсивны выборочные отклонения нового фамильного состояния от прежнего, а величина разброса выборочных колебаний определяет “темпер” дивергенции фамильных состояний от начального.

В первом поколении вероятность попадания в выборку  $i$ -й фамилии равна  $p_i$  при каждом из  $N(1)/2$  испытаний (при каждом выборе фамилии для потомка). Вероятности  $\{p_i\}$  равны концентрациям фамилий в начальном фамильном состоянии популяции  $\mathbf{p}$ . При описанной схеме распределение результатов выборки является полиномиальным, как говорилось выше.

Пусть вектор  $\mathbf{x} = \mathbf{x}(1)$  с концентрациями фамилий  $\{x_i\}$  обозначает состояние популяции в первом поколении. Если рассматривать только какую-либо одну из координат вектора-состояния  $\mathbf{x}$  (концентрацию отдельной, скажем,  $i$ -й фамилии), то вероятность попадания этой фамилии в выборку (“успеха”) в результате одного из  $N(1)/2$  испытаний при формировании первого поколения равна  $p_i$ . Концентрация  $i$ -й фамилии  $x_i$  в выборке (в следующем поколении) является результатом деления количества успехов на размер выборки  $N(1)/2$  (т. е. деления суммы  $N(1)/2$  независимых биномиальных переменных с вероятностью успеха  $p_i$ , равной концентрации фамилии среди родоначальников). Соответствующее распределение количества успехов в выборке является биномиальным.

Ожидаемым (средним) значением концентрации  $x_i$  для  $i$ -й фамилии в новом поколении будет прежнее значение  $p_i$ , а дисперсия выборочных отклонений  $x_i$  от  $p_i$  равна  $p_i(1 - p_i)/\frac{N(1)}{2}$  в соответствии со свойствами биномиальных испытаний. Таким образом, *у случайного дрейфа нет преимущественного направления* (ожидаемое значение концентраций фамилий в следующем поколении совпадает с предыдущим значением).

При отсутствии направления у динамики фамильного состояния в результате случайного дрейфа ее можно характеризовать разбросом возможных отклонений состояний от начального значения, увеличивающимся в силу накопления выборочных ошибок в ряду поколений, т. е. характеризовать степенью и темпом дивергенции от исходного положения. Величина фамильной дивергенции за поколение, “темпер” ненаправленной эволюции, измеряемая, скажем, средним абсолютным отклонением или средним квадратическим

отклонением, или средним квадратом отклонения (дисперсией), равным  $V(x_i) = p_i(1 - p_i)/\frac{N(1)}{2}$  (а не просто средним отклонением, которое при ненаправленной эволюции равно нулю), зависит, как видим, от значения рассматриваемой концентрации  $p_i$  в родительской популяции. При одинаковых прочих условиях выборочная дисперсия  $x_i$  как характеристика скорости ненаправленной дивергенции определяется значением  $p_i$  (дисперсия пропорциональна  $p_i(1 - p_i)$ ). Тем самым темп дивергенции для концентрации  $i$ -го аллеля (фамилии) зависит от текущего значения  $p_i$  и со временем меняется вместе с ним, как подчеркивалось многими исследователями.

В один и тот же момент времени в одной и той же популяции темп дивергенции для другой фамилии с другой концентрацией будет в общем случае иным. Выборочные дисперсии характеризуют случайную ненаправленную динамику и могут служить показателем скорости дивергенции. Однако затруднительно определить, из-за чего различия в величине выборочного отклонения разных фамилий достигают наблюдаемого значения — объясняется ли это только темпом дивергенции, зависимым от их концентраций среди родителей, или причиной является, например, давление некоторого фактора. Кроме того, одинаковые значения отклонений характеризуются по-разному при разных значениях  $p$ .

При одновременном изучении концентраций множества фамилий разброс выборочных отклонений характеризуется матрицей ковариаций, зависящей от значений концентраций фамилий в родительской популяции. В случае разных концентраций фамилий темп дивергенции от начального состояния отличается как по различным направлениям (по разным осям координат, на которых откладываются концентрации соответствующих фамилий), так и по отдельной оси в зависимости от значения концентрации. Здесь возникает задача оптимального объединения без потери информации данных по отдельным фамилиям для получения единой характеристики дивергенции фамильного состояния от начального значения и между разными популяциями с общим происхождением.

На этом пути желательно использовать такое преобразование координат, когда темп дивергенции стабилен и не зависит ни от направления, ни от текущего состояния. Тогда упрощается построение обобщенной характеристики динамики отклонений состояния популяции от начальной точки для получения единого показателя, облегчающего сравнение популяций. Такими полезными статистическими особенностями дивергенции обладает, например, случай, когда выборочные отклонения имели бы стандартное многомерное нормальное

распределение с единичными дисперсиями по каждой из независимо и случайно изменяющихся переменных состояния.

Для получения подобного показателя рассмотрим в следующем разделе обобщение преобразования  $\theta = \arccos \sqrt{p}$  (облегчающего изучение динамики отдельной фамилии подобно используемому в [16]) на случай анализа одновременных изменений множества фамилий (см., например, [15]). При рассматриваемом обобщении достигается изотропность темпа отклонения (дивергенции) популяции от начального состояния в результате случайного дрейфа.

### ПРОСТРАНСТВО СОСТОЯНИЙ ПОПУЛЯЦИИ С МНОЖЕСТВОМ ФАМИЛИЙ КАК ОБЛАСТЬ ГИПЕРСФЕРЫ

В связи с зависимостью скорости дивергенции от состояния популяции  $x$  возникает задача добиться, чтобы характер случайных выборочных колебаний был бы одним и тем же для любого вектора концентраций  $x$  различных фамилий (аллелей), т. е. не зависел от состояния. Начать можно с такого преобразования отдельных концентраций, при котором на дисперсию преобразованной биномиальной переменной не влияет вероятность успеха. Для этой цели Р. Фишером были предложены арксинус-преобразование и преобразование  $\cos \theta = 1 - 2p$  [16, 17]. Последнее было использовано им для анализа генного дрейфа по концентрации одного аллеля, его применение стабилизировало выборочную дисперсию, принимающую постоянное значение. У нас подобный подход означает изучение свойств случайной динамики концентраций фамилий по отдельности. Результаты такого изучения остаются корректными в качестве части общей динамики при исследовании всего множества фамилий, так как процесс случайного дрейфа допускает произвольную группировку фамилий, предельным случаем которой будет группа из одной фамилии.

Для стабилизации темпа дивергенции по концентрации *одной* из фамилий можно использовать угловую переменную  $\theta$ , получаемую преобразованием  $\theta = \arccos \sqrt{p}$ , при котором выборочная дисперсия  $\theta$  не зависит от концентрации  $p$ . Геометрически углу  $\theta$  соответствует согласно школьному курсу тригонометрии точка на единичной окружности с центром в начале координат (точнее, у нас точка на части этой окружности в первой четверти) или радиус-вектор данной точки. При этом на оси абсцисс откладывается  $\sqrt{p}$ , косинус угла между радиус-вектором указанной точки  $p$  и осью абсцисс, а по оси ординат  $\sqrt{1 - p}$  (косинус угла с осью ординат) — см. рис. 1. Угол  $\theta$  измеряем

в радианах. На тригонометрической окружности он совпадает с длиной дуги между осью абсцисс и  $y = \sqrt{p}$ . При малых отклонениях  $y_1$  от  $y = y_0$  длина дуги между этими точками приближенно равна хорде между  $y_0$  и  $y_1$  или расстоянию между  $y_0$  и соответствующей  $y_1$  точкой на касательной прямой в точке  $y_0$  к окружности, что используется в дальнейшем. Каждая точка  $y = (\sqrt{p}, \sqrt{1-p})$  на окружности служит геометрическим образом фамильного состояния популяции при наличии только двух фамилий, а пространством состояний является часть тригонометрической окружности в первой четверти.

При обобщении этой картины на случай  $k$  фамилий (групп) в популяции [14] получим, что в пространстве состояний (фазовом пространстве) будет  $k$  осей координат, на которых откладывают  $k$  значений  $\{\sqrt{x_i}\}$ , направляющих косинусов  $y_i = \sqrt{x_i}$  радиуса-вектора состояния в  $k$ -мерном пространстве (косинусов углов  $\theta_i$  между радиус-вектором и  $i$ -й осью координат). На этом пути перейдем к более строгому изучению стабилизации темпа дивергенции в последовательности выборочных отклонений при случайному дрейфе популяции с несколькими группами однофамильцев. При этом через  $\mathbf{p}$  обозначаем начальное состояние популяции, а  $\mathbf{x}$  относим к состоянию с учетом выборочного отклонения.

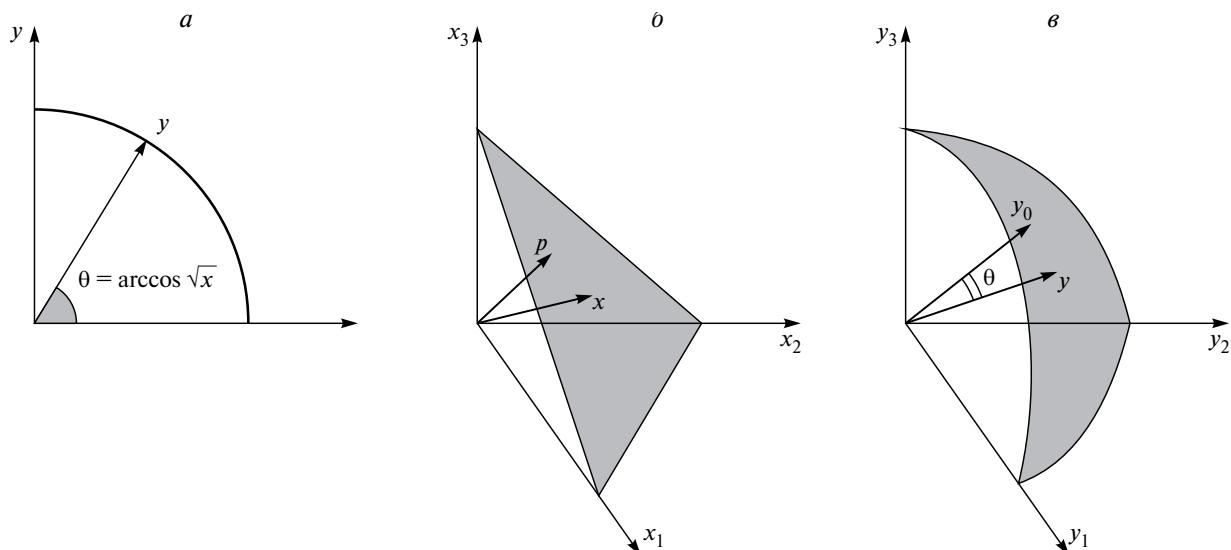
Итак, перейдем к изучению свойств преобразования координат  $\{y = \sqrt{x_i}\}$ , при котором по  $i$ -й оси откладывается корень квадратный из концентрации  $i$ -й фамилии. Все множество  $\{y\}$  состояний популяции в новых координатах состоит из точек части поверхности гиперсферы (с радиусом  $R = 1$  и с центром в начале координат), которая находится над полуосями неотрицательных координат с границей из состояний, у которых имеется нулевая координата. Например, для трех переменных (концентраций) множество состояний популяции состоит из точек  $\{y\}$  вида

$$\mathbf{y} = (y_1, y_2, y_3)^T = (\sqrt{x_1}, \sqrt{x_2}, \sqrt{x_3})^T,$$

$$R^2 = y_1^2 + y_2^2 + y_3^2 = x_1 + x_2 + x_3 = 1,$$

что иллюстрируется рис. 1. Оказывается, что дивергенция состояний, кроме примыкающих к границе фазового пространства, обладает желательными свойствами стабилизации (правда, за такое преимущество приходится платить ограничением на величину промежутка времени, когда преимущество существует).

Пусть  $\Theta(\mathbf{y}, \mathbf{y}_0)$  обозначает угол между двумя векторными состояниями  $\mathbf{y}$  и  $\mathbf{y}_0$  (соответствующими  $\mathbf{x}$  и  $\mathbf{p}$ , см. рис. 1). Отклонения  $\mathbf{x}$  от начального состояния  $\mathbf{p}$  ( $\mathbf{y}$  от  $\mathbf{y}_0$ ) можно измерять различными способами, из которых в координатах  $\mathbf{y}$  длина дуги большого круга на гиперсфере привлекательна своими статистическими и геометрическими свойствами. Она аналогична прямой в Евклидовом пространстве в том смысле, что также дает кратчайшее



**Рис. 1.** Пространство состояний популяции в различных системах координат.  
 а – затененный угол  $\theta$  между радиус-вектором  $y_0$  и осью абсцисс;  
 б – затененная часть плоскости как пространство состояний популяции в терминах концентраций групп;  
 в – затененная часть сферы как пространство состояний популяции в терминах квадратных корней из концентраций.  
 Объяснения см. в тексте.

расстояние между двумя точками (теперь на гиперсфере). Кроме того, длина дуги на единичной гиперсфере совпадает с угловым расстоянием  $\theta(\mathbf{x}, \mathbf{p}) = \theta(\mathbf{y}_1, \mathbf{y}_0)$ , которое, как говорилось, измеряем в радианах

$$\theta(\mathbf{x}, \mathbf{p}) \equiv \theta(\mathbf{y}_1, \mathbf{y}_0) \equiv \arccos \left( \sum_{i=1}^k \sqrt{x_i} \sqrt{p_i} \right). \quad (1)$$

Рассматриваемое преобразование было предложено в [18] с точки зрения, главным образом, изучения выборочных свойств статистики  $\cos \theta$ , ее связи с критерием хи-квадрат и др. В [14, 19] доказана изотропность пространства выборочных отклонений на гиперсфере. Если сдвинуть одинаковым образом как точку  $\mathbf{x}$ , так и  $\mathbf{p}$ , то обычное (Евклидов) расстояние между ними в Евклидовом пространстве останется прежним. Аналогично угловое расстояние  $\theta(\mathbf{x}, \mathbf{p})$  на гиперсфере не изменится при соответствующем сдвиге  $\mathbf{x}$  и  $\mathbf{p}$ . Описанная картина верна для любой популяции как совокупности, состоящей из непересекающихся групп однотипных объектов.

### АНАЛИЗ РАЗБРОСА УГОЛОВОГО ОТКЛОНЕНИЯ

При случайном фамильном дрейфе по одной фамилии с начальной концентрацией  $p$  в популяции результат добавления в выборку потомка (у нас в выборку размера  $N/2$  потомков мужского пола) является случайной величиной со значениями 1 (если у потомка окажется рассматриваемая фамилия, вероятность этого “успеха” равна  $p$ ) и 0 (в противном случае). Распределение суммы случайных величин (количество успехов во всей выборке), получаемой при этом, является *биномиальным*. Повторим, что согласно известным свойствам биномиального распределения с вероятностью успеха  $p$  (см., например, [20]) у полученной концентрации  $x$  в случайной выборке размера  $N/2$  математическое ожидание (среднее значение)  $E\{x\}$  и дисперсия  $V\{x\}$  равны соответственно

$$E\{x\} = p, \quad V\{x\} = p(1-p)/\frac{N}{2}.$$

Согласно центральной предельной теореме сумма достаточно большого количества сравнительно малых случайных величин ведет себя как нормальная случайная величина, т. е. при большом  $N/2$ , где  $N$  – размер популяции с учетом обоих полов, распределение  $x$  (суммы “успехов”, деленной на  $N/2$ ) является приближенно нормальным с приведенными значениями  $E\{x\}$  и  $V\{x\}$ , обозначаемым как  $N(p, p(1-p)/\frac{N}{2})$ .

При использовании преобразования  $\theta(p) = \arccos \sqrt{p}$  угла  $\theta(x)$  также будет приближенно нормальное распределение. Покажем,

что у него математическое ожидание приближенно равно  $\theta(p)$ , а дисперсия не зависит от  $p$ . У нас  $p$  обозначает исходную концентрацию рассматриваемой фамилии, а в контексте популяционной генетики  $p$  имеет смысл концентрации рассматриваемого аллеля аутосомного локуса в родоначальной популяции. Напомним, что когда речь идет об аллелях, для углового отклонения используем обозначение “ $\theta$ ”, а когда имеются в виду фамилии, к “ $\theta$ ” добавляем “ $s$ ”, т. е. значение  $\theta s(p)$  характеризует фамильное состояние популяции.

Найдем приближенно дисперсию значений  $\theta s(x)$  в новом поколении с помощью известного  $\delta$ -метода (см., например, [20]) следующим образом. Новая концентрация в поколении потомков получается прибавлением к  $p$  случайного выборочного отклонения  $\delta p$  с нулевым математическим ожиданием и дисперсией  $p(1-p)/\frac{N}{2}$ . Для получения примерного значения  $\theta s(x) = \theta s(p + \delta p)$  в следующем поколении используем член первого порядка по  $\delta p$  в разложении Тейлора  $\theta s(p)$ :  $\theta s(x) \approx \theta s(p) + (d\theta s(p)/dp)\delta p$ . Здесь  $\theta s(p)$  – константа,  $d\theta s(p)/dp$  – постоянный множитель при случайной переменной  $\delta p = x - p$  с нулевой средней величиной и с дисперсией  $p(1-p)/\frac{N}{2}$ , соответствующей дисперсии концентрации фамилии в следующем поколении.

Как известно, дисперсия произведения константы на случайную величину  $\delta p$  равна произведению дисперсии  $\delta p$  на квадрат константы. Отсюда вычисление приближенной (межпопуляционной) дисперсии  $V(\theta s)$  теоретически мыслимых вариантов фамильных состояний  $\theta s$  популяции в следующем поколении дает известное значение

$$V(\theta s) \approx (d\theta s(p)/dp)^2 p(1-p)/\frac{N}{2} = \frac{2}{N}.$$

Таким образом, при замене  $p$  на  $\theta s = \theta s(p)$  выборочное отклонение новой переменной  $\theta s$  от фамильного состояния родительской популяции  $\theta s(p)$  при достаточной величине  $N$  приближенно имеет нормальное распределение с нулевой средней и дисперсией  $\frac{2}{N}$ , независимой от значения  $p$ .

**Ремарка 1.** Отметим, что при выводе дисперсии  $V(\theta s)$  мы в разложении Тейлора ограничились членом с первой производной  $d\theta s(p)/dp$  и пренебрегли следующими. Однако уже вторая производная неограниченно растет, когда  $p$  стремится к нулю. Поэтому можно пользоваться полученной аппроксимацией дисперсии, когда  $p$  превышает надлежащий порог, выбираемый из условия малости эффекта следующего члена в разложении Тейлора для  $\theta s$ .

Обоснуем более строго свойства углового расстояния при случайном дрейфе генов и фамилий.

**Результат 2.** Пусть концентрация  $x$  аллеля (фамилии) в популяции с неперекрывающимися поколениями определяется при каждой смене поколений случайной выборкой с возвращением из совокупности аллелей рассматриваемого локуса (из совокупности фамилий) родительского поколения. Положим, что размер такой выборки на шаге  $\tau$  равен  $2Ne(\tau)\left(\frac{Ne(\tau)}{2}\right)$ , где  $Ne(\tau)$  обозначает эффективную численность популяции в поколении  $\tau$  (см., например, [11, 12]).

Пусть  $t$  поколений тому назад концентрация данного аллеля (фамилии) в популяции была равна  $p$ , и расстояние (отклонение) между текущим  $x$  и начальным  $p$  состояниями в угловых координатах  $\theta(x) = \arccos(\sqrt{x})$  находится как  $|\theta(\sqrt{x}) - \theta(\sqrt{p})|$  (соответственно как  $|\theta_s(\sqrt{x}) - \theta_s(\sqrt{p})|$ ).

Тогда при  $\frac{\widetilde{Ne}(t)}{t} \rightarrow \infty$  и при  $\sqrt{p}$ , превышающем надлежащим образом выбранный порог, асимптотические распределения для  $\theta(\sqrt{x(t)})$  и  $|\theta(\sqrt{x(t)}) - \theta(\sqrt{p})|^2$ , а также для  $\theta_s$  (переменной фамильного состояния, аналогичной  $\theta$ ) и их параметры имеют вид:

$$\begin{aligned} E\{\theta(\sqrt{x(t)})\} &= \theta(\sqrt{p}), \quad E\left\{\left|\theta(\sqrt{x(t)}) - \theta(\sqrt{p})\right|^2\right\} = \frac{t}{8\widetilde{Ne}(t)}, \\ \theta(\sqrt{x(t)}) &= N\left(\sqrt{p}, \frac{t}{8\widetilde{Ne}(t)}\right), \quad \frac{8\widetilde{Ne}(t)}{t} \left|\theta(\sqrt{x(t)}) - \theta_s(\sqrt{p})\right|^2 = \chi^2, \\ E\{\theta_s(\sqrt{x(t)})\} &= \theta_s(\sqrt{p}), \quad E\left\{\left|\theta_s(\sqrt{x(t)}) - \theta_s(\sqrt{p})\right|^2\right\} = \frac{t}{2\widetilde{Ne}(t)}, \\ \theta_s(\sqrt{x(t)}) &= N\left(\sqrt{p}, \frac{t}{2\widetilde{Ne}(t)}\right), \quad \frac{2\widetilde{Ne}(t)}{t} \left|\theta_s(\sqrt{x(t)}) - \theta_s(\sqrt{p})\right|^2 = \chi^2. \end{aligned} \quad (2)$$

Здесь  $N(m, V)$  – символ нормального распределения с математическим ожиданием  $m$  и дисперсией  $V$ ,  $\widetilde{Ne}(t) = t / \sum_{\tau=1}^t \frac{1}{Ne(\tau)}$ ,  $\widetilde{Ne}(t)$  и  $\chi^2$  обозначают среднюю гармоническую численность популяции для ряда  $\{Ne(\tau), \tau = 1, 2, \dots, t\}$  и распределение хи-квадрат (одной степенью свободы) соответственно.

**Доказательство** проведем для конкретности в случае анализа фамильных состояний. Рассмотрим последовательность нескольких поколений, в  $t$ -м из которых фамильное состояние популяции представляет собой выборку  $Ne(\tau)/2$  фамилий из предыдущего поколения. Пусть на первом шаге реализовалось состояние  $\theta_s(1) = \theta_s(x(1))$ . Случайное отклонение  $\delta\theta_s(1)$  нового значения  $\theta_s$  от начального состояния  $\theta_s(p)$ , как говорилось ранее, приближенно имеет нормальное распределение с нулевой средней и независимой от  $p$  дисперсией  $\frac{1}{2Ne(1)}$ . На втором шаге следующее выборочное отклонение  $\delta\theta_s(2)$  не коррелирует с предыдущим и будет

нормальным с нулевой средней и независимой от  $x(1)$  дисперсией  $\frac{1}{2Ne(2)}$ , где  $Ne(2)$  – очередной эффективный размер популяции. Распределение суммы двух нормально распределенных некоррелирующих случайных отклонений с нулевыми средними и дисперсиями  $\frac{1}{2Ne(1)}$  и  $\frac{1}{2Ne(2)}$  является нормальным распределением с нулевой средней и дисперсией  $\frac{1}{2Ne(1)} + \frac{1}{2Ne(2)}$ . Продолжая эти рассуждения, мы получим, что в поколении  $t$  итоговое суммарное отклонение значения  $\theta_s$  от начальной величины приближенно распределено нормально с нулевой средней и дисперсией  $\sum_{\tau=1}^t \frac{1}{2Ne(\tau)} = \frac{t}{2\widetilde{Ne}(t)}$ , складывающейся из дисперсий отклонений на отдельных шагах (поколениях). Здесь  $\widetilde{Ne}(t) = t / \sum_{\tau=1}^t \frac{1}{Ne(\tau)}$ ,  $Ne(\tau)$  обозначает эффективный размер популяции в поколении  $\tau$ ,  $\widetilde{Ne}(t)$  – средняя гармоническая численность популяции для ряда  $\{Ne(\tau)\}$ , равная обратной величине к среднему арифметическому  $(\sum_{\tau=1}^t \frac{1}{Ne(\tau)})/t$  для  $\{\frac{1}{Ne(\tau)}\}$ . Отсюда вытекает, что у нормированного квадрата углового расстояния  $\frac{2\widetilde{Ne}(t)}{t} \theta_s^2$  будет распределение хи-квадрат с одной степенью свободы. ▲

Распределение отклонения  $\theta_s(t)$  от начального значения в простом частном случае постоянно-го размера  $N$  у популяции приближенно является нормальным  $N(0, \frac{t}{2N})$  с нулевым математическим ожиданием и дисперсией  $\frac{t}{2N}$ , а нормированный квадрат углового расстояния  $\frac{2N}{t} \theta_s^2$  имеет распределение хи-квадрат с одной степенью свободы. Еще раз напомним, что приведенные результаты корректны, когда значение  $t$  мало по сравнениюю  $\widetilde{Ne}(t)$ .

Хотя далее у нас речь идет о совокупности фамилий, результаты имеют общий характер и применимы к случайным выборкам из любой совокупности дискретных объектов, сгруппированных согласно их типам с соответствующими вероятностями попадания типов в выборку. Таким образом, если слово “фамилия” заменить на название объекта, то выводы останутся верными для такого случая, например для концентраций аллелей. Доказываемые факты относительно выборочных свойств угла  $\theta$  являются вариантом результата, полученного [18] в области статистики, адаптированным применительно к дрейфу фамилий в духе анализа

генного дрейфа в [14]. Сформулируем эти факты более строго для многомерного случая.

**Результат 3.** Пусть дана случайная выборка с возращением размера  $N/2$  из популяции с  $k$  вариантами фамилий. Положим, вероятность извлечения  $i$ -й фамилии равна ее концентрации  $x_i > 0$ ,  $i = 1, 2, \dots, k$ ,  $\sum_{i=1}^k x_i = 1$ . Вектор с координатами  $\{x_i\}$  обозначим как  $\mathbf{x}$ , а случайный вектор концентраций фамилий в выборке как  $\mathbf{x}_1 = \mathbf{x}(1)$  и определим преобразование  $\mathbf{y}(\mathbf{x})$  как

$$\mathbf{y}(\mathbf{x}) = (y_1, y_2, \dots, y_k)^T \equiv (\sqrt{x_1}, \sqrt{x_2}, \dots, \sqrt{x_k})^T, \\ \sum_{i=1}^k y_i^2 = 1, \quad \mathbf{y}(1) \equiv \mathbf{y}(\mathbf{x}(1)), \quad \mathbf{y}_0 = \mathbf{y}(\mathbf{x}_0). \quad (3)$$

Тогда асимптотически при  $N \rightarrow \infty$

$$\mathbf{y}(1) = N \left( \mathbf{y}_0, \frac{1}{2N} \mathbf{W}(\mathbf{y}_0) \right), \quad \mathbf{W}(\mathbf{y}_0) \equiv \mathbf{I} - \mathbf{y}_0 \mathbf{y}_0^T. \quad (4)$$

Здесь  $N(\mathbf{m}, \mathbf{V})$  – символ многомерного нормального распределения с вектором математического ожидания  $\mathbf{m}$  и матрицей ковариаций  $\mathbf{V}$ ,  $\mathbf{I}$  – единичная матрица.

**Доказательство.** Распределение фамильного состава выборки, получаемой при сделанных предположениях, является полиномиальным (мультиномиальным). Согласно известным свойствам полиномиального распределения с вероятностями  $\{x_i\} = \mathbf{x}$  (см., например, [20]) у полученных в случайной выборке размера  $\frac{N}{2}$  концентраций  $\mathbf{x}_1$  (и их отклонений от  $\mathbf{x}$ ) матрица ковариаций  $\mathbf{V}(\mathbf{x})$  имеет вид

$$\mathbf{V}(\mathbf{x}) = \frac{2}{N} \left[ x_i (\delta_{ij} - x_j) \right] = \frac{2}{N} \left( \mathbf{D}(\mathbf{x}) - \mathbf{x} \mathbf{x}^T \right) = \\ = \frac{2}{N} \left( \mathbf{D}^2(\mathbf{y}) - (\mathbf{D}(\mathbf{y}) \mathbf{y})(\mathbf{y}^T \mathbf{D}(\mathbf{y})) \right),$$

где  $\delta_{ij}$  обозначает символ Кронекера ( $\delta_{ij} = 1$  при  $i = j$  и нулю в противном случае),  $\mathbf{D}(\mathbf{x}) \mathbf{D}(\mathbf{y})$  – диагональная матрица с координатами вектора  $\mathbf{x}$  (соответственно  $\mathbf{y}$ ) на главной диагонали. Дальнейшее доказательство разобьем на пункты.

1. При преобразовании (3) можно приближенно найти математическое ожидание  $E\{\mathbf{y}(\mathbf{x})\}$  и матрицу ковариаций  $\mathbf{V}(\mathbf{y})$  для координат  $\{y_i\}$  вектора  $\mathbf{y} = \mathbf{y}(\mathbf{x})$  с помощью  $\delta$ -метода, использующего члены первого порядка в разложении Тейлора  $\mathbf{y}(\mathbf{x})$ . Применим его как к дисперсиям (см., скажем, [20]), так и ковариациям. Повторим, что когда  $y(x)$  получается

преобразованием случайной переменной  $x$  с математическим ожиданием  $x_0$ , то

$$y(x) = y(x_0 + (x - x_0)) \equiv y(x_0 + \delta x) \approx y(x_0) + (dy(x_0)/dx)\delta x, \\ \delta y \equiv y(x) - y(x_0) \approx (dy(x_0)/dx)\delta x,$$

где  $dy(x_0)/dx$  – константа, а  $\delta x$  – случайное отклонение  $x$  от  $x_0$  с нулевым математическим ожиданием ( $E\{\delta x\} = 0$ ). Отсюда

$$E\{y(x)\} \approx E\{y(x_0)\} + (dy(x_0)/dx)E\{\delta x\} = y(x_0), \\ V(y) \equiv E\{\delta y^2\} \approx (dy(x_0)/dx)^2 V(x).$$

2. Когда  $\mathbf{y}(\mathbf{x})$  получен преобразованием случайного вектора  $\mathbf{x}$  с математическим ожиданием  $\mathbf{x}_0$ , то для  $i$ -й координаты  $y_i(\mathbf{x})$  вектора  $\mathbf{y}(\mathbf{x})$  имеем

$$y_i(\mathbf{x}) = y_i(x_0 + (x - x_0)) \equiv y_i(x_0 + \delta x) \approx \\ \approx y_i(x_0) + \sum_j (\partial y_i(x_0)/\partial x_j) \delta x_j,$$

где  $\partial y_i(x_0)/\partial x_j$  – константы, а  $\delta x_j$  – случайные отклонения координат  $\mathbf{x}$  от координат  $\mathbf{x}_0$  с нулевым математическим ожиданием ( $E\{\delta x_j\} = 0$ ).

В векторно-матричном виде это соотношение можно переписать как

$$\mathbf{y}(\mathbf{x}) = \mathbf{y}(\mathbf{x}_0 + \delta \mathbf{x}) \approx \mathbf{y}(\mathbf{x}_0) + [\partial_{ij} \mathbf{y}(\mathbf{x}_0)] \delta \mathbf{x}, \\ \delta \mathbf{y} \equiv \mathbf{y}(\mathbf{x}) - \mathbf{y}(\mathbf{x}_0) \approx [\partial_{ij} \mathbf{y}(\mathbf{x}_0)] \delta \mathbf{x}, [\partial_{ij} \mathbf{y}(\mathbf{x}_0)] \equiv [\partial y_i(x_0)/\partial x_j].$$

Так как константы  $\partial_{ij} \mathbf{y}(\mathbf{x}_0)$  можно выносить за знак математического ожидания  $E$  и согласно полученному выше  $E\{\delta \mathbf{x}\} = \mathbf{0}$ , то

$$E\{\mathbf{y}(\mathbf{x})\} \approx E\{\mathbf{y}(\mathbf{x}_0)\} + [\partial_{ij} \mathbf{y}(\mathbf{x}_0)] E\{\delta \mathbf{x}\} = \mathbf{y}(\mathbf{x}_0).$$

3. Теперь обратимся к вычислению матрицы ковариаций  $\mathbf{V}(\mathbf{y})$  случайного вектора-столбца  $\mathbf{y}$ . По определению  $\mathbf{V}(\mathbf{y}) \equiv E\{\delta \mathbf{y} \times \delta \mathbf{y}^T\}$ ,  $\mathbf{V}(\mathbf{x}) \equiv E\{\delta \mathbf{x} \times \delta \mathbf{x}^T\}$ , подстановка в  $\mathbf{V}(\mathbf{y})$  приведенного выше значения  $\delta \mathbf{y}$  дает

$$\mathbf{V}(\mathbf{y}) = \mathbf{V}(\delta \mathbf{y}) \approx E\{[\partial_{ij} \mathbf{y}(\mathbf{x}_0)] \delta \mathbf{x} \times \delta \mathbf{x}^T [\partial_{ij} \mathbf{y}(\mathbf{x}_0)]^T\} = \\ = [\partial_{ij} \mathbf{y}(\mathbf{x}_0)] E\{\delta \mathbf{x} \times \delta \mathbf{x}^T\} [\partial_{ij} \mathbf{y}(\mathbf{x}_0)]^T = \\ = [\partial_{ij} \mathbf{y}(\mathbf{x}_0)] \mathbf{V}(\mathbf{x}) [\partial_{ij} \mathbf{y}(\mathbf{x}_0)]^T. \quad (5)$$

Напомним, что при анализе фамильной структуры размер случайной выборки равен  $N/2$ ,  $\mathbf{y}(\mathbf{x}) = \{\sqrt{x_i}\}$  и

$$\mathbf{V}(\mathbf{x}) = \frac{2}{N} (\mathbf{D}(\mathbf{x}) - \mathbf{x} \mathbf{x}^T) = \frac{2}{N} (\mathbf{D}^2(\mathbf{y}) - (\mathbf{D}(\mathbf{y}) \mathbf{y})(\mathbf{y}^T \mathbf{D}(\mathbf{y}))),$$

$$[\partial_{ij} \mathbf{y}(\mathbf{x})] \equiv [\partial y_i / \partial x_j] = \frac{1}{2} \left[ \frac{\delta_{ij}}{\sqrt{x_i}} \right] = \frac{1}{2} \mathbf{D}^{-1}(\mathbf{y}).$$

Подстановка этих выражений в формулу (5) для  $\mathbf{V}(\mathbf{y})$  дает

$$\begin{aligned}\mathbf{V}(\mathbf{y}) &\approx \frac{1}{2} \mathbf{D}^{-1}(\mathbf{y}) \left( \frac{2}{N} (\mathbf{D}^2(\mathbf{y}) - (\mathbf{D}(\mathbf{y})\mathbf{y})(\mathbf{y}^T \mathbf{D}(\mathbf{y}))) \right) \frac{1}{2} \mathbf{D}^{-1}(\mathbf{y}) = \\ &= \frac{1}{2N} (\mathbf{I} - \mathbf{y}\mathbf{y}^T) \equiv \frac{1}{2N} \mathbf{W}(\mathbf{y}),\end{aligned}$$

где  $\mathbf{I}$  – единичная матрица,  $\mathbf{W}(\mathbf{y}) = [\delta_{ij} - y_i y_j] = \mathbf{I} - \mathbf{y}\mathbf{y}^T$ . Значит, в результате замены (3) приближенно матрица ковариаций  $\mathbf{V}(\mathbf{y})$  для новых переменных  $\mathbf{y}$  (и для их отклонений  $\delta\mathbf{y}$  от математического ожидания  $\mathbf{y}_0$ ) пропорциональна

$$\mathbf{W}(\mathbf{y}) = [\delta_{ij} - y_i y_j] \text{ с множителем } \frac{1}{2N} \text{ при анализе фамильного дрейфа (и } \frac{1}{8N} \text{ при дрейфе генов).}$$

**4.** Чем больше размер выборки, тем теснее располагаются выборочные отклонения вблизи нулевого значения и тем лучше аппроксимируется их распределение многомерным нормальным, причем асимптотически

$$\begin{aligned}\mathbf{x}_1 &= N\left(\mathbf{x}, \frac{2}{N}(\mathbf{D}(\mathbf{x}) - \mathbf{x}\mathbf{x}^T)\right), \quad \mathbf{y} = N\left(\mathbf{y}_0, \frac{1}{2N} \mathbf{W}(\mathbf{y}_0)\right), \\ \delta\mathbf{y} &= \mathbf{y} - \mathbf{y}_0 = N\left(0, \frac{1}{2N} \mathbf{W}(\mathbf{y}_0)\right).\end{aligned}$$

Теперь покажем, что распределение квадратов Евклидова  $|\delta\mathbf{y}|$  и углового  $\theta_s$  расстояний между  $\mathbf{y}_1$  и  $\mathbf{y}_0$  является широко употребляемым в биометрии распределением хи-квадрат.

**Результат 4.** В рамках предыдущего результата Евклидово  $|\delta\mathbf{y}|$  и угловое  $\theta_s$  расстояния между  $\mathbf{y}_1$  и  $\mathbf{y}_0$  удовлетворяют

$$2N|\delta\mathbf{y}|^2 \equiv 2N(\mathbf{y}_1 - \mathbf{y}_0, \mathbf{y}_1 - \mathbf{y}_0) = 2N(\delta\mathbf{y}, \delta\mathbf{y}) = \chi_{k-1}^2,$$

$$2N\theta_s^2(\mathbf{y}_1, \mathbf{y}_0) = \chi_{k-1}^2; \theta_s = \arccos\left(\sum_{i=1}^k \sqrt{x_i(t)} \sqrt{p_i}\right), \quad (6)$$

где  $\chi_{k-1}^2$  обозначает распределение хи-квадрат с  $k-1$  степенями свободы.

**Доказательство.** При сделанных ранее предположениях угловое расстояние  $\theta_s$  аппроксимирует Евклидово расстояние на касательной плоскости к гиперсфере. Соответственно распределения квадратов этих расстояний приближенно одинако-

вы. Матрица ковариаций для  $\delta\mathbf{y}$  равна  $\frac{1}{2N} \mathbf{W}(\mathbf{y}_0)$ ,  $\mathbf{W}(\mathbf{y}) = \mathbf{I} - \mathbf{y}\mathbf{y}^T$ ;  $\mathbf{y}^T \mathbf{y} = 1$  и для  $\sqrt{2N} \delta\mathbf{y}$  она равна  $\mathbf{W}(\mathbf{y}_0)$ . Здесь  $\mathbf{I}$  – единичная матрица размера  $k$ .

Заметим, что матрица  $\mathbf{W}(\mathbf{y})$  является идемпотентной (т. е., как можно легко проверить,

$\mathbf{W}^2(\mathbf{y}) = \mathbf{W}(\mathbf{y})$ ). Кроме того, приближенно  $E\{\delta\mathbf{y}\} = 0$ . Известно, что для такого случая

$$(\sqrt{2N} \delta\mathbf{y}, \sqrt{2N} \delta\mathbf{y}) = 2N(\delta\mathbf{y}, \delta\mathbf{y}) = \chi_{\text{tr} \mathbf{W}}^2 = \chi_{k-1}^2,$$

где  $\text{tr} \mathbf{W}$  обозначает след матрицы  $\mathbf{W}$  (сумму ее диагональных элементов, равную у нас  $k-1$ ). Следовательно, произведение  $2N$  на квадрат расстояния (отклонения  $|\delta\mathbf{y}|$ ) между фамильными состояниями  $\mathbf{y}_1 = \mathbf{y}(1)$  и  $\mathbf{y}_0$  приближенно имеет распределение хи-квадрат с  $k-1$  степенями свободы. При больших  $N$  отклонения  $\delta\mathbf{y}$  с близкой к единице вероятностью малы, значения углового расстояния  $\theta_s^2(\mathbf{y}_1, \mathbf{y}_0)$  аппроксимируются величинами  $|\delta\mathbf{y}|^2$  и асимптотически  $2N\theta_s^2(\mathbf{y}_1, \mathbf{y}_0) = \chi_{k-1}^2$ . ▶

Обратимся к более наглядной геометрической картине приведенного результата [14]. Напомним, что переход (3) от координат пространства фамильных состояний  $\mathbf{x}$ ,

$$\mathbf{x} = (x_1, x_2, \dots, x_k)^T, \quad \sum_{i=1}^k x_i = (\mathbf{x}, \mathbf{e}) = 1, \quad \mathbf{e} \equiv (1, 1, \dots, 1)^T,$$

как части гиперплоскости над полуосами неотрицательных координат, к координатам  $\{y_i = \sqrt{x_i}\}$  геометрически означает преобразование пространства фамильных состояний (симплекс) в часть гиперсферы с единичным радиусом (см. рис. 1 в трехмерном случае). При этом матрица ковариаций для выборочных отклонений  $\delta\mathbf{y}$  новых переменных приближенно равна  $\mathbf{V}(\mathbf{y})$  со следующими легко проверяемыми свойствами:

$$\begin{aligned}\mathbf{V}(\mathbf{y}) &= \frac{1}{2N} \mathbf{W}(\mathbf{y}), \quad \mathbf{W}(\mathbf{y}) \equiv \mathbf{I} - \mathbf{y}\mathbf{y}^T = [\delta_{ij} - y_i y_j], \\ \mathbf{y}^T \mathbf{y} &= 1; \quad \mathbf{W}(\mathbf{y})\mathbf{y} = 0, \quad \mathbf{W}(\mathbf{y})\mathbf{v} = \mathbf{v}, \quad \mathbf{v}^T(\mathbf{y}, \mathbf{v}) = 0.\end{aligned}$$

Таким образом, вектор  $\mathbf{y}$  является собственным вектором матриц  $\mathbf{V}(\mathbf{y})$  и  $\mathbf{W}(\mathbf{y})$  с собственным числом  $\lambda = 0$ , остальные собственные векторы  $\mathbf{W}(\mathbf{y})$  ортогональны  $\mathbf{y}$  с равными единице собственными числами. Очевидно,  $\mathbf{y}$  является вектором единичной нормали к гиперсфере (3).

Перейдем к новой системе координат, в которой на одной из осей лежит вектор нормали, а остальные оси располагаются в касательной плоскости в точке  $\mathbf{y}$ , образуя ортонормированную систему. Данное преобразование с ортонормированной матрицей перехода не изменяет распределение выборочных отклонений и собственных чисел матрицы ковариаций. В новой системе координат она является диагональной, ее главная диагональ состоит из дисперсий по новым координатам (из собственных чисел  $\lambda$ ). Дисперсия (собственное число) по нормали равна нулю, т. е. выборочные отклонения по

направлению вектора нормали невозможны. Это означает, что вся выборочная изменчивость сконцентрирована на гиперсфере. Будем аппроксимировать ее в касательной плоскости. Повторим, что равенство  $\mathbf{W}\mathbf{v} = \mathbf{v}$  означает, что любой вектор  $\mathbf{v}$ , ортогональный  $\mathbf{y}$  (лежащий в касательной плоскости к гиперсфере в точке  $\mathbf{y}$ ), будет собственным для матрицы  $\mathbf{W}$  с  $\lambda = 1$  [14]. Поэтому на касательной плоскости дисперсии выборочных отклонений одинаковы по любому направлению (изотропность). Займемся изучением свойств  $\theta_s$  в ряду поколений.

### АППРОКСИМАЦИЯ УГОЛОВОГО ОТКЛОНЕНИЯ ПО МНОЖЕСТВУ ФАМИЛИЙ ПРИ ОТНОСИТЕЛЬНО МАЛОМ КОЛИЧЕСТВЕ ПОКОЛЕНИЙ

Большинство из описанных свойств углового расстояния по множеству фамилий соответствуют свойствам выборки, которая с точки зрения случного дрейфа рассматривается как характеристика только одного шага в цепи изменений популяции в поколениях. Иная точка зрения фокусируется по рекомендации Р. Фишера на свойствах динамики углового расстояния в ряду поколений в результате случайных выборочных колебаний состояния популяции. Она была широко популяризована Л. Кавалли-Сфорца с соавт. [21, 12] в отношении генного дрейфа на относительно малом промежутке времени, не приводившими, однако, теоретических обоснований. В дальнейшем анализ данной ситуации был представлен в [14], изложен также в [22].

Геометрические свойства процесса случного дрейфа, аппроксимируемого диффузионным процессом на гиперсфере в римановом пространстве, и его асимптотика на небольших промежутках времени рассматривались в [23, 24]. Общий случай асимптотики диффузионных процессов в римановом пространстве на небольших временах рассмотрен в предположении невырожденной матрицы диффузии внутри и на границе фазового пространства в [25]. В монографии [26] проанализированы разносторонние информационно-геометрические свойства модели Райта – Фишера, одним из примеров которой является рассматриваемый случай фамильного дрейфа.

Приведенные выше результаты о свойствах углового расстояния для множества фамилий в контексте динамики фамильного состояния соответствуют однократной смене поколений, сопровождаемой случным выборочным отклонением фамильного состава популяции от  $\mathbf{y}_0$  до  $\mathbf{y}(1)$ . Конечно, наибольший интерес представляет динамика в течение не одного, а ряда поколений. Эта динамика описывается результатами последовательности вложенных выборок, соответствующими последовательности смены поколений популяции.

**Ремарка 5.** Проанализируем второе поколение под другим углом зрения. Рассмотрим его как гипотетический ансамбль популяций, состоящий из возможных вариантов популяций-потомков популяций первого поколения. Этот ансамбль можно интерпретировать как иерархически подразделенную метапопуляцию. Она разбивается на группы, происходящие от разных популяций первого поколения. В силу ненаправленного характера случного дрейфа ожидаемые концентрации фамилий в каждой отдельной группе популяций совпадают с концентрациями у породившей группу популяции первого поколения. Поэтому межгрупповая дисперсия распределения концентраций отдельной фамилии *по группам* второго поколения такая же, как дисперсия распределения ее концентраций *по популяциям* первого поколения. Аналогично *межгрупповая матрица ковариаций*  $\mathbf{Vs}_{betw}(\mathbf{y}(2)|\mathbf{y}(0))$  совпадает с матрицей ковариаций распределения концентраций по популяциям первого поколения, равной согласно (4) значению  $\frac{1}{2Ne(1)}\mathbf{W}(\mathbf{y}_0)$ :

$$\mathbf{Vs}_{betw}(\mathbf{y}(2) | \mathbf{y}(0)) = \mathbf{Vs}(\mathbf{y}(1) | \mathbf{y}(0)) = \frac{1}{2Ne(1)}\mathbf{W}(\mathbf{y}_0).$$

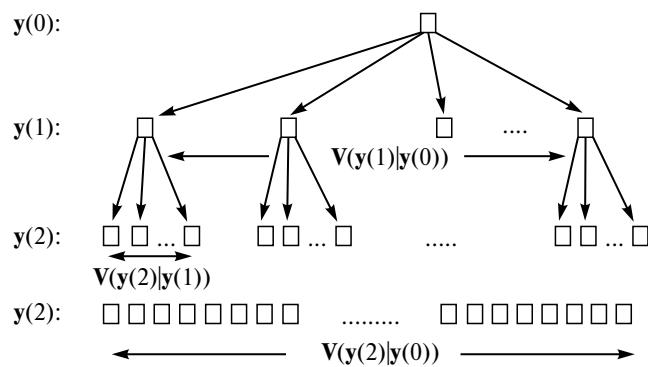
В каждой отдельной группе популяций, происходящих от некоторой популяции первого поколения с концентрациями фамилий  $\mathbf{y}(1)$ , матрица ковариаций согласно (4) равна  $\frac{1}{2Ne(2)}\mathbf{W}(\mathbf{y}_1)$  и рассматривается как *внутригрупповая матрица ковариаций*.

Итак, все популяции второго поколения образуют метапопуляцию, подразделенную на группы (см. рис. 2). Дивергенция групп между собой по концентрации отдельной фамилии характеризуется *межгрупповой дисперсией* распределения ее концентраций по группам. Дивергенция популяций внутри групп метапопуляции характеризуется *средней внутригрупповой дисперсией* распределения концентраций фамилии по популяциям внутри групп.

Дивергенция популяций всего ансамбля характеризуется *общей (полной) дисперсией* распределения концентраций фамилий по всем популяциям второго уровня. По *правилу сложения дисперсий* (см., например, его применение к фамильной структуре в [27]) полная дисперсия равна сумме межгрупповой и средней внутригрупповой дисперсий. Это правило остается верным в случае не одной, а множества фамилий, если слово “дисперсия” заменить на “матрица ковариаций”.

Рассмотрим динамику матрицы ковариаций  $\mathbf{Vs}(\mathbf{y}(t))$ . Начнем изучение со случая двух поколений случного дрейфа.

**Результат 6.** Пусть рассматривается последовательность двух независимых вложенных случайных



**Рис. 2.** Межгрупповая, внутригрупповые и полная матрицы ковариаций концентраций состояния популяций:

□ – обозначение популяции; \$y(0)\$ – состояние родоначальной популяции; \$y(1)\$ – случайные состояния ее потомков, популяций первого поколения, разброс возможных состояний \$y(1)\$ характеризуется матрицей ковариаций \$V(y(1)|y(0))\$, служащей межгрупповой матрицей ковариаций для популяций следующего поколения, где \$y(0)\$ фиксировано; \$y(2)\$ – случайные состояния популяций, потомков родоначальной популяции во втором поколении. Они образуют метапопуляцию, состоящую из групп с происхождением от отдельных популяций первого поколения и с внутригрупповыми матрицами ковариаций \$V(y(2)|y(1))\$. Здесь \$y(1)\$ случайно варьирует между группами; следующая строка относится ко второму поколению без разбиения на группы его популяций, разброс которых характеризуется полной матрицей ковариаций \$Vs(y(2))\$. Стрелки, направленные сверху вниз, соединяют родительскую популяцию с популяцией потомков.

выборок с возвращением из популяции с \$k\$ вариантами фамилий, где их концентрации превышают надлежащий порог. Положим, что размер такой выборки (эффективный размер популяции) в поколении \$\tau = 1, 2\$ равен \$\frac{Ne(\tau)}{2}\$, а вероятность попадания в выборку фамилии \$i\$-го типа равна \$x\_i(\tau) > 0\$, \$\sum\_{i=1}^k x\_i(\tau) = 1\$, где \$x\_i\$ – ее концентрация в родительской популяции.

Тогда во втором поколении матрица ковариаций \$Vs(y(2))\$ случайного вектора \$y(2) \equiv \{\mathbf{y}\_i(2) \equiv \sqrt{x\_i(2)}\}\$ приближенно выражается как

$$\begin{aligned} Vs(y(2)) &\approx \left( \frac{1}{2Ne(1)} + \frac{1}{2Ne(2)} \right) (\mathbf{I} - \mathbf{y}(0)\mathbf{y}^T(0)) \equiv \\ &\equiv \left( \frac{1}{2Ne(1)} + \frac{1}{2Ne(2)} \right) \mathbf{W}(\mathbf{y}_0). \end{aligned}$$

**Доказательство.** Учтем, что при смене поколений к фамильному состоянию популяции

добавляется случайное выборочное отклонение \$\delta\$, причем \$E\{\delta\} = \mathbf{0}\$ независимо от номера поколения. Для первого поколения имеем

$$\begin{aligned} E\{y(1)|y(0)\} &= E\{y(0) + \delta(1)|y(0)\} = y(0), \\ E\{y(2)|y(1)\} &= E\{y(1) + \delta(2)|y(1)\} = y(1), \\ Vs(y(1)|y(0)) &= Vs(\delta(1))|y(0)) = \\ &= \frac{1}{2Ne(1)} (\mathbf{I} - \mathbf{y}(0)\mathbf{y}^T(0)) \equiv \frac{1}{2Ne(1)} \mathbf{W}(\mathbf{y}_0). \end{aligned}$$

Отсюда очевидно, что матрица ковариаций во втором поколении при условии \$y(1)\$ равна  $\frac{1}{2Ne(2)} \mathbf{W}(\mathbf{y}_1)$ .

Рассмотрим ансамбль популяций второго поколения, интерпретируемый как иерархически подразделенная метапопуляция, состоящая из групп. Каждая группа порождается соответствующей популяцией первого поколения и содержит возможные варианты популяций-потомков с разными фамильными состояниями. В нашем случае межгрупповая матрица ковариаций согласно предыдущей remarке имеет вид

$$Vs_{betw}(y(2)|y(0)) = Vs(y(1)|y(0)) = \frac{1}{2Ne(1)} \mathbf{W}(\mathbf{y}_0).$$

Дисперсия распределения концентрации отдельной фамилии по популяциям внутри какой-либо группы, порождаемой популяцией первого поколения с фамильным состоянием \$y(1)\$ (при условии \$y(1)\$), является внутригрупповой дисперсией (\$Vs\_{in}\$) для этой группы, а при рассмотрении множества фамилий вместо дисперсии имеем матрицу ковариаций \$Vs\_{in}(y(2)|y(1))\$ концентраций фамилий, являющуюся внутригрупповой матрицей ковариаций. Согласно правилу сложения дисперсий (ковариаций) полные (для всей метапопуляции-ансамбля) значения данных характеристик изменчивости равны сумме межгруппового и среднего внутригруппового значений этих показателей, т. е. полная матрица ковариации \$Vs\_{tot}(y(2))\$ концентраций фамилий популяций второго поколения имеет вид

$$\begin{aligned} Vs_{tot}(y(2)|y(0)) &= Vs_{betw}(y(2)|y(0)) + \\ &+ E_{y(1)}\{Vs_{in}(y(2)|y(1))\} = \\ &= Vs(y(1)|y(0)) + E_{y(1)}\{Vs(y(2)|y(1))\} = \\ &= \frac{1}{2Ne(1)} \mathbf{W}(\mathbf{y}_0) + E_{y(1)}\left\{\frac{1}{2Ne(2)} \mathbf{W}(\mathbf{y}_1)|y(0)\right\}. \end{aligned}$$

Рассмотрим второе слагаемое. Вспомним, что для любой случайной величины \$x\$

$$V(x) \equiv E\{(x - E\{x\})^2\} = E\{x^2\} - (E\{x\})^2, \\ E\{x^2\} = (E\{x\})^2 + V(x).$$

Аналогично для любого случайного вектора  $\mathbf{y}$  имеем

$$\mathbf{V}(\mathbf{y}) \equiv E\{\{(\mathbf{y} - E\{\mathbf{y}\})(\mathbf{y} - E\{\mathbf{y}\})^T\}\} = E\{\mathbf{y}\mathbf{y}^T\} - E\{\mathbf{y}\}E\{\mathbf{y}^T\}, \\ E\{\mathbf{y}\mathbf{y}^T\} = \mathbf{V}(\mathbf{y}) + E\{\mathbf{y}\}E\{\mathbf{y}^T\}, \quad (7)$$

где  $\mathbf{V}(\mathbf{y})$  – матрица ковариаций вектора  $\mathbf{y}$ .

Таким образом, у нас средняя внутригрупповая матрица ковариации имеет вид

$$E_{\mathbf{y}(1)}\{\mathbf{Vs}_{in}(\mathbf{y}(2))|\mathbf{y}(1)\} = E_{\mathbf{y}(1)}\left\{\frac{1}{2Ne(2)}\mathbf{W}(\mathbf{y}(1))|\mathbf{y}(0)\right\} = \\ \equiv E_{\mathbf{y}(1)}\left\{\frac{1}{2Ne(2)}(\mathbf{I} - \mathbf{y}(1)\mathbf{y}^T(1))|\mathbf{y}(0)\right\} = \\ = \frac{1}{2Ne(2)}\mathbf{I} - \frac{1}{2Ne(2)}E_{\mathbf{y}(1)}\{\mathbf{y}(1)\mathbf{y}^T(1)|\mathbf{y}(0)\}$$

Подставим сюда  $E_{\mathbf{y}(1)}\{\mathbf{y}(1)\mathbf{y}^T(1)\}$  согласно (7) и продолжим равенства

$$= \frac{1}{2Ne(2)}\mathbf{I} - \frac{1}{2Ne(2)}\mathbf{V}(\mathbf{y}(1)|\mathbf{y}(0)) + E\{\mathbf{y}(1)\}E\{\mathbf{y}^T(1)\} = \\ = \frac{1}{2Ne(2)}\mathbf{I} - \frac{1}{2Ne(2)}\left(\frac{1}{2Ne(1)}\mathbf{W}(\mathbf{y}(0)) + \mathbf{y}(0)\mathbf{y}^T(0)\right) \approx \\ \approx \frac{1}{2Ne(2)}(\mathbf{I} - \mathbf{y}(0)\mathbf{y}^T(0)) \equiv \frac{1}{2Ne(2)}\mathbf{W}(\mathbf{y}_0).$$

Здесь мы пренебрегли членом малой величины, содержащим произведение  $\frac{1}{Ne(1)Ne(2)}$ .

В итоге получаем

$$\mathbf{Vs}(\mathbf{y}(2)|\mathbf{y}(0)) = \frac{1}{2Ne(1)}\mathbf{W}(\mathbf{y}_0) + E_{\mathbf{y}(1)}\left\{\frac{1}{2Ne(2)}\mathbf{W}(\mathbf{y}_1)|\mathbf{y}(0)\right\} \\ \approx \left(\frac{1}{2Ne(1)} + \frac{1}{2Ne(2)}\right)\mathbf{W}(\mathbf{y}_0).$$

совокупности дискретных объектов (в частности фамилий), то нижеследующий результат сформулируем для произвольной совокупности (в случае фамилий в нем под  $\mathbf{V}$  подразумевается  $\mathbf{Vs}$ ). Дадим модификацию обоснования в [14] динамики матрицы ковариаций в этом случае.

**Результат 7.** Пусть рассматриваются последовательности независимых вложенных случайных выборок с возвращением из совокупности с  $k$  типами объектов (в нашем случае фамилий из популяции) и с такой группировкой объектов, когда минимальная концентрация среди групп превышает надлежащий порог. Положим, что размер выборки (эффективный размер популяции) на шаге  $\tau$  (в поколении  $\tau$ ) равен  $\frac{Ne(\tau)}{2}$ , а вероятность попадания в выборку объекта  $i$ -го типа равна  $x_i(\tau) > 0$ ,  $\sum_{i=1}^k x_i(\tau) = 1$ , где  $x_i$  – доля (концентрация) группы объектов  $i$ -го типа в “родительской” совокупности. Тогда на шаге  $t$  (в поколении  $t$ ) матрица ковариации  $\mathbf{V}(\mathbf{y}(t))$  случайного вектора  $\mathbf{y}(t) = \{y_i(t) = \sqrt{x_i(t)}\}$  при относительно малом  $t$  (при  $\frac{2\widetilde{N}e(t)}{t} \gg 1$ , где  $\widetilde{N}e(t)$  – среднее гармоническое значение для рассматриваемого ряда эффективных размеров популяции  $\{Ne(\tau), \tau = 1, 2, \dots, t\}$ ) приближенно выражается как

$$\mathbf{V}(\mathbf{y}(t)) = \left(\sum_{\tau=1}^t c(\tau)\right)(\mathbf{I} - \mathbf{y}(0)\mathbf{y}^T(0)) = \\ = \frac{t}{2\widetilde{N}e(t)}(\mathbf{I} - \mathbf{y}(0)\mathbf{y}^T(0)) = \frac{t}{2\widetilde{N}e(t)}\mathbf{W}(\mathbf{y}_0), \quad (8)$$

$$c(\tau) \equiv \frac{1}{2Ne(\tau)}, \quad \widetilde{N}e(t) \equiv t \left/ \sum_{\tau=1}^t \frac{1}{Ne(\tau)} \right..$$

Доказательство проведем по индукции. Мы подметили закономерность динамики  $\mathbf{V}(\mathbf{y}(t))$  в двух первых поколениях, согласно которой существует  $t$  ( $t = 2$ ), когда (8) выполняется для любой последовательности длиной не больше  $t$ . Покажем, что то же самое верно и при  $t + 1$ . Для этого рассмотрим ансамбль популяций в поколении  $t + 1$  как метапопуляцию, подразделенную на группы. Каждая группа порождается соответствующей популяцией первого поколения, т. е. состоит из возможных вариантов ее популяций-потомков в поколении  $t + 1$  с разными фамильными состояниями.

Как и при анализе двух поколений, ожидаемые концентрации фамилий в группах совпадают с фамильным состоянием порождающей популяции, принимающим значения  $\{y(1)\}$ , а дисперсия  $Vs(y(1))$  распределения концентраций отдельной фамилии по

популяциям первого поколения является межгрупповой дисперсией  $Vs_{betw}$  для ансамбля на любом другом из последующих поколений. Аналогично матрица ковариаций для распределения по популяциям первого поколения концентраций множества фамилий является межгрупповой матрицей ковариаций, которая в нашем случае имеет согласно (4)

$$\text{вид } Vs_{betw} = Vs(y(1)|y(0)) = \frac{1}{2Ne(1)} W(y_0).$$

Для отдельной группы в поколении  $t + 1$ , происходящей от популяции с фамильным состоянием  $y(1)$ , матрица ковариации  $Vs(y(t + 1)|y(1))$  является полной. По предположению индукции доказываемая формула верна в случае до  $t$  поколений, отделяющих ансамбль от порождающей популяции. Она верна как для  $Vs_{tot}(y(t)|y(0))$ , так и для  $Vs_{tot}(y(t + 1)|y(1))$ , поскольку для них число поколений, отделяющих  $y(t)$  от  $y(0)$  и  $y(t + 1)$  от  $y(1)$ , одно и то же (равно  $t$ ), т. е. матрица  $Vs_{tot}(y(t + 1)|y(1))$  находится согласно предположению индукции по формуле (8). В метапопуляции она представляет собой одну из *внутригрупповых матриц*. По правилу сложения ковариаций и с учетом предположения индукции применительно к матрице  $Vs_{tot}(y(t + 1)|y(1))$ , рассматриваемой как внутригрупповая, получаем

$$\begin{aligned} Vs_{tot}(y(t+1) | y(0)) &= Vs_{betw}(y(1) | y(0)) + \\ &+ E_{y(1)} \{Vs_{in}(y(t+1)) | y(1)\} = \\ &= \frac{1}{2Ne(1)} W(y_0) + \left( \sum_{\tau=2}^{t+1} c(\tau) \right) E_{y(1)} \{W(y(1)) | y(0)\}, \\ E_{y(1)} \{W(y(1)) | y(0)\} &\equiv E_{y(1)} \{(\mathbf{I} - y(1)y^T(1)) | y(0)\} = \\ &= \mathbf{I} - E_{y(1)} \{(y(1)y^T(1)) | y(0)\}. \end{aligned}$$

Подставим сюда  $E_{y(1)} \{y(1)y^T(1)\}$  согласно (7) и продолжим равенства

$$\begin{aligned} E_{y(1)} \{W(y(1)) | y(0)\} &= \mathbf{I} - (\mathbf{V}(y(1)|y(0)) + E \{y(1)\} E \{y^T(1)\}) \\ &= \mathbf{I} - (c(1)W(y_0) + y(0)y^T(0)) = (1 - c(1))W(y_0). \end{aligned}$$

Продлим преобразование  $Vs_{tot}(y(t + 1)|y(0))$  подстановкой найденного значения

$$E_{y(1)} \{W(y(1)) | y(0)\} :$$

$$\begin{aligned} Vs_{tot}(y(t+1) | y(0)) &= c(1)W(y_0) + \left( \sum_{\tau=2}^{t+1} c(\tau) \right) (1 - c(1))W(y_0) \approx \\ &\approx \left( \sum_{\tau=1}^{t+1} c(\tau) \right) W(y_0) = \frac{t}{2\widetilde{N}e(t)} W(y_0), \quad c(\tau) \equiv \frac{t}{2Ne(\tau)}. \end{aligned}$$

Здесь мы пренебрегли слагаемыми, содержащими произведение  $c(\tau)c(1) = \frac{1}{2Ne(\tau)} \frac{1}{2Ne(1)}$ , и учли, что по определению  $\widetilde{N}e(t) = t / \sum_{\tau=1}^t \frac{1}{Ne(\tau)}$  как обратная величина к среднему арифметическому для  $\left\{ \frac{1}{Ne(\tau)} \right\}$ .

Таким образом, выполняется переход индукции: если  $V(y((t)))$  зависит от  $t$  по предлагаемой формуле, то она верна при  $t + 1$  и, значит, при любом (относительно малом)  $t$ . ▶

**Следствие 8.** В условиях предыдущего результата при  $\frac{2\widetilde{N}e(t)}{t} \rightarrow \infty$ , где  $t$  – количество поколений, и при такой группировке фамилий, когда минимальная концентрация среди групп превышает надлежащий порог, асимптотическое распределение фамильного состояния  $y(t)$  является многомерным нормальным с математическим ожиданием  $y(0)$  и матрицей ковариации, пропорциональной идемпотентной матрице

$$\begin{aligned} W(y_0) &= \mathbf{I} - y(0)y^T(0); \\ \mathbf{y}(t) &= N \left( y(0), \frac{t}{2\widetilde{N}e(t)} (\mathbf{I} - y(0)y^T(0)) \right), \quad (\mathbf{I} - y(0)y^T(0))^2 = \\ &= (\mathbf{I} - y(0)y^T(0))^2 = \mathbf{I} - y(0)y^T(0). \end{aligned}$$

Распределение квадрата Евклидова расстояния  $\Delta(y(t))$  между  $y(t)$  и  $y(0)$ , т. е. величины  $\Delta^2(y(t)) = \|y(t) - y(0)\|^2 = \sum_{i=1}^k (y_i(t) - y_i(0))^2$ , удовлетворяет  $\frac{2\widetilde{N}e(t)}{t} \Delta^2(y(t)) \equiv \frac{2\widetilde{N}e(t)}{t} \|y(t) - y(0)\|^2 = \chi_{k-1}^2$ ,  $\frac{2\widetilde{N}e(t)}{t} \gg 1$ .

Для углового расстояния  $\theta s(y(t))$  между текущим фамильным состоянием  $\mathbf{x}(t)$  и начальным  $\mathbf{p}$  (между  $y(t)$  и  $y_0$  на гиперсфере), определяемого согласно (1) как:

$$\theta s(t) = \theta s(\mathbf{x}(t), \mathbf{p}) = \theta s(y(t), y_0) \equiv \arccos \left( \sum_{i=1}^k \sqrt{x_i(t)} \sqrt{p_i} \right),$$

асимптотически выполняется

$$\frac{2\widetilde{N}e(t)}{t} \theta s^2(\mathbf{x}(t), \mathbf{p}) = \chi_{k-1}^2$$

Здесь  $\widetilde{N}e(t)$  и  $\chi_{k-1}^2$  обозначают среднюю гармоническую численность популяции для ряда  $\{Ne(\tau)\}$  длиной  $t$  поколений и распределение хи-квадрат с  $k - 1$  степенями свободы соответственно.

**Доказательство** вытекает из центральной предельной теоремы, приближенного равенства Евклидова и углового расстояний в предположениях

следствия и ненаправленного характера случайного дрейфа. ▶

## ОБСУЖДЕНИЕ

Интерес автора к рассматриваемым проблемам мотивирован наблюдениями малых деревень России, где резко преобладала одна или несколько фамилий. Такая ситуация описывалась не раз в художественной литературе упоминаниями типа “у нас в деревне все Смирновы”. Картина, когда большинство жителей деревни оказываются однофамильцами, поражает. Например, при изучении автором популяций европейского севера России [28] встретилась деревня с 91% однофамильцев при общем количестве жителей 126 человек. Для городского жителя такая ситуация парадоксальна, и возникает желание дать теоретическое объяснение наблюдаемым различиям между фамильным и генетическим разнообразием. Поэтому настоящая статья фокусируется на теоретических аспектах анализа фамильной структуры.

Результаты анализа фамильной структуры популяции важны также своими параллелями с анализом генетической структуры в силу сходства патрилинейной передачи фамилии и генетической информации потомкам. В популяциях конечного размера такая передача сопровождается случайными флуктуациями как концентраций фамилий, так и аллелей. Напомним используемые нами подходы к изучению этой ситуации. При анализе флуктуаций мы используем предположение о не-перекрывании поколений, т. е. пренебрегаем существованием возрастной структуры у человека. Это предположение можно рассматривать как аппроксимацию реальной ситуации, часто применяемую в популяционной генетике, например при использовании закона Харди – Вайнберга.

При моделировании флуктуаций частот аллелей аутосомного локуса нередко используется модель Райта – Фишера, которую можно сформулировать в виде процесса вложенных выборок, формирующих состав нового поколения как случайную выборку аллелей поколения родителей. Та же самая картина получается при рассмотрении мужского компонента популяции, фамилии которого наследуются по поколениям при их патрилинейной передаче от отца к сыну. Данный процесс называем процессом случайного дрейфа (генов и фамилий одновременно в одной и той же популяции). Понятно, что качественные свойства случайного генного дрейфа и дрейфа фамилий одинаковы, но количественно различаются. Суть в том, что при генном дрейфе новое поколение с численностью  $N$  формируется как случайная выборка с возвращением  $2N$  гамет из пула родительских гамет, а состав мужского компонента нового поколения как выборка  $N/2$  фамилий отцов. Размер выборки

определяет интенсивность флуктуаций, которые больше в 4 раза для фамилий.

Понятно, что сходство процессов генного дрейфа и дрейфа фамилий означает возможность использования для анализа фамильного дрейфа методов, наработанных в популяционной генетике в течение длительного времени. В данной статье внимание концентрируется на стабилизации темпа дивергенции фамильного состояния популяции (вектора концентраций фамилий) от начального положения. Решение этой задачи опирается на известное нелинейное преобразование  $y_i = \sqrt{x_i}$  концентраций фамилий  $x_i$  (координат фамильного состояния) и анализ углового расстояния  $\theta$  между состояниями. Приведены модификации соответствующих популяционно-генетических подходов, расширены и углублены обоснования аппроксимации свойств случайного дрейфа на относительно небольших промежутках времени в поколениях.

В одной и той же популяции случайный дрейф приводит к фамильной дивергенции, вчетверо превосходящей дивергенцию генетическую на относительно небольшом промежутке времени в поколениях. Этот вывод характеризует с другой точки зрения результат о четырехкратном различии между стандартным коэффициентом инбридинга и его фамильным аналогом, полученный автором в [15] при анализе фамильного дрейфа в терминах концентраций фамилий  $\{x_i\}$ . Преимуществом изложенного подхода к анализу в пространстве  $\{y_i = \sqrt{x_i}\}$

является независимость дивергенции от начального состояния и постоянный темп ее увеличения за поколение в случае неизменяющегося эффективного размера популяции. Кроме того, дивергенция, отражаемая средним квадратом углового расстояния от начального положения, обратно пропорциональна среднему гармоническому эффективному размеру популяции и прямо пропорциональна количеству поколений дивергенции. Данный результат важен для сравнения популяций и их систем и для решения микротаксономических задач.

В заключение напомним основные условия корректности этих результатов на рассматриваемом промежутке времени:

отклонениями от патрилинейной передачи фамилии можно пренебречь;

размеры популяции должны быть достаточно велики, чтобы выборочные изменения фамильного состояния аппроксимировались нормальным распределением;

на рассматриваемом промежутке времени концентрации фамилий не должны быть слишком малыми во избежание нарушения свойств преобразования  $y_i = \sqrt{x_i}$ ;

сам рассматриваемый промежуток времени в поколениях должен быть небольшим по сравнению

со средним гармоническим эффективным разме-  
ром популяции на нем.

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

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

### СПИСОК ЛИТЕРАТУРЫ

1. *Lasker G.W. Surnames and Genetic Structure.* Cambridge Univ. Press, 2005. 148 p.
2. *King T.E., Jobling M.A. What's in a name Y chromosomes, surnames and the genetic genealogy revolution // Trends in Genetics.* 2009. V. 25. Iss. 8. P. 351–360.
3. *Jobling M.A. In the name of the father- surnames and genetics // Trends in Genetics.* 2001. V. 17. № 6. P. 353–357.
4. *Балановская Е.В., Балановский О.П. Русский гено- фонд на Русской равнине.* М.: Луч, 2007. 415 с.
5. *Сорокина И.Н., Чурносов М.И., Балтуцкая И.В. и др. Антропогенетическое изучение населения Центральной России.* М.: Изд-во РАМН, 2014. 336 с.
6. *Colantonio S.E., Lasker G.W., Kaplan B.A., Fuster V. Use of surname models in human population biology: A review of recent developments // Human Biology.* 2003. V. 75. № 6. P. 785–807.
7. *Crow J.F., Mange A.P. Measurement of inbreeding from the frequency of marriages between persons of the same surname // Social Biology.* 1982. V. 29. № 1/2. P. 101–105.
8. *Crow J.F. The estimation of inbreeding from isonymy // Human Biology.* 1980. V. 52. № 1. P. 1–14.
9. *Crow J.F. The estimation of inbreeding from isonymy (reprint) with an update // Human Biology.* 1989. V. 61. № 5/6. Special issue on foundations of anthropological genetics. P. 935–948.
10. *Rogers A.R. Doubts about isonymy // Human Biology.* 1991. V. 63. № 5. P. 663–668.
11. *Ли Ч. Введение в популяционную генетику.* М.: Мир, 1978. 555 с. (*Li C.C. First course in population genetics.* California: Boxwood Press Pacific Grove, 1976).
12. *Кимура М. Молекулярная эволюция: теория ней-  
тральности.* М.: Мир. 1985. 394 с. (*Kimura M. The Neutral Theory of Molecular Evolution.* Cambridge: Cambr. Univ. Press., 1983)
13. *Хедрик Ф. Генетика популяций.* М.: Техносфера. 2003. 592 с. (*Hedrick P.W. Genetics of Populations.* 2nded. Boston: Jones and Bartlett Publ., 2000. 553 pp.)
14. *Малютов М.Б., Пасеков В.П. Об одной статисти-  
ческой задаче популяционной генетики // Теория вероятностей и ее применения.* 1971. Т. 16. Вып. 3. С. 579–581. (*Mal'yutov M.B., Pasekov V.P. On one statistical problem of population genetics // Theory of Probability and its Applications.* 1971. Iss. V. 16. № 3. P. 559–566)
15. *Пасеков В.П. К анализу случайных процессов изо-  
нимии. I. Структура изонимии // Генетика.* 2021. Т. 57. № 10. С. 1194–1204.  
doi: 10.31857/S001667582110009X (*Passekov V.P. To the Analysis of Random Processes of Isonymy: I. Isonymic Structure // Rus. J. Genet.* 2021. V. 57. № 10, P. 1214–1222.  
doi: 10.1134/S1022795421100094)
16. *Fisher R.A. On the dominance ratio // Proc. R. Soc. Edinb.* 1922. V. 42. P. 321–341 (Bull. Math. Biol. 1990. V. 52. № 1–2. P. 297–318)
17. *Fisher R.A. The Genetical Theory of Natural Selection.* Oxford: Clarendon Press, 1930. 272 p.
18. *Bhattacharyya A. On a measure of divergence between two multinomial populations // Sankhya.* 1946. V. 7. Part 4. P. 401–406.
19. *Edwards A.W.F. Distances between populations on the basis of gene frequencies // Biometrics.* 1971. V. 27. № 4. P. 873–881.
20. *Вейр Б. Анализ генетических данных: дискрет-  
ные генетические признаки.* М.: Мир, 1995. 400 с. (*Weir B.S. Genetic data analysis: Methods for discrete population genetic data.* Sunderland: Sinauer, 1990.)
21. *Cavalli-Sforza L.L., Edwards A.W.F. Phylogenetic analysis. Models and estimation procedures // Am. J. Hum. Genet.* 1967. V. 19. P. 233–257 (Evolution. 1967. V. 21. № 3. P. 550–570).
22. *Свирежев Ю.М., Пасеков В.П. Основы математи-  
ческой генетики.* М.: Наука, 1982. 511 с. (*Svirezhev Y.M., Passekov V.P. Fundamentals of mathematical evolutionary genetics.* Kluwer Acad. Publ., Dordrecht et al., 1990. 395 p.)
23. *Antonelli P.L., Strobeck C. The geometry of random drift. I. Stochastic distance and diffusion // Adv. Appl. Probab.* 1977. V. 9. № 2. P. 238–249.
24. *Papangelou F. The large deviations of a multi-allele Wright–Fisher process mapped on the sphere // Ann. Appl. Prob.* 2000. V. 10. № 4. P. 1259–1273.
25. *Молчанов С.А. Диффузионные процессы и риманова геометрия // УМН.* 1975. Т. 30. Вып. 1(181). С. 3–59. (*Molchanov S.A. Diffusion processes and Riemannian geometry // Russ. Math. Surveys.* 1975. V. 30. Iss. 1. P. 1–63)
26. *Hofrichter J., Jost J., Tran T.D. Information geometry and population genetics: The mathematical structure of the Wright–Fisher model.* Springer, 2017. 320 p.
27. *Пасеков В.П. Описание дивергенции субпопу-  
ляций в иерархической системе при анализе изонимии. I. Дисперсия как показатель дивер-  
генции // Генетика.* 2022. Т. 58. № 6. С. 713–727  
doi: 10.31857/S0016675822060054 (*Passekov V.P. Description of Divergence of Subpopulations in the*

- Hierarchical System When Analyzing Isonymy: I. 28. *Пасеков В.П., Ревазов А.А.* К популяционной генетике населения европейского севера СССР. Сообщение I. Данные по структуре шести деревень Архангельской области // Генетика, Т. 11. № 7. 1975. С. 145–155.  
doi: 10.1134/S1022795422060059)

## On Stabilizing the Rate of Isonomy Divergence

V. P. Passekov<sup>1,\*</sup>

<sup>1</sup>*Federal Research Center “Computer Science and Control”, Russian Academy of Sciences, Moscow 119991 Russia*  
*\*e-mail: pass40@mail.ru*

A theoretical analysis of the surname state of the population (the vector of namesake concentrations in the male component of the population) and its dynamics as a result of random surname drift is presented. An approximation of such a process by the Wright-Fisher model of a population with non-overlapping generations without selection pressure is used, i.e., an approximation by a sequence of nested random samples with the replacement from fathers' surnames in the population. The sample size is  $N/2$  according to the size of the male component in the population of size  $N$ . In the same population, processes of random drift of both surnames and genes simultaneously occur. Their cardinal difference is that the sample size of surnames is four times smaller than the sample size of autosomal locus alleles. The analysis of random drift is simplified when moving from concentration coordinates to the square roots of them. As generations change, the state receives a sample deviation, measured by angular distance, and its mean square gives the rate of divergence, stabilizing in the new coordinates. An adaptation (in relation to the analysis of surname drift) of a known in population genetics result about the nature of divergence at a stage of a relatively small number of generations compared to the size of the population is given. The divergence of surnames occurs four times faster than the divergence of allele concentrations.

**Keywords:** random surname drift, divergence of surname and allele concentrations, isonymy, angular distances, stabilization of the divergence rate.

---

## КРАТКИЕ СООБЩЕНИЯ

---

УДК 575.18:598.241.2

# ПАРИТЕТНОЕ СООТНОШЕНИЕ САМЦОВ И САМОК В ВЫВОДКАХ КРАСАВКИ *Anthropoides virgo* В КАЛМЫКИИ И БУРЯТИИ

© 2024 г. Е. А. Мудрик<sup>1, 2\*</sup>, Е. И. Ильяшенко<sup>1, 3</sup>, К. Д. Кондракова<sup>1, 3</sup>,  
А. А. Абушин<sup>4</sup>, Л. Д. Базаров<sup>5</sup>, Ц. З. Доржиев<sup>6, 7</sup>, Д. В. Политов<sup>1, 2</sup>

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

<sup>2</sup>Всероссийский научно-исследовательский институт охраны окружающей среды, Москва, 117628 Россия

<sup>3</sup>Институт проблем экологии и эволюции им. А. Н. Северцова Российской академии наук, Москва, 119011 Россия

<sup>4</sup>Государственный природный биосферный заповедник "Черные земли", Республика Калмыкия,  
п. Комсомольский, 359240 Россия

<sup>5</sup>Национальный парк "Тункинский", Республика Бурятия, с. Кырен, 671010 Россия

<sup>6</sup>Бурятский государственный университет им. Д. Банзарова, Улан-Удэ, 670000 Россия

<sup>7</sup>Институт общей и экспериментальной биологии Сибирского отделения Российской академии наук,  
Улан-Удэ, 670000 Россия

\*e-mail: mudrik@vigg.ru

Поступила в редакцию 26.06.2024 г.

После доработки 09.07.2024 г.

Принята к публикации 15.07.2024 г.

Анализировали соотношение полов в потомстве журавля красавки (*Anthropoides virgo* Linneaus, 1758) в двух удаленных друг от друга гнездовых группировках европейской (Республика Калмыкия) и азиатской (Республика Бурятия) частей ареала в 2023 г. С применением молекулярно-генетического маркера пола EE0.6 определен пол у 46 птенцов (24 самца и 22 самки) из 30 выводков. В суммарной выборке различия в пропорции самок и самцов не были значимы (0.521,  $P = 0.092$ ). Соотношение полов в каждой из изученных гнездовых группировок также статистически не отличалось от паритетного, хотя отмечено преобладание в полных выводках и в целом в потомстве красавки в Калмыкии самцов (0.650,  $P = 0.523$ ), а в Бурятии – самок (0.423,  $P = 0.781$ ). Полученные данные свидетельствуют в пользу стратегии воспроизводства у красавки равного количества потомков обоего пола и могут косвенно указывать на то, что у этого вида энергетические затраты на выращивание самцов и самок равны, а воздействие условий среды не настолько жесткое, чтобы существующий баланс изменялся в пользу какого-либо из полов.

**Ключевые слова:** журавль, соотношение полов, молекулярно-генетическое определение пола, птенец, EE0.6.

**DOI:** 10.31857/S00166758241200114 **EDN:** VZVCAL

Соотношение полов в популяциях животных является важной характеристикой репродуктивной биологии вида. В норме долгоживущие моногамные виды, у которых заботу о потомстве осуществляют оба родителя, воспроизводят равное количество самцов и самок [1]. Ранее мы показали, что у журавля красавки *Anthropoides virgo* Linneaus, 1758 (Aves: Gruiformes: Gruidae), гнездовая часть ареала которого охватывает степные и полупустынные зоны Евразии, соотношение полов в выводках некоторых европейских (азово-черноморской, прикаспийской, волго-уральской) и азиатских (зуральской, алтайской, хакасской и забайкальской) гнездовых группировок на территории Российской Федерации в отдельные годы исследований (2017–2019 гг.) достоверно не отличалось от равного,

хотя в большинстве случаев демонстрировало тенденцию к преобладанию самок [2, 3]. Вместе с тем изменения условий обитания, отражающиеся на доступности пищевых ресурсов, могут провоцировать в разных популяциях одного вида и в одних и тех же популяциях в разные репродуктивные сезоны отклонения в соотношении полов в потомстве. У многих видов животных в неблагоприятные годы такие сдвиги происходят в пользу воспроизводства менее ресурсопотребляющего и уязвимого пола [4], которым у журавлей являются самки.

Красавка в России относится к видам с сокращающейся численностью вследствие ухудшения условий гнездования, гибели от отравлений сельскохозяйственными ядохимикатами и

**Таблица 1.** Соотношение полов птенцов в полных и неполных выводках в Калмыкии и Бурятии в 2023 г.

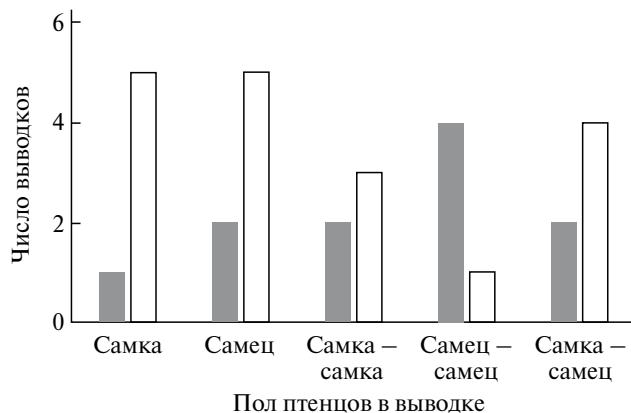
Количество и пол изученных птенцов	Калмыкия	Бурятия
Всего птенцов	20	26
самцы	13	11
самки	7	15
соотношение полов	0.650, $P = 0.523$	0.423, $P = 0.781$
Выводки с одним птенцом	3	10
самцы	2	5
самки	1	5
соотношение полов	0.666, $P = 1$	0.5, $P = 1$
Выводки с двумя птенцами	8	8
самцы	10	6
самки	6	10
соотношение полов	0.625, $P = 0.722$	0.375, $P = 0.722$

браконьерства на путях миграции [5, 6]. Изучение вопросов репродуктивной биологии и воспроизведения в разных гнездовых группировках красавки актуально для мониторинга ее состояния, особенно на границе ареала. У птенцов журавлей нет полового диморфизма, поэтому определение пола для них осуществляют молекулярно-генетическими методами. В данном сообщении мы приводим новые сведения по соотношению полов в выводках красавки в двух удаленных друг от друга гнездовых группировках из европейской (Северный Прикаспий, Республика Калмыкия) и азиатской (Юго-Западное Забайкалье и Тункинская котловина, Республика Бурятия) частей ареала в 2023 г.

В работе использовали биологический материал (растущие перья из оперения груди или шеи), полученный в результате собственных экспедиций в Калмыкии (Яшкульский, Черноземельский, Юстинский, Кетченеровский районы) [7] и Бурятии (Иволгинский, Селенгинский, Мухоршибирский, Джидинский, Тункинский районы) [8] в июне – июле 2023 г. Размер выборок составил 20 и 26 птенцов соответственно (табл. 1). У красавки – выводковый тип онтогенетического развития, при котором птенцы покидают гнездо практически сразу после вылупления. В норме кладка состоит из двух яиц. Отлов птенцов, еще не способных к полету, проводили в соответствии с разрешениями Федеральной службы по надзору в сфере природопользования (Росприроднадзор) № 21 от 15.02.2023 г. и № 63 от 02.05.2023 г. ручным способом. После забора биоматериала, занимающего 5–10 мин, птенцов отпускали к родителям. Растущие перья помещали в пробирки 1.5 мл типа Эпендорф с раствором Лонгмайра, транспортировали в лабораторию при комнатной температуре,

а затем хранили в морозильной камере при  $-20^{\circ}\text{C}$ . Выделение ДНК из растущих перьев осуществляли набором “К-сорб” (НПК Синтол, Россия) по протоколу производителя. В качестве молекулярно-генетического маркера пола использовали последовательность W-хромосомы EE0.6. Для ПЦР-амплификации применяли комбинацию праймеров AWS05/NRD4 [9] и фрагментный анализ ампликонов по описанной ранее [10, 11] и хорошо зарекомендовавшей себя при диагностике пола у красавки [2, 3] методике. Соотношение полов определяли общепринятым способом как долю самцов по отношению к общему количеству птенцов [12]. Для сравнения выборок самцов и самок использовали точный тест Фишера для выборок малого размера, рассчитанный в базовом пакете *stats* среди R версии 3.6.3 [13].

Всего было обнаружено 12 и 18 пар красавок с птенцами в Калмыкии и Бурятии соответственно. Из 12 выводков в Калмыкии на момент встречи девять были полными (представлены двумя птенцами) и три – неполными (одним птенцом). В одном полном выводке не удалось взять пробу от одного из сибсов, поэтому его анализировали только в общем соотношении полов в группировке и исключили из анализа соотношения полов в полных выводках. Среди 18 выводков в Бурятии неполными были 10, полными – 8. Таким образом, в каждой гнездовой группировке мы анализировали одинаковое количество полных выводков (по 16 птенцов). В Калмыкии из восьми полных выводков было два разнополых, четыре представлены только самцами, два – только самками. В Бурятии из восьми полных выводков было четыре разнополых, в одном – только самцы, в трех – только самки (рис. 1).



**Рис. 1.** Количество самцов и самок в выводках красавки в Калмыкии (закрашенные столбцы) и Бурятии (незакрашенные столбцы) в 2023 г.

В общей выборке из 46 птенцов двух гнездовых группировок идентифицировано 24 самца и 22 самки; показатель соотношения полов в ней составил 0.521,  $P = 0.092$ . В Калмыкии определено 13 самцов и 7 самок, в Бурятии – 11 самцов и 15 самок. Во всех выводках различия в соотношении полов не были значимы, несмотря на тенденцию к преобладанию в полных выводках и в целом в потомстве красавки в Калмыкии самцов, а в Бурятии – самок (табл. 1, рис. 1). Соотношение полов среди всех птенцов красавки в Калмыкии и Бурятии достоверно не отличалось от равного (0.5) и составило 0.650,  $P = 0.523$  и 0.423,  $P = 0.781$  соответственно.

Представленные данные дополняют немногочисленные сведения о размножении красавки и журавлей в целом. В природе соотношение полов в потомстве журавлей практически не изучено, а исследования воспроизведения некоторых видов при вольерном разведении показывают, что этот показатель не всегда находится в равновесии. Так, искусственно созданные условия могут провоцировать у журавлей стрессовые реакции, сопровождающиеся воспроизведением преимущественно самок (например у японского журавля *Grus japonensis*) [14, 15], но могут и не оказывать такого действия (например у стерха *Leucogeranus leucogeranus*) [15, 16]. Информация по полу птенцов в потомстве красавки в Юго-Западном Забайкалье получена нами впервые. В прикаспийской гнездовой группировке в 2017 г. при анализе более представительной выборки (61 птенец из Калмыкии, Дагестана, Астраханской и Волгоградской областей) мы выявили равное соотношение полов с тенденцией к преимуществу самок в полных выводках и в целом в группировке [2], при этом непосредственно в Калмыкии из 32 птенцов 16 были определены как самцы и 16 как самки. Таким образом, перечисленные факты свидетельствуют в пользу эволюционно предпочтительной для красавки стратегии

паритетного воспроизводства потомков обоего пола [3] и, возможно, косвенно указывают на то, что у этого вида энергетические затраты родителей на выращивание самцов и самок равны, а воздействие условий среды не настолько жесткое, чтобы существующий баланс изменялся в пользу какого-либо из полов.

Авторы благодарят В. Ю. Ильяшенко, А. Ю. Скрипниченко и Е. Н. Бадмаеву за помощь в проведении полевых работ.

Исследование выполнено за счет гранта Российского научного фонда № 23-24-00613, <https://rscf.ru/project/23-24-00613/>.

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных соблюdenы. Исследование одобрено Локальным комитетом по биоэтике при Федеральном государственном бюджетном учреждении науки Институт общей генетики им. Н.И. Вавилова Российской академии наук (протокол № 1 от 18.05.2023 г.).

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

- Donald P.F. Adult sex ratios in wild bird populations // *Ibis*. 2007. V. 149. P. 671–692.
- Мудрик Е.А., Ильяшенко Е.И., Джамирзоев Г.С. и др. Соотношение полов у птенцов красавок (*Anthropoides virgo* Linnaeus, 1758) прикаспийской гнездовой группировки // Генетика. 2018. Т. 54. Приложение. С. S54–S57. <https://doi.org/10.1134/S0016675818130143>
- Мудрик Е.А., Ильяшенко Е.И., Постельных К.А. и др. Соотношение полов в потомстве моногамного вида птиц – красавки *Anthropoides virgo* // Генетика. 2022. Т. 58. № 11. С. 1273–1279. <https://doi.org/10.31857/S0016675822110078> [Mudrik E.A., Ilyashenko E.I., Postelnykh K.A. et al. Sex ratio in the offspring of monogamous bird species (Demoiselle crane *Anthropoides virgo*) // Rus. J. Genet. 2022. V. 58. № 11. P. 1317–1322. [https://doi.org/10.1134/S1022795422110072\]](https://doi.org/10.1134/S1022795422110072)
- Trivers R.L., Willard D.E. Natural selection of parental ability to vary the sex ratio of offspring // Science. 1973. V. 179. № 4068. P. 90–92. <https://doi.org/10.1126/science.179.4068.90>
- Ильяшенко Е.И. Журавль-красавка *Anthropoides virgo* (Linnaeus, 1758) / Красная книга Российской Федерации. Животные, 2-ое изд. М.: ФГБУ “ВНИИ Экология”, 2021 г. С. 689–691.
- Ильяшенко Е.И., Ильяшенко В.Ю. Угрозы популяциям серых журавлей *Grus grus* и красавок *Anthropoides virgo* // Рус. орнитол. журн. (Экспресс-выпуск 2354). С. 4662–4663.

7. Абушин А.А., Ильяшенко В.Ю., Кондракова К.Д., Ильяшенко Е.И. Обследование мест обитания красавки в Республике Калмыкия в 2023 г. // Инф. бюллетень Рабочей группы по журавлям Евразии им. В.Е. Флинта. 2024 г. № 18. С. 18–23.
8. Ильяшенко Е.И., Доржсэев Ц.З., Кондракова К.Д. и др. Обследование мест обитания журавлей в Предбайкалье и Юго-Западном Забайкалье в 2023 г. // Инф. бюллетень Рабочей группы по журавлям Евразии им. В.Е. Флинта 2024 г. № 18. С. 24–34.
9. Bao W.B., Wu S.L., Zhang H.X. Sex identification of seven species of cranes in China by PCR // J. Anim. Vet. Adv. 2009. V. 8. № 6. P. 1137–1140.
10. Мудрик Е.А., Кащенцева Т.А., Гамбург Е.А., Политов Д.В. Определение пола у десяти видов журавлей с помощью ДНК-маркера EE0.6 // Генетика. 2013. Т. 49. № 12. С. 1436–1439. <https://doi.org/10.7868/S0016675813120060>
11. Мудрик Е.А., Кащенцева Т.А., Гамбург Е.А. и др. Неинвазивный метод идентификации пола птенцов журавлей по ДНК из капиллярных сосудов аллантоиса // Онтогенез. 2013. Т. 44. № 5. С. 372–376. <https://doi.org/10.7868/S0475145013050054> [Mudrik E.A., Kashentseva T.A., Gamburg E.A. et al. Non-invasive method of sex identification of crane chicks by the DNA from capillary vessels of allantois // Rus. J. Dev. Biol. 2013. V. 44. № 5. P. 279–282. <https://doi.org/10.1134/S1062360413050056>
12. Wilson K., Hardy I.C. Statistical analysis of sex ratios: An introduction // Sex Ratios: Concepts and Research Methods. London: Cambridge Univ. Press, 2002. P. 48–92. <https://doi.org/10.1017/CBO9780511542053.004>
13. Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria: 2020.
14. Нестеренко О.Н., Кащенцева Т.А. Преобладание самок в потомстве японских журавлей и его возможный адаптивный характер // Тр. Окского заповедника. Вып. 34. Рязань: 2015. С. 250–254.
15. Нестеренко О.Н., Кащенцева Т.А. Соотношение полов в потомстве японских журавлей и стерхов в Питомнике Окского заповедника, Россия // Журавли Евразии (распространение, биология). 2021. Вып. 6. М.: Тов-во науч. изданий КМК, С. 523–533.
16. Мудрик Е.А., Кащенцева Т.А., Постельных К.А. и др. Соотношение полов в потомстве искусственной популяции стерха (*Grus leucogeranus* Pallas) // Генетика. 2015. Т. 51. № 12. С. 1439–1443. <http://doi.org/10.7868/S0016675815120085>

## Parity Ratio of Males and Females in Broods of the Demoiselle Crane *Anthropoides virgo* in Kalmykia and Buryatia

**E. A. Mudrik<sup>a, b\*</sup>, E. I. Ilyashenko<sup>a, c</sup>, K. D. Kondrakova<sup>a, c</sup>, A. A. Abushin<sup>d</sup>,  
L. D. Bazarov<sup>e</sup>, Ts. Z. Dorzhiev<sup>f, g</sup>, D. V. Politov<sup>a, b</sup>**

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

<sup>b</sup>*All-Russian Research Institute for Environmental Protection, Moscow, 117628 Russia*

<sup>c</sup>*Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, 119011 Russia*

<sup>d</sup>*State Nature Biosphere Reserve "Chernye Zemli", Republic of Kalmykia, Komsomolsky, 359240 Russia*

<sup>e</sup>*Tunkinsky National Park, Republic of Buryatia, Kyren, 671010 Russia*

<sup>f</sup>*Buryat State University named after D. Banzarov, Ulan-Ude, 670000 Russia*

<sup>g</sup>*Institute of General and Experimental Biology, Siberian Branch of the Russian Academy of Sciences,  
Ulan-Ude, 670000 Russia*

*\*e-mail: mudrik@vigg.ru*

The offspring sex ratio of the Demoiselle crane (*Anthropoides virgo* Linneaus, 1758) was analyzed in two remote breeding groups of the European (Republic of Kalmykia) and Asian (Republic of Buryatia) parts of the range in 2023. Using the molecular genetic marker EE0.6, the sex was determined in 46 chicks (24 males and 22 females) from 30 broods. In the total sample, the deviations from equal proportion of females and males were not significant (0.521,  $P = 0.092$ ). The sex ratio in each of the studied breeding groups also did not statistically differ from the parity, although males (0.650,  $P = 0.523$ ) prevailed in two-chick broods and in general in the offspring of the Demoiselle crane in Kalmykia (0.423,  $P = 0.781$ ), while females prevailed in two-chick broods and in general in the offspring of the Demoiselle crane in Buryatia (0.423,  $P = 0.781$ ). The obtained data support the strategy of reproduction of parity number of both sexes in the Demoiselle crane and may indirectly indicate that in this species the energy cost on raising males and females is equal, and the pressure of environmental conditions is not so strong to cause bias of the existing balance towards any sex.

**Keywords:** crane, sex ratio, molecular genetic sex determination, chick, EE0.6.

---

## КРАТКИЕ СООБЩЕНИЯ

---

УДК 575.162:575.167:57.024

# ВЛИЯНИЕ ГЕНОТИПА ПО ГЕНУ АЛКОГОЛЬДЕГИДРОГЕНАЗЫ *ADH1B* НА УРОВЕНЬ БИОМАРКЕРА ЗЛОУПОТРЕБЛЕНИЯ АЛКОГОЛЕМ СДТ

© 2024 г. А. А. Ким<sup>1</sup>, А. С. Гуреев<sup>1</sup>, А. В. Рубанович<sup>1,\*</sup>, С. А. Боринская<sup>1,\*\*</sup>

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

\*e-mail: rubanovich@vigg.ru

\*\*e-mail: borinskaya@vigg.ru

Поступила в редакцию 30.07.2024 г.

После доработки 20.08.2024 г.

Принята к публикации 30.08.2024 г.

Проблема злоупотребления алкоголем актуальна для общественного здравоохранения и связана с сокращением продолжительности жизни, многочисленными заболеваниями, социальными и экономическими проблемами. Методы выявления индивидов, злоупотребляющих алкоголем, необходимы в клинической практике для определения стратегий профилактики и лечения и вос требованы в судебно-медицинской экспертизе. Получение сведений путем опроса часто ведет к снижению доз потребляемого алкоголя, поэтому необходимы объективные лабораторные методы диагностики хронического злоупотребления алкоголем и острых алкогольных эксцессов. Одним из наиболее специфичных биомаркеров злоупотребления алкоголем является углевод-дефицитный трансферрин (СДТ), представляющий собой изоформы трансферрина с пониженным количеством сиаловых остатков. СДТ образуется за счет нарушения гликозилирования продуктами метаболизма этанола. Основной путь окисления экзогенного этанола – преобразование его в ацетальдегид под действием фермента алкогольдегидрогеназы *ADH1B*. В данном исследовании впервые показано, что у носителей аллеля *ADH1B\*48His* (rs1229984) гена алкогольдегидрогеназы, ведущего к образованию ацетальдегида в более высоких концентрациях, повышен уровень СДТ по сравнению с индивидами, не имеющими этого аллеля, при одинаковом уровне потребления алкоголя. Этот эффект необходимо учитывать в клинической и судебно-медицинской практике, а также в исследованиях эффектов потребления алкоголя с применением менделевской рандомизации по полиморфному локусу rs1229984.

**Ключевые слова:** биомаркеры злоупотребления алкоголем, углевод-дефицитный трансферрин, СДТ, алкогольдегидрогеназа, ген *ADH1B*.

**DOI:** 10.31857/S0016675824120125 **EDN:** VZRIDW

Злоупотребление алкоголем порождает многочисленные социальные и экономические проблемы и представляет серьезную проблему общественного здравоохранения. В России это одна из основных причин снижения продолжительности жизни, особенно заметного у мужчин трудоспособного возраста [1, 2].

Определение индивидуального уровня потребления алкоголя необходимо при лечении алкогольной зависимости, при решении о возврате прав, изъятых за вождение в нетрезвом виде, перед некоторыми хирургическими операциями и в ряде других медицинских вмешательств, а также востребовано судебной медициной и криминалистической экспертизой. Уровень потребления, определенный путем опроса респондентов, не является достоверным. Показано, что при опросе респонденты занижают количество потребляемого

алкоголя более чем в два раза [3, 4]. Объективные данные об уровне потребления могут быть получены при использовании биомаркеров, отражающих количество алкоголя, потребленного за тот или иной период. К прямым биомаркерам относится содержание этанола в крови, выдыхаемом воздухе и моче, а также продукты метаболизма алкоголя, такие как ацетальдегид (составляет 95–98% среди метаболитов этанола), этилглюкуронид (< 0.2%), этилсульфат (< 0.1%), фосфатидилэтанол и другие. Однако время, в течение которого можно обнаружить их следы в биологических жидкостях, составляет немногим более суток после потребления умеренных доз алкоголя, и только фосфатидилэтанол может быть обнаружен на протяжении более длительного времени [5]. Наиболее часто используемые непрямые биомаркеры злоупотребления алкоголем – ферменты печени

аланин-аминотрансфераза (ALT), аспарагин-аминотрансфераза (AST), гамма-глутамил трансфераза (GGT, более специфичен, чем первые два), средний объем эритроцитов (MCV) и углевод-дефицитный трансферрин (CDT) [5]. Указанные биомаркеры имеют повышенный уровень при злоупотреблении алкоголем, который сохраняется на протяжении недель. Но их повышение может быть вызвано и иными причинами, такими как панкреатит, холецистит, цирроз печени, вирусные гепатиты, анемия и ряд других заболеваний, а также прием некоторых лекарственных препаратов.

Из перечисленных биомаркеров наибольшей чувствительностью и специфичностью обладает биомаркер CDT, определяемый стандартно в сыворотке крови [6]. Трансферрин переносит железо и является гликопротеидом, имеющим два сайта N-гликозилирования, олигосахариды в которых несут в преобладающей изоформе трансферрина 4 сиаловых остатка. Злоупотребление алкоголем приводит к нарушению гликозилирования и снижению количества сиаловых остатков на трансферрине с образованием углевод-дефицитных изоформ (CDT). CDT был случайно открыт в 1976 г. нейрологом Х. Стиблер при обследовании пациентов, часть из которых злоупотребляла алкоголем и имела измененную структуру трансферрина — меньшее количество сиаловых остатков [7]. В норме в сыворотке крови преобладает тетрасиалотрансферрин (~80%) в сочетании с другими изоформами — пентасиало- (~14%), трисиало- (~4%), дисиало- (~1%) и гексасиалотрансферрин (~1%) [8].

Как было показано, при потреблении алкоголя на уровне 40–60 г в день в течение 1–2 недель происходит нарушение гликозилирования, и трансферрин теряет сиализацию, образуя углевод-дефицитные изоформы с двумя сиаловыми остатками, реже с другим количеством вплоть до полного отсутствия (асиалотрансферрин) [5, 9, 10]. Наружение сиализации трансферрина связано с ингибированием ферментов, участвующих в гликозилировании этанолом и продуктами его метаболизма, в первую очередь ацетальдегидом [11, 12]. Повышение уровня CDT, согласно ряду работ, выше 1.3% указывает на злоупотребление алкоголем [6, 13]. Однако рабочая группа по стандартизации измерения CDT при Международной федерации клинической химии и лабораторной медицины (IFCC WG-CDT) указывает, что верхней границей нормы следует считать 1.7% CDT [10].

Экзогенный этанол метаболизируется в основном алкогольдегидрогеназой ADH1B. Замена нуклеотида в полиморфном локусе rs1229984 гена ADH1B соответствует двум аллелям, определяющим различия в активности фермента и скорости образования ацетальдегида. Аллель ADH1B\*48Arg (rs1229984-G) преобладает у индивидов

европейского происхождения, тогда как аллель ADH1B\*48His (rs1229984-A), определяющий более высокую активность фермента и образование ацетальдегида в более высоких концентрациях, распространен в Восточной и Юго-Восточной Азии [14]. Частота аллеля ADH1B\*48His у русских составляет 5–7%, а частота носителей этого аллеля, соответственно, 10–12% [14]. Мы предположили, что более высокие концентрации ацетальдегида у носителей аллеля ADH1B\*48His могут влиять на уровень образования CDT, модулируя влияние потребляемого этанола.

В настоящей статье мы проверили эту гипотезу и подтвердили ее, сравнив уровень CDT у русских мужчин — носителей разных аллелей, с известным уровнем потребления алкоголя. Носители аллеля ADH1B\*48His имеют повышенный уровень CDT по сравнению с теми, кто этого аллеля не имеет, при одинаковом уровне потребления алкоголя. Такой эффект показан впервые.

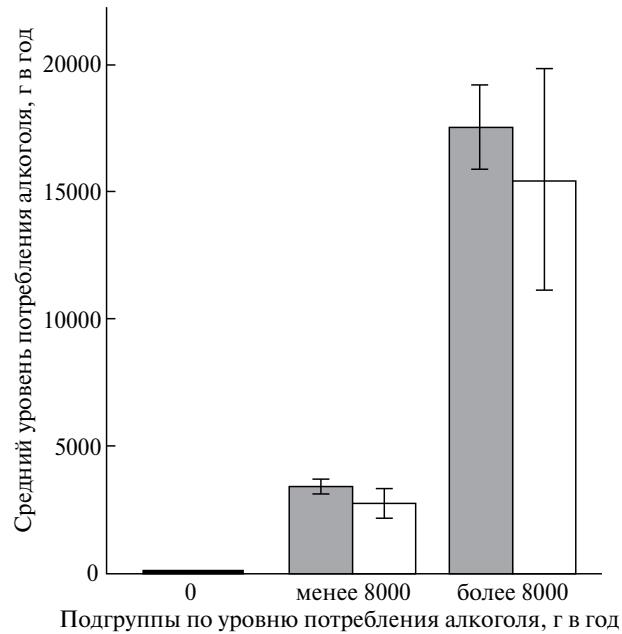
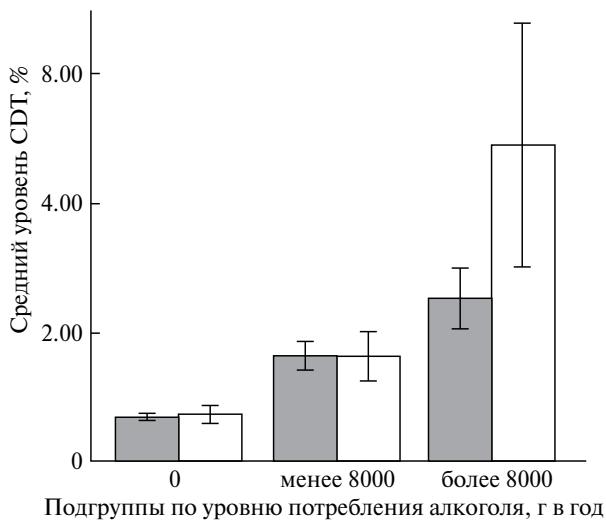
Материалом для анализа послужили данные Ижевского исследования семей [1, 2] об уровне CDT у 642 русских мужчин [15] в возрасте от 22 до 59 лет с известными генотипами по полиморфизму rs1229984 гена ADH1B [16, 17]. Статистические вычисления и визуализация данных проводились с использованием пакета программ SPSS v.19. Для сравнений различий количественных признаков в подгруппах использовалась непараметрическая статистика Манна — Уитни. Разброс данных на рисунках охарактеризован удвоенной ошибкой среднего ( $\pm$  SE).

Средний уровень CDT равномерно возрастает в подгруппах с увеличивающимся уровнем потребления алкоголя. Корреляция по Пирсону для выборки в целом составила 0.298 при  $p < 0.0001$ . Для носителей аллеля ADH1B\*48His эта корреляция была существенно выше: 0.601 при  $p < 0.0001$ . В исследованной выборке (642 русских мужчин) 68 человек (10.6%) являются гетерозиготными носителями аллеля ADH1B\*48His. Мужчины-носители аллеля ADH1B\*48His потребляют алкоголь в среднем на 17–20% меньше, чем обладатели генотипа ADH1B\*Arg/Arg (из-за неприятных симптомов, связанных с повышенным уровнем ацетальдегида после приема алкоголя) [16]. Для оценки уровня CDT в подгруппах с разным уровнем потребления выборка была разделена на три подгруппы — трезвенники (не потребляли алкоголь в течение года перед исследованием), потреблявшие менее 8000 г в год в пересчете на этанол (эквивалентно потреблению одной бутылки водки в неделю) и потреблявшие более 8000 г этанола в год (табл. 1).

На рис. 1 представлен уровень потребления алкоголя обладателями разных генотипов по гену ADH1B в изученной выборке, на рис. 2 — уровень CDT в тех же подгруппах. Носители аллеля

**Табл. 1.** Распределение мужчин в подгруппах по уровню потребления алкоголя

Уровень потребления алкоголя, г в год	Генотип по <i>ADH1B</i> (%)		Всего
	<i>Arg/Arg</i>	<i>Arg/His</i>	
0	76 (91.6)	7 (8.4)	83
Менее 8000	329 (88.4)	43 (11.6)	372
Более 8000	169 (90.4)	18 (9.6)	187
Всего:	574 (89.4)	68 (10.6)	642

**Рис. 1.** Уровень потребления алкоголя в зависимости от генотипа. Серые столбики – носители генотипа *ADH1B\*Arg/Arg*, белые – *ADH1B\*Arg/His*.**Рис. 2.** Уровень CDT в зависимости от генотипа по *ADH1B* в подгруппах с разным уровнем потребления алкоголя. Серые столбики – носители генотипа *ADH1B\*Arg/Arg*, белые – *ADH1B\*Arg/His*.

*ADH1B\*48His* в подгруппе трезвенников имеют незначимо более высокий уровень CDT, чем обладатели генотипа *Arg/Arg* – 0.70 и 0.61 соответственно. В подгруппе потребляющих менее 8000 г в год носители аллеля *ADH1B\*48His* потребляют в среднем на 20% алкоголя меньше (2741 г в год), чем не имеющие этого аллеля мужчины (3398 г в год). Эти различия в нашей выборке статистически незначимы. Согласно опубликованным данным по 56 исследованиям носители аллеля *ADH1B\*48His* потребляют в среднем на 17.2% алкоголя меньше (95%ДИ 15.6–18.9%), чем не имеющие этого аллеля [17]. Поэтому можно полагать, что различия не случайны, а отражают общую тенденцию, не достигая статистической значимости из-за недостаточного размера выборки. В нашей выборке эти различия не были значимыми из-за недостаточной мощности теста (для мощности 80% необходим общий объем выборки 2880 человек). Однако средний уровень CDT (1.60 и 1.59%) одинаков у носителей обоих генотипов.

В подгруппе с высоким уровнем потребления алкоголя носители аллеля *ADH1B\*48His* потребляют в среднем на 12% меньше алкоголя (15452 г в год), чем те, у кого этого аллеля нет (17557 г в год). Однако уровень CDT у носителей аллеля *ADH1B\*48His* и в этой подгруппе почти в два раза выше – 4.88%, по сравнению с 2.51% у не имеющих данного аллеля ( $p = 0.003$  по критерию Манна – Уитни). Полученные данные указывают, что у носителей данного аллеля повышен уровень CDT, причем тем больше, чем выше уровень потребления алкоголя. Аллель *ADH1B\*48His* не является редким – доля его носителей составляет среди русских 10–12%, в населении Кавказа – 20–25%, а в Южной Сибири (алтайцы, тувинцы, буряты) – до 40–50% [14]. Поэтому повышенный уровень CDT у носителей аллеля *ADH1B\*48His* необходимо учитывать в экспертной и клинической практике.

Таким образом, в проведенном нами исследовании впервые показано, что уровень используемого в клинической и судебно-медицинской практике биомаркера злоупотребления алкоголем CDT у мужчин-носителей аллеля *ADH1B\*48His* гена алкогольдегидрогеназы повышен по сравнению с теми, кто не имеет этого аллеля. Повышение тем больше, чем выше доза потребляемого алкоголя.

Можно предполагать, что эффект опосредован нарушением гликозилирования из-за более высокой концентрации ацетальдегида, образующегося у носителей аллеля *ADH1B\*48His* после приема алкоголя. В результате при определении уровня CDT у носителей данного аллеля, не злоупотребляющих алкоголем, могут быть получены ложнопозитивные результаты и ошибочно диагностировано злоупотребление. Этот эффект необходимо учитывать не только в клинической практике и судебно-медицинской экспертизе, но и в исследованиях, в которых влияние потребления алкоголя на здоровье и поведение изучается с применением менделевской рандомизации по полиморфному локусу rs1229984 *ADH1B\*48His*.

Метод менделевской рандомизации используют в обсервационных эпидемиологических исследованиях для формирования рандомизированных выборок, отличающихся только по изучаемому параметру [18, 19]. В качестве инструмента рандомизации используется выделение подгрупп по генетическим вариантам, тесно связанным с изучаемым фактором. Проблемой таких исследований является множество конфаундеров, связанных с образом жизни и иными параметрами, которые приводят к тому, что рандомизированные подгруппы, например, с разным потреблением алкоголя отличаются не только уровнем потребления, но и социально-экономическим статусом и многими другими характеристиками, что затрудняет выявление причинно-следственных связей.

Для исследования влияния алкоголя на здоровье и поведение в качестве инструмента рандомизации используют варианты гена алкогольдегидрогеназы, так как надежно показано сниженное потребление алкоголя у носителей аллеля *ADH1B\*48His* [17, 20]. Эффект влияния генотипа по *ADH1B* на уровень CDT при высоком уровне потребления алкоголя обнаружен нами впервые и требует подтверждения на независимо сформированных выборках большего объема с разным уровнем потребления алкоголя. Если бы аллель *ADH1B\*48His* воздействовал на уровень CDT только путем снижения количества потребляемого алкоголя, то у носителей этого аллеля уровень CDT был бы ниже, чем у не имеющих его индивидов. Но нарушение гликозилирования под действием усиленного образования ацетальдегида после приема этанола дает более сильный эффект, чем снижение уровня потребления алкоголя, в результате итоговый эффект оказывается противоположным; и уровень CDT у носителей аллеля *ADH1B\*48His* повышен. Нарушение гликозилирования под действием ацетальдегида касается не только трансферрина, но и иных гликопротеинов, например, аполипопротеина E [11], и может затрагивать различные процессы. Обнаруженный нами эффект повышения уровня CDT у носителей аллеля *ADH1B\*48His* имеет

фундаментальное значение для исследований влияния алкоголя на здоровье и другие фенотипические признаки, а также важен для применения в клинической и судебно-медицинской практике при выявлении лиц, злоупотребляющих алкоголем.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (соглашение № 122022600161-3).

Исследование одобрено Этическим комитетом Института общей генетики им. Н.И. Вавилова РАН 15.07.2024, протокол № 23.

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и 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. Все обследованные – совершеннолетние.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Андреев Е.М., Кирьянов Н.А., Леон Д. и др. Злоупотребление алкоголем и преждевременная смертность в России на примере Ижевска // Наркология. 2008. № 7. С. 38–52.
2. Leon D.A., Saburova L., Tomkins S. et al. Hazardous alcohol drinking and premature mortality in Russia: A population based case-control study // Lancet. 2007. V. 369. № 9578. P. 2001–2009. [https://doi.org/10.1016/S0140-6736\(07\)60941-6](https://doi.org/10.1016/S0140-6736(07)60941-6)
3. Midanik L.T. Validity of self-reported alcohol use: A literature review and assessment // Br. J. Addict. 1988. V. 83. № 9. P. 1019–1030. <https://doi.org/10.1111/j.1360-0443.1988.tb00526.x>
4. Немцов А.В., Андриенко Ю.В. Самоотчеты населения России о потреблении алкоголя // Наркология. 2007. № 5. С. 58–61.
5. Jones A.W. Brief history of the alcohol biomarkers CDT, EtG, EtS, 5-HTOL, and Peth // Drug Test Anal. 2024. V. 16. № 6. P. 570–587. <https://doi.org/10.1002/dta.3584>
6. Брюн Е.А., Мягкова М.А., Сокольчик Е.И. и др. Методика диагностического тестирования на предмет хронического злоупотребления алкоголем. Методические рекомендации. № 16. М.: 2011. 40 с.
7. Stibler H., Kjellin K.G. Isoelectric focusing and electrophoresis of the CSF proteins in tremor of different origins // J. Neurol. Sci. 1976. V. 30. P. 269–285. [https://doi.org/10.1016/0022-510x\(76\)90133-7](https://doi.org/10.1016/0022-510x(76)90133-7)

8. Bergström J.P., Helander A. Influence of alcohol use, ethnicity, age, gender, BMI, and smoking on the serum transferrin glycoform pattern: Implications for use of carbohydrate-deficient transferrin (CDT) as alcohol biomarker // Clin. Chim. Acta. 2008. V. 388. № 1–2. P. 59–67.  
<https://doi.org/10.1016/j.cca.2007.10.011>
9. Stibler H. Carbohydrate-deficient transferrin in serum: A new marker of potentially harmful alcohol consumption reviewed // Clin. Chem. 1991. V. 37. № 12. P. 2029–2037.
10. Helander A., Wielders J., Anton R. et al. International Federation of Clinical Chemistry and Laboratory Medicine Working Group on Standardisation of Carbohydrate-Deficient Transferrin (IFCC WG-CDT). Reprint of standardisation and use of the alcohol biomarker carbohydrate-deficient transferrin (CDT) // Clin. Chim. Acta. 2017. V. 467. P. 15–20.  
<https://doi.org/10.1016/j.cca.2017.03.018>
11. Lakshman M.R., Rao M.N., Marmillot P. Alcohol and molecular regulation of protein glycosylation and function // Alcohol. 1999. V. 19. № 3. P. 239–247.  
[https://doi.org/10.1016/s0741-8329\(99\)00041-5](https://doi.org/10.1016/s0741-8329(99)00041-5)
12. Flahaut C., Michalski J.C., Danel T. et al. The effects of ethanol on the glycosylation of human transferrin // Glycobiology. 2003. V. 13. № 3. P. 191–198.  
<https://doi.org/10.1093/glycob/cwg016>
13. Bortolotti F., Sorio D., Bertaso A., Tagliaro F. Analytical and diagnostic aspects of carbohydrate deficient transferrin (CDT): A critical review over years 2007–2017 // J. Pharmaceutical and Biomed. An. 2018. V. 147. P. 2–12.  
<https://doi.org/10.1016/j.jpba.2017.09.006>
14. Borinskaya S., Kal'ina N., Marusin A. et al. Distribution of alcohol dehydrogenase *ADH1B*\*47His allele in Eurasia // Am. J. Hum. Genet. 2009. V. 84. № 1. P. 89–92.  
<https://doi.org/10.1016/j.ajhg.2008.12.007>
15. McDonald H., Borinskaya S., Kiryanov N. et al. Comparative performance of biomarkers of alcohol consumption in a population sample of working-aged men in Russia: The Izhevsk Family Study // Addiction. 2013. V. 108. № 9. P. 1579–1589.  
<https://doi.org/10.1111/add.12251>
16. Боринская С.А., Ким А.А., Рубанович А.В., Янковский Н.К. Влияние аллелей гена *ADH1B* и уровня образования на характер потребления алкоголя у российских мужчин // Acta Naturae. 2013. Т. 5. № 3 (18). С. 103–110.
17. Holmes M.V., Dale C.E., Zuccolo L. et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data // BMJ. 2014. V. 349. P. 1–16.  
<https://doi.org/10.1136/bmj.g4164>
18. Sanderson E., Glymour M.M., Holmes M.V. et al. Mendelian randomization // Nat. Rev. Meth. Primers. 2022. V. 2. P. 6.  
doi: 10.1038/s43586-021-00092-5
19. Skrivankova V.W., Richmond R.C., Woolf B.A. et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): Explanation and elaboration. BMJ. 2021. V. 375.  
doi:10.1136/bmj.n2233
20. Jennings M.V., Martínez-Magaña J.J., Courchesne-Krak N.S. et al. A genome-wide association and Mendelian randomisation study of alcohol use variants in a diverse cohort comprising over 3 million individuals // EBioMedicine. 2024. V. 103.  
doi: 10.1016/j.ebiom.2024.105086

## Effect of Genotype at *ADH1B* Alcoholdehydrogenase Gene on Level of CDT Alcohol Abuse Marker

A. A. Kim<sup>1</sup>, A. S. Gureev<sup>1</sup>, A. V. Rubanovich<sup>1, \*</sup>, S. A. Borinskaya<sup>1, \*\*</sup>

<sup>1</sup>Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia

\*e-mail: rubanovich@vigg.ru

\*\*e-mail:borinskaya@vigg.ru

Alcohol abuse poses an important challenge to public health and is associated with a shorter lifespan, numerous disorders, and social and economic problems. Methods to identify heavy drinkers are necessary for choosing the prevention and treatment strategies in medicine and are in demand in forensics. Alcohol consumption is often underestimated in self-reported data, and objective laboratory tests are therefore essential to employ in diagnosing chronic alcohol abuse and acute alcoholic excess. Carbohydrate-deficient transferrin (CDT) is one of the most specific biomarkers of alcohol abuse. CDT is a set of transferrin isoforms with a lower content of sialic acid residues and is found when glycosylation is impaired by ethanol metabolites. Oxidation of exogenous ethanol to acetaldehyde by alcohol dehydrogenase 1B (ADH1B) is a major pathway of ethanol metabolism. This study showed for the first time that carriers of the allele *ADH1B\*48His* (rs1229984), which determines acetaldehyde production to higher concentrations, have greater CDT levels as compared with noncarriers, alcohol consumption being the same. The difference should be taken into account in medicine, forensics, and studies where Mendelian randomization with respect to the polymorphic locus rs1229984 is performed to address the effects of drinking alcohol.

**Keywords:** alcohol abuse biomarkers, carbohydrate-deficient transferrin, CDT, alcohol dehydrogenase, *ADH1B* gene.

ПЕРСОНАЛИИ

ПОЗДРАВЛЕНИЕ С 90-ЛЕТИЕМ АКАДЕМИКА РАН  
ШЕСТАКОВА СЕРГЕЯ ВАСИЛЬЕВИЧА



**Глубокоуважаемый Сергей Васильевич!**

Уже более полувека Ваша научная деятельность в нашей стране и в мире посвящена развитию генетики – молекулярной и радиационной, генетики и генетической инженерии микроорганизмов, биотехнологии. Ваши научные интересы охватывают широкий диапазон фундаментальных и прикладных аспектов академической науки – введение в генетические исследования новых цианобактериальных объектов, создание для них методов генетического анализа и генной инженерии, изучение закономерностей мутационного процесса, рекомбинации и репарации.

Научная общественность знает Вас не только как талантливого ученого, но и как успешного руководителя и организатора науки: Вашей рукой написан первый Устав Межрегиональной общественной организации «Вавиловское общество генетиков и селекционеров»; многие годы Вы заведовали

кафедрой генетики Биологического факультета МГУ им. М.В. Ломоносова, возглавляли Институт общей генетики им. Н.И. Вавилова РАН и научный журнал «Генетика». Вы являетесь автором более 300 научных работ, авторских свидетельств и патентов на изобретения в области промышленной микробиологии и биотехнологии. Под Вашим научным руководством защищены десятки кандидатских и докторских диссертаций.

Результаты Вашего многолетнего труда получили признание Правительства и научной общественности: Вы – лауреат Государственной премии СССР, заслуженный деятель науки Российской Федерации, заслуженный работник высшей школы Российской Федерации, заслуженный профессор МГУ им. М.В. Ломоносова, лауреат премии им. М.В. Ломоносова, награждены Золотой медалью им. Н.И. Вавилова РАН, медалью «За заслуги» Московского общества испытателей природы и именным серебряным нагрудным знаком МГУ им. М.В. Ломоносова «Naturaliae Eminentis».

Вы внесли колossalный вклад в развитие генетической науки и образования как выдающийся ученый, руководитель, педагог, и как человек, глубоко преданный науке и нашей Родине.

**Дорогой Сергей Васильевич!**

Поздравляем Вас с 90-летним юбилеем и в этот замечательный день от всей души желаем здоровья, благополучия, дальнейших творческих свершений и радости от них!

С уважением  
от коллектива Биологического факультета МГУ им. М.В. Ломоносова,  
коллектива Института общей генетики им. Н.И. Вавилова РАН,  
коллектива Института цитологии и генетики СО РАН,  
членов Вавиловского общества генетиков и селекционеров.

**Поправки:**

К статье «ТЕСТИРОВАНИЕ МИКРОСАТЕЛЛИТНЫХ ЛОКУСОВ КРЫЖОВНИКА НА ЧЁРНОЙ И КРАСНОЙ СМОРОДИНЕ»

*А.В. Пикунова, А.А. Павленко, М.А. Должикова, О.Д. Голяева, С.Д. Князев*

ж. «Генетика», 2024. Т. 60. № 10. С. 117–121.

На стр. 120, правая колонка, следует читать: Работа выполнена при поддержке гранта РНФ «Секвенирование и аннотация ядерного и хлоропластного генома красной смородины (*Ribes rubrum* L.)».

К статье «ГЕНОМНЫЙ АНАЛИЗ ОКРАСКИ МЕХА СОБОЛЯ (*Martes zibellina* L.), ПОИСК МУТАЦИЙ, ОПРЕДЕЛЯЮЩИХ ОТСУТСТВИЕ ПИГМЕНТАЦИИ (АЛЬБИНОС)»

ж. «Генетика», 2024. Т. 60. № 11. С. 112–117.

П.А. Филимонов, А.Д. Манахов, М.И. Митина, А.А. Онохов, И.Е. Чернова,  
Л.В. Максимова, С.Н. Каштанов, Е.И. Рогаев

На стр. 115, левая колонка, следует читать: Исследования проводились с привлечением средств Фонда РНФ, грант № 23-26-00233 от 13.01.2023 г.